

Física Teórica 3 – segundo cuatrimestre de 2023

Guía 3: proceso binario en tiempo discreto

■ **Proceso binario I.** El estado de un proceso de Markov en pasos discretos puede ser **a** o **b**. Las probabilidades de transición en pasos consecutivos, $p(i, n|j, n-1)$, entre uno y otro estado están dadas esquemáticamente por

$$\begin{bmatrix} p(a \rightarrow a) & p(a \rightarrow b) \\ p(b \rightarrow a) & p(b \rightarrow b) \end{bmatrix} = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}. \quad (1)$$

Por simplicidad se asume que α y β no son ni 0 ni 1.

a) Encuentre las ecuaciones que dan las probabilidades $p_a(n+1)$ y $p_b(n+1)$ luego de $n+1$ pasos en términos de las probabilidades $p_a(n)$ y $p_b(n)$ del paso anterior.

Como hay sólo dos estados, p_b puede eliminarse del problema, ya que $p_b = 1 - p_a$. Entonces, en todo lo que sigue es suficiente con escribir y resolver las ecuaciones para p_a .

b) Encuentre la distribución estacionaria, es decir, aquella para la cual

$$p_a(n+1) = p_a(n) \equiv P_a. \quad (2)$$

c) Suponga que la probabilidad se parametriza del siguiente modo:

$$p_a(n) = P_a + \Delta(n). \quad (3)$$

Encuentre y resuelva la ecuación de evolución para $\Delta(n)$ a partir de una condición inicial $\Delta(0)$ en $t_0 = 0$. Estudie la convergencia a la distribución estacionaria. ¿Qué sucede si $\alpha = \beta = \frac{1}{2}$?

d) Calcule las cuatro probabilidades de transición de n pasos, $p(i, k+n|j, k)$. *Sugerencia:* puesto que las probabilidades de transición no dependen del tiempo,

$$p(i, k+n|j, k) = p(i, n|j, 0). \quad (4)$$

Estas cuatro probabilidades condicionales se obtienen fijando de manera adecuada las condiciones iniciales en los resultados de los ítems anteriores.

e) Como aplicación del problema 1, suponga que $\alpha = 1/8$ y $\beta = 1/2$. El sistema parte del estado **a**, y es observado luego de diez pasos en el mismo estado **a**. Dada esta información, calcule exactamente las probabilidades de que el sistema haya pasado por uno u otro estado en el quinto paso.

f) Grafique $p(i, k|a, 0; a, n)$, con $0 \leq k \leq n$, como función de k , para valores fijos de n . ¿Qué pasa a medida que k se aleja de 0 o de n ?

■ **Solución.** Queremos encontrar las probabilidades $p_a(n+1)$ y $p_b(n+1)$ en términos de las probabilidades del paso anterior. Hay dos maneras en las que el proceso puede estar en el estado a a tiempo $n+1$: i) si estaba en el estado a a tiempo n y permanece en ese estado, o ii) si estaba en el estado b a tiempo n y realiza una transición. Formalmente,

$$p_a(n+1) = p_a(n)p(a \rightarrow a) + p_b(n)p(b \rightarrow a) = (1 - \alpha)p_a(n) + \beta p_b(n). \quad (5)$$

De manera análoga,

$$p_b(n+1) = p_a(n)p(a \rightarrow b) + p_b(n)p(b \rightarrow b) = \alpha p_a(n) + (1 - \beta)p_b(n). \quad (6)$$

Puesto que hay sólo dos estados, podemos eliminar p_b escribiendo

$$p_b(k) = 1 - p_a(k). \quad (7)$$

Así, la Ec. (5) se lee como

$$p_a(n+1) = (1 - \alpha)p_a(n) + \beta[1 - p_a(n)] = \beta + (1 - \alpha - \beta)p_a(n). \quad (8)$$

El sistema está en el equilibrio luego de n pasos si $p_a(n+1) = p_a(n)$. Para $k \geq n$, $p_a(k) = P_a$, donde P_a satisface la ecuación

$$P_a = \beta + (1 - \alpha - \beta)P_a. \quad (9)$$

De aquí obtenemos

$$P_a = \frac{\beta}{\alpha + \beta}. \quad (10)$$

Si $p_a = P_a$, entonces p_b debe ser igual a

$$P_b = 1 - P_a = \frac{\alpha}{\alpha + \beta}. \quad (11)$$

Para resolver la Ec. (8), resulta práctico escribir $p_a(n)$ en la siguiente forma:

$$p_a(n) = P_a + \Delta(n). \quad (12)$$

Reemplazando en la Ec. (8), resulta una ecuación para Δ ,

$$\Delta(n+1) = \beta - (\alpha + \beta)P_a + (1 - \alpha - \beta)\Delta(n) = (1 - \alpha - \beta)\Delta(n). \quad (13)$$

A partir de la condición inicial $\Delta(0)$ es sencillo propagar la solución para todo $n \geq 0$,

$$\Delta(n) = \gamma^n \Delta(0), \quad (14)$$

donde hemos definido

$$\gamma = 1 - \alpha - \beta. \quad (15)$$

De acuerdo a las hipótesis del problema,

$$-1 < \gamma < 1. \quad (16)$$

Si $\alpha + \beta = 1$, por ejemplo, si $\alpha = \beta = \frac{1}{2}$, entonces $\gamma = 0$ y para $n \geq 1$

$$\Delta(n) = 0. \quad (17)$$

Es decir, si $\gamma = 0$, el sistema alcanza el equilibrio en un solo paso. Si $\gamma \neq 0$,

$$\lim_{n \rightarrow \infty} \Delta(n) = 0. \quad (18)$$

El sistema tiende al equilibrio cuando $n \rightarrow \infty$ tanto más rápido cuanto menor es $|\gamma|$.

Para dejar escrita la Ec. (14) en términos de $p_a(0)$, notemos que

$$p_a(0) = P_a + \Delta(0), \quad (19)$$

de modo que

$$\Delta(n) = \gamma^n [p_a(0) - P_a]. \quad (20)$$

Luego, sumando P_a a ambos lados de la ecuación,

$$p_a(n) = (1 - \gamma^n) P_a + \gamma^n p_a(0). \quad (21)$$

Los mismos argumentos de antes muestran que el sistema tiende al equilibrio, independientemente de la condición inicial. Además,

$$p_b(n) = (1 - \gamma^n) P_b + \gamma^n p_b(0). \quad (22)$$

Una cuestión interesante es calcular las probabilidades de transición de n pasos, es decir, la probabilidad de que el estado vaya de i a j transcurridos n pasos a partir de cualquier instante inicial k ,

$$p(i, k + n | j, k). \quad (23)$$

Como el sistema tiene simetría de traslación en el tiempo (esto es, las probabilidades de transición de un paso no dependen de n), el valor de k es irrelevante, de manera que alcanza con calcular

$$p(i, n | j, 0). \quad (24)$$

Esto no requiere hacer ningún cálculo nuevo. Sabemos cómo calcular la probabilidad de que el sistema esté en un dado estado para cualquier condición inicial sobre las probabilidades. El hecho de que el sistema esté en el estado j a tiempo $t = 0$ significa que

$$p_j(0) = 1. \quad (25)$$

Entonces, lo que debemos hacer es reescribir las Ecs. (21) y (22) especificando las condiciones iniciales de esa manera:

$$\begin{aligned}
 p(a, n|a, 0) &= (1 - \gamma^n) P_a + \gamma^n = P_a + \gamma^n P_b, \\
 p(a, n|b, 0) &= (1 - \gamma^n) P_a, \\
 p(b, n|a, 0) &= (1 - \gamma^n) P_b, \\
 p(b, n|b, 0) &= (1 - \gamma^n) P_b + \gamma^n = P_b + \gamma^n P_a.
 \end{aligned}
 \tag{26}$$

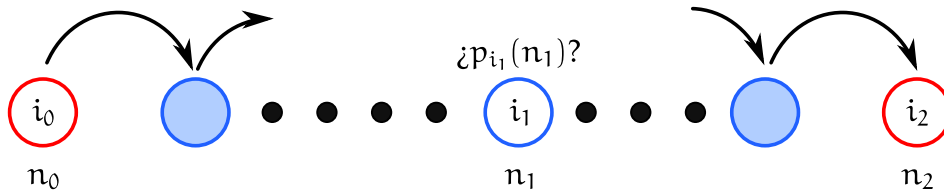
Podemos escribir esto más sintéticamente como

$$p(i, n|j, 0) = (1 - \gamma^n) P_i + \delta_{ij} \gamma^n. \tag{27}$$

Las últimas preguntas que propone el problema son acerca de la probabilidad de que el proceso haya pasado por el estado i_1 a tiempo n_1 sabiendo que a tiempo n_2 se lo ha observado en el estado i_2 , y a tiempo n_0 en el estado i_0 , con $n_0 \leq n_1 \leq n_2$. Es decir, cuál es la probabilidad condicional

$$p(i_1, n_1 | i_2, n_2; i_0, n_0). \tag{28}$$

En otras palabras: sabemos con certeza el estado actual del proceso y el estado inicial. Queremos averiguar cuál es la probabilidad de que haya pasado por el estado i_1 en un paso intermedio a tiempo n_1 .



Para desenredar esta cuestión, llamemos a los eventos involucrados con las letras A, B y C, ordenados temporalmente. Por ejemplo, $A = (i_0, n_0)$. Lo que queremos averiguar es

$$p(B|A, C). \tag{29}$$

Hay que tratar de expresar esto en términos de las probabilidades de transición directas de dos estados. Empecemos por escribir la probabilidad conjunta de los tres eventos y usemos la definición de la probabilidad condicional para hacer aparecer la probabilidad que queremos calcular:

$$p(A, B, C) = p(B|A, C) p(A, C). \tag{30}$$

Por lo tanto, la probabilidad buscada es

$$p(B|A, C) = \frac{p(A, B, C)}{p(A, C)} = \frac{p(A, B, C)}{p(C|A) p(A)}. \quad (31)$$

El denominador puede escribirse como

$$p(A, B, C) = p(C|A, B) p(A, B). \quad (32)$$

La propiedad de Markov implica $P(C|A, B) = P(C|B)$. Entonces,

$$p(A, B, C) = p(C|B) p(A, B) = p(C|B) p(B|A) p(A). \quad (33)$$

Volviendo a la Ec. (31), resulta

$$p(B|A, C) = \frac{p(C|B) p(B|A)}{p(C|A)}. \quad (34)$$

Explícitamente, para el proceso binario es

$$p(i_1, n_1 | i_2, n_2; i_0, n_0) = \frac{p(i_2, n_2 | i_1, n_1) p(i_1, n_1 | i_0, n_0)}{p(i_2, n_2 | i_0, n_0)}. \quad (35)$$

Usando la simetría ante las traslaciones temporales,

$$p(i_1, n_1 | i_2, n_2; i_0, n_0) = \frac{p(i_2, n_2 - n_1 | i_1, 0) p(i_1, n_1 - n_0 | i_0, 0)}{p(i_2, n_2 - n_0 | i_0, 0)}. \quad (36)$$

Por ejemplo, la probabilidad de haber pasado por a en n_1 si se sabe que el proceso partió de a en tiempo $n_0 = 0$ y estuvo en a a tiempo n_2 es

$$p(a, n_1 | a, n_2; a, 0) = \frac{p(a, n_2 - n_1 | a, 0) p(a, n_1 | a, 0)}{p(a, n_2 | a, 0)} \quad (37)$$

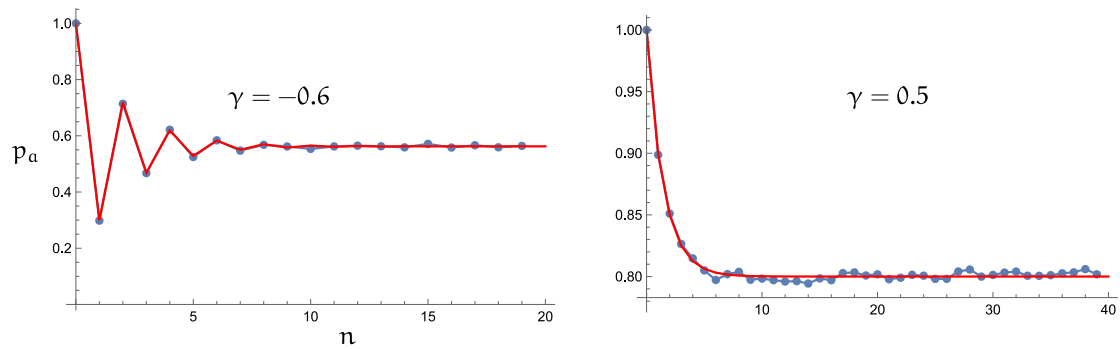
Usando la Ec. (26), resulta

$$p(a, n_1 | a, n_2; a, 0) = \frac{(P_a + \gamma^{n_2 - n_1} P_b) (P_a + \gamma^{n_1} P_b)}{P_a + \gamma^{n_2} P_b}. \quad (38)$$

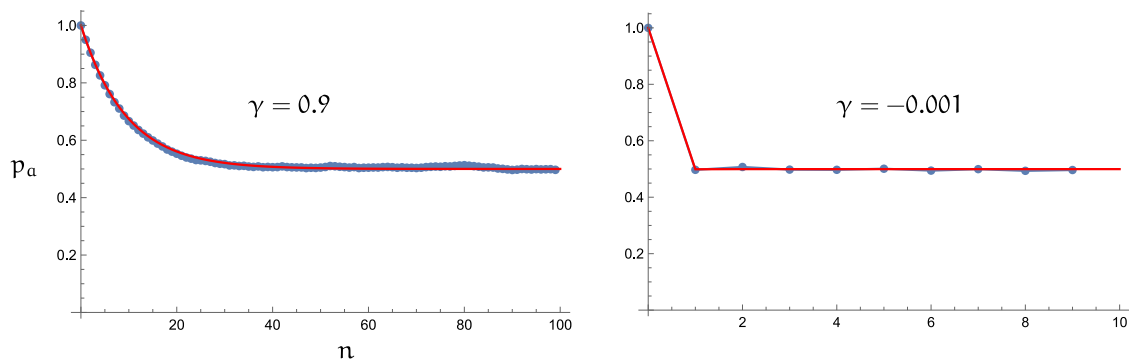
La probabilidad $p(b, n_1 | b, n_2; b, 0)$ se obtiene intercambiando P_a con P_b . Con una sustitución similar, el resto de los casos puede obtenerse a partir de los siguientes tres:

$$\begin{aligned} p(b, n_1 | a, n_2; a, 0) &= \frac{(1 - \gamma^{n_2 - n_1}) (1 - \gamma^{n_1}) P_a P_b}{P_a + \gamma^{n_2} P_b}, \\ p(a, n_1 | a, n_2; b, 0) &= \frac{(P_a + \gamma^{n_2 - n_1} P_b) (1 - \gamma^{n_1})}{(1 - \gamma^{n_2})}, \\ p(b, n_1 | a, n_2; b, 0) &= \frac{(1 - \gamma^{n_2 - n_1}) (P_b + \gamma^{n_1} P_a)}{(1 - \gamma^{n_2})}. \end{aligned} \quad (39)$$

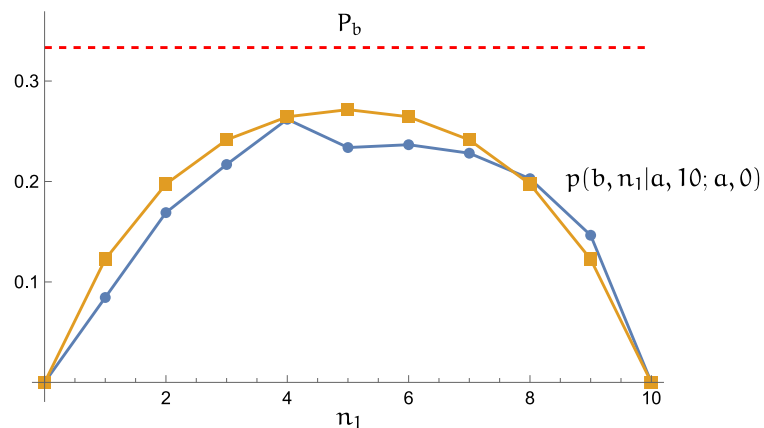
En la página siguiente hay algunos gráficos.



En estas figuras, el proceso parte con certeza del estado α . Si γ es negativo, la probabilidad tiende a la probabilidad de equilibrio P_α , pero oscila por encima y por debajo de P_α . Los puntos indican el resultado de simulaciones numéricas. Para estimar las probabilidades, se tomó una muestra de 10^4 cadenas.



Este par de figuras ilustra cómo el valor de γ fija la escala de tiempo de relajación. Un valor de $|\gamma|$ cercano a uno implica un tiempo de relajación grande. En la figura de la derecha, γ es muy próximo a cero. Luego de un solo paso, el proceso está prácticamente en el estado de equilibrio.



En esta figura se muestra la probabilidad de pasar por el estado \mathbf{b} a tiempo n_1 si se parte de del estado \mathbf{a} a tiempo $n_0 = 0$ y se llega a \mathbf{a} a tiempo $n_2 = 10$. En amarillo, los puntos teóricos. En azul, la estimación de la probabilidad hecha en base a una muestra de sólo 500 cadenas. Se tomó un número bajo de cadenas para mostrar la divergencia entre el resultado teórico y el del experimento numérico. Notar que a medida que el punto intermedio se aleja de los extremos, la probabilidad de encontrar al proceso en el estado \mathbf{b} se acerca a su valor en el equilibrio.

■ **Proceso binario II.** El problema anterior puede resolverse siguiendo un método más general que no depende de que el sistema tenga únicamente dos estados. Las probabilidades de transición en pasos consecutivos para un proceso discreto con N estados pueden representarse mediante una matriz \mathbf{M} de $N \times N$, con elementos

$$M_{ji} = p(i, n+1 | j, n). \quad (40)$$

En general, \mathbf{M} podría depender de n ; supondremos que no. Para el proceso binario,

$$\mathbf{M} = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}. \quad (41)$$

a) Demuestre que si $\mathbf{p}(n)$ es el vector cuyas componentes dan la probabilidad de cada estado luego de n pasos, entonces $\mathbf{p}(n+1) = \mathbf{p}(n) \cdot \mathbf{M}$ y, por lo tanto,

$$\mathbf{p}(n) = \mathbf{p}(0) \cdot \mathbf{M}^n. \quad (42)$$

¿Qué ecuación satisfacen las distribuciones estacionarias?

Como \mathbf{M} no es necesariamente simétrica, sus autovectores no tienen por que ser ortogonales. Sin embargo, puede demostrarse que los autovectores por izquierda y derecha, \mathbf{v}_r y \mathbf{w}_r , definidos respectivamente por $\mathbf{v}_r \cdot \mathbf{M} = \lambda_r \mathbf{v}_r$ y $\mathbf{M} \cdot \mathbf{w}_r = \lambda_r \mathbf{w}_r$, sí lo son; es decir $\mathbf{v}_r \cdot \mathbf{w}_s \propto \delta_{rs}$. (Notar que los dos conjuntos de autovectores comparten el mismo conjunto de autovalores. ¿Por qué?)

- b) Encuentre los autovectores y autovalores para la matriz (41) del proceso binario y verifique la propiedad de ortogonalidad $\mathbf{v}_r \cdot \mathbf{w}_s \propto \delta_{rs}$.
- c) Usando la condición de suma de probabilidades igual a 1, muestre que $\lambda = 1$ es un autovalor, para cualquier matriz de transición, y que un autovector por derecha es $\mathbf{w} \propto (1, 1, \dots, 1)$. De los N autovectores, puede elegirse que éste sea el primero. Tomaremos $\mathbf{w}_1 = (1, 1, \dots, 1)$.
- d) El autovector \mathbf{v}_1 asociado a \mathbf{w}_1 se normaliza de manera que sus componentes sumen 1. Verificar entonces que $\mathbf{w}_1 \cdot \mathbf{v}_1 = 1$. El resto de los autovectores puede normalizarse de modo que $\mathbf{w}_i \cdot \mathbf{v}_j = \delta_{ij}$, en la forma que resulte más conveniente. En los ítems siguientes se asume esa normalización.
- e) La propiedad de ortogonalidad entre los dos conjuntos de autovectores asegura que cada conjunto $\{\mathbf{v}_r\}$ y $\{\mathbf{w}_r\}$ es linealmente independiente. Asumiendo que es posible encontrar N autovectores \mathbf{w} distintos, demuestre que cualquier vector \mathbf{u} en N dimensiones puede escribirse como

$$\mathbf{u} = \sum_{r=1}^N (\mathbf{u} \cdot \mathbf{w}_r) \mathbf{v}_r, \quad (43)$$

y que, entonces,

$$\mathbf{p}(n) = \sum_{r=1}^N \mathbf{p}(0) \cdot \mathbf{w}_r \lambda_r^n \mathbf{v}_r \quad \text{y} \quad (\mathbf{M}^n)_{ij} = \sum_{r=1}^N \lambda_r^n (\mathbf{w}_r)_i (\mathbf{v}_r)_j. \quad (44)$$

El par de ecuaciones anteriores dan una manera práctica de calcular $\mathbf{p}(n)$ y \mathbf{M}^n .

f) Por fin, como aplicación, escriba los autovectores normalizados para el proceso binario con la matriz de transición (41) y verifique que la ecuación (44) reproduce para $\mathbf{p}(n) = (p_a(n), p_b(n))$ los resultados obtenidos siguiendo el método más elemental del problema anterior.

■ **Solución.** En general, si un proceso puede asumir N estados, numerados del 1 al N , la probabilidad a tiempo $n + 1$ de estar en el estado i es

$$p_i(n + 1) = \sum_{j=1}^N p(i, n + 1|j, n) p_j(n). \quad (45)$$

Asumiendo que las probabilidades de transición de un paso son independientes de n , podemos definir la matriz \mathbf{M} tal que

$$(\mathbf{M})_{ji} = p(i, n + 1|j, n) = p(i, 1|j, 0). \quad (46)$$

La Ec. (45) se reescribe como

$$p_i(n + 1) = \sum_{j=1}^N p_j(n) M_{ji}. \quad (47)$$

Introduciendo el vector de probabilidades $\mathbf{p}(n)$, con componentes

$$[\mathbf{p}(n)]_i = p_i(n), \quad (48)$$

la Ec. (47) puede quedar expresada en forma matricial,

$$\mathbf{p}(n + 1) = \mathbf{p}(n) \cdot \mathbf{M}. \quad (49)$$

Iterando a partir de la condición inicial $\mathbf{p}(0)$, queda

$$\mathbf{p}(n) = \mathbf{p}(0) \cdot \mathbf{M}^n. \quad (50)$$

Una distribución estacionaria es aquella que satisface la ecuación

$$\mathbf{P} = \mathbf{P} \cdot \mathbf{M}. \quad (51)$$

Las características que distinguen a una matriz de transición son dos: primero, que todos sus elementos son no negativos, porque representan probabilidades; y, segundo, que

la suma de los elementos de cada fila es igual a uno. En efecto,

$$\sum_{i=1}^N M_{ji} = \sum_{i=1}^N p(i, n+1|j, n) = 1. \quad (52)$$

Una matriz con estas propiedades recibe el nombre de matriz estocástica.

El problema probabilístico se reduce al problema algebraico de calcular \mathbf{M}^n . Estamos habituados al caso de las matrices simétricas. Para estas matrices es posible encontrar un conjunto de N autovectores ortogonales, tales que

$$\begin{aligned} \mathbf{M} \cdot \mathbf{v}_k &= \lambda_k \mathbf{v}_k, \\ \mathbf{v}_j \cdot \mathbf{v}_k &\propto \delta_{jk}. \end{aligned} \quad (53)$$

Para definir la acción de un operador lineal es suficiente especificar su acción sobre una base del espacio de vectores. El conjunto de vectores \mathbf{v}_k forma una base del espacio. Es fácil ver que el operador definido por

$$\mathbf{A} = \sum_{k=1}^N \frac{\lambda_k}{\mathbf{v}_k \cdot \mathbf{v}_k} \mathbf{v}_k \mathbf{v}_k \quad (54)$$

tiene la siguiente propiedad:

$$\mathbf{A} \cdot \mathbf{v}_j = \lambda_j \mathbf{v}_j. \quad (55)$$

De manera que $\mathbf{A} = \mathbf{M}$, porque tienen la misma acción sobre una base de vectores. Veremos que escribir \mathbf{M} como en la Ec. (54) vuelve trivial el cálculo de \mathbf{M}^n .

En la Ec. (54) hemos introducido notación de diadas. Suelen aparecer también en mecánica clásica y en electromagnetismo. En mecánica clásica, para escribir el tensor de inercia; en electromagnetismo, para escribir el tensor de Maxwell. Una diada se simboliza con la yuxtaposición de dos vectores, por ejemplo,

$$\mathbf{D} = \mathbf{a}\mathbf{b}. \quad (56)$$

Una diada es un operador lineal. Su acción sobre los vectores está definida como

$$\begin{aligned} \mathbf{v} \cdot \mathbf{D} &= (\mathbf{v} \cdot \mathbf{a})\mathbf{b}, \\ \mathbf{D} \cdot \mathbf{v} &= \mathbf{a}(\mathbf{b} \cdot \mathbf{v}). \end{aligned} \quad (57)$$

Las diadas dan una manera práctica de escribir los operadores lineales sin necesidad de recurrir a su definición por componentes. Solemos decir que el símbolo \mathbf{M} representa una matriz. Sería mejor decir que \mathbf{M} es un operador, cuya acción, habiendo elegido una base del espacio vectorial, puede definirse en términos de las componentes M_{ij} de una matriz.

Para no multiplicar la notación, lo habitual es usar el mismo símbolo para el operador y para la matriz.

A partir de la descomposición (54), es fácil calcular las potencias de \mathbf{M} . Por ejemplo,

$$\mathbf{M}^2 = \mathbf{M} \cdot \mathbf{M} = \sum_{k=1}^N \sum_{j=1}^N \frac{\lambda_k}{\mathbf{v}_k \cdot \mathbf{v}_k} \frac{\lambda_j}{\mathbf{v}_j \cdot \mathbf{v}_j} \mathbf{v}_k (\mathbf{v}_k \cdot \mathbf{v}_j) \mathbf{v}_j = \sum_{k=1}^N \frac{\lambda_k^2}{\mathbf{v}_k \cdot \mathbf{v}_k} \mathbf{v}_k \mathbf{v}_k. \quad (58)$$

En general,

$$\mathbf{M}^n = \sum_{k=1}^N \frac{\lambda_k^n}{\mathbf{v}_k \cdot \mathbf{v}_k} \mathbf{v}_k \mathbf{v}_k. \quad (59)$$

Esto da una manera muy eficiente de calcular \mathbf{M}^n , sin necesidad de diagonalizar explícitamente.

Tratándose de procesos de Markov, la matriz de transición \mathbf{M} no tiene por qué ser simétrica. Sin embargo, se puede calcular \mathbf{M}^n recurriendo a un método similar al anterior. Lo que puede demostrarse es que siempre es posible encontrar dos conjuntos de autovectores, por izquierda y por derecha, respectivamente,

$$\mathbf{v}_r \cdot \mathbf{M} = \lambda_r \mathbf{v}_r, \quad \mathbf{M} \cdot \mathbf{w}_r = \lambda_r \mathbf{w}_r. \quad (60)$$

Los autovectores \mathbf{v}_r no son ortogonales entre sí, y tampoco lo son los autovectores \mathbf{w}_r entre sí. Lo que vale ahora es la ortogonalidad entre ambos conjuntos de autovectores,

$$\mathbf{v}_r \cdot \mathbf{w}_s \propto \delta_{rs}. \quad (61)$$

Esto es todo lo que necesitamos para poder calcular las potencias de \mathbf{M} . La condición de ortogonalidad entre ambos conjuntos de autovectores implica que ambos conjuntos, considerados separadamente, son linealmente independientes (demostrarlo). Por lo tanto, la matriz \mathbf{M} queda completamente especificada si definimos su acción sobre los vectores \mathbf{v}_r o sobre los vectores \mathbf{w}_r , puesto que forman bases. Es fácil ver que

$$\mathbf{M} = \sum_{r=1}^N \frac{\lambda_r}{\mathbf{v}_r \cdot \mathbf{w}_r} \mathbf{w}_r \mathbf{v}_r. \quad (62)$$

En efecto, su acción sobre los vectores \mathbf{w}_k es

$$\left(\sum_{r=1}^N \frac{\lambda_r}{\mathbf{v}_r \cdot \mathbf{w}_r} \mathbf{w}_r \mathbf{v}_r \right) \cdot \mathbf{w}_k = \sum_{r=1}^N \frac{\lambda_r}{\mathbf{v}_r \cdot \mathbf{w}_r} \mathbf{w}_r (\mathbf{v}_r \cdot \mathbf{w}_k) = \lambda_k \mathbf{w}_k. \quad (63)$$

El análogo de la Ec. (59) en el caso de las matrices estocásticas es

$$\mathbf{M}^n = \sum_{r=1}^N \frac{\lambda_r^n}{\mathbf{v}_r \cdot \mathbf{w}_r} \mathbf{w}_r \mathbf{v}_r. \quad (64)$$

La propiedad (52),

$$\sum_{i=1}^N M_{ji} = 1, \quad (65)$$

inmediatamente implica que uno de los autovectores por derecha es

$$\mathbf{w}_1 = (1, 1, \dots, 1). \quad (66)$$

El autovalor asociado es

$$\lambda_1 = 1. \quad (67)$$

Es convencional normalizar el autovector \mathbf{v}_1 de modo que sus componentes sumen uno. De esa forma

$$\mathbf{v}_1 \cdot \mathbf{w}_1 = \sum_{j=1}^N (\mathbf{v}_1)_j = 1. \quad (68)$$

Conviene normalizar el resto de los autovectores de modo que

$$\mathbf{v}_r \cdot \mathbf{w}_s = \delta_{rs}. \quad (69)$$

No hay una manera única de hacer esto. En todo lo que sigue supondremos que vale la normalización anterior.

Si existe n_0 tal que \mathbf{M}^{n_0} tiene elementos estrictamente positivos, se dice que la matriz es *regular*. Para estas matrices, puede demostrarse que el autovalor $\lambda_1 = 1$ tiene multiplicidad igual a uno y que el autovector \mathbf{v}_1 puede elegirse con todos sus elementos positivos. Además, todos los otros autovalores tienen módulo estrictamente menor que uno. En tal caso, si separamos el primer término en la descomposición (64), obtenemos

$$\mathbf{M}^n = \mathbf{w}_1 \mathbf{v}_1 + \sum_{j=2}^N \lambda_j^n \mathbf{w}_j \mathbf{v}_j. \quad (70)$$

Al hacer tender n a infinito, el único término que sobrevive es el primero,

$$\lim_{n \rightarrow \infty} \mathbf{M}^n = \mathbf{w}_1 \mathbf{v}_1. \quad (71)$$

Cualquiera sea la condición inicial,

$$\lim_{n \rightarrow \infty} \mathbf{p}(n) = \lim_{n \rightarrow \infty} \mathbf{p}(0) \cdot \mathbf{M}^n = [\mathbf{p}(0) \cdot \mathbf{w}_1] \mathbf{v}_1. \quad (72)$$

Pero como $\mathbf{w}_1 = (1, 1, \dots, 1)$ y además $\mathbf{p}(0)$ es un vector de probabilidades

$$\mathbf{p}(0) \cdot \mathbf{w}_1 = \sum_{j=1}^N p_j(0) = 1. \quad (73)$$

En definitiva,

$$\lim_{n \rightarrow \infty} \mathbf{p}(n) = \mathbf{v}_1. \quad (74)$$

De esta forma identificamos al autovector \mathbf{v}_1 con la única distribución de probabilidades estacionaria.

Tomemos como ejemplo el proceso binario definido por la matriz

$$\mathbf{M} = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}. \quad (75)$$

Si α y β no son cero ni uno, la matriz es regular. El polinomio característico para el problema de autovalores y autovectores es

$$(1 - \alpha - \lambda)(1 - \beta - \lambda) - \alpha\beta = \lambda^2 - 2 \left(1 - \frac{\alpha + \beta}{2}\right) \lambda + 1 - \alpha - \beta = 0. \quad (76)$$

De aquí resultan los autovalores

$$\lambda = 1 - \frac{\alpha + \beta}{2} \pm \sqrt{\left(1 - \frac{\alpha + \beta}{2}\right)^2 - 1 + \alpha + \beta} = 1 - \frac{\alpha + \beta}{2} \pm \frac{\alpha + \beta}{2}. \quad (77)$$

Previsiblemente, uno de los autovalores es $\lambda_1 = 1$. El otro es

$$\lambda_2 = 1 - \alpha - \beta \equiv \gamma. \quad (78)$$

Ya sabemos que uno de los autovectores por derecha es $\mathbf{w}_1 = (1, 1)$. Las componentes del otro autovector están determinadas por la ecuación

$$\beta w_1 + \alpha w_2 = 0. \quad (79)$$

Podemos elegir

$$\mathbf{w}_2 = (\alpha, -\beta). \quad (80)$$

Los autovectores por izquierda de la matriz \mathbf{M} son los autovectores por derecha de la matriz transpuesta \mathbf{M}^t . La ecuación para las componentes del autovector \mathbf{v}_1 es

$$-\alpha v_1 + \beta v_2 = 0. \quad (81)$$

Con la condición de que la suma de sus componentes sea uno, queda

$$\mathbf{v}_1 = \frac{1}{\alpha + \beta} (\beta, \alpha). \quad (82)$$

Para el segundo autovector,

$$\beta v_1 + \alpha v_2 = 0. \quad (83)$$

Bajo la condición de normalización $\mathbf{v}_2 \cdot \mathbf{w}_2 = 1$, resulta

$$\mathbf{v}_2 = \frac{1}{\alpha + \beta}(1, -1). \quad (84)$$

Dada una distribución de probabilidad inicial, $\mathbf{p}(0)$, la distribución a tiempo n será

$$\mathbf{p}(n) = \mathbf{v}_1 + \gamma^n [\mathbf{p}(0) \cdot \mathbf{w}_2] \mathbf{v}_2. \quad (85)$$

De aquí podemos leer cada componente,

$$\begin{aligned} p_a(n) &= \frac{\beta}{\alpha + \beta} + \frac{\gamma^n}{\alpha + \beta} [\alpha p_a(0) - \beta p_b(0)], \\ p_b(n) &= \frac{\alpha}{\alpha + \beta} - \frac{\gamma^n}{\alpha + \beta} [\alpha p_a(0) - \beta p_b(0)]. \end{aligned} \quad (86)$$

Definiendo, como en el problema anterior,

$$\begin{aligned} P_a &= \frac{\beta}{\alpha + \beta}, \\ P_b &= \frac{\alpha}{\alpha + \beta}, \end{aligned} \quad (87)$$

tendremos

$$\begin{aligned} p_a(n) &= P_a + \gamma^n [P_b p_a(0) - P_a p_b(0)], \\ p_b(n) &= P_b - \gamma^n [P_b p_a(0) - P_a p_b(0)]. \end{aligned} \quad (88)$$

Evidentemente se verifica que

$$p_a(n) + p_b(n) = P_a + P_b = 1. \quad (89)$$

Usando estas propiedades, las Ecs. (88) pueden escribirse exactamente igual que en el problema anterior:

$$\begin{aligned} p_a(n) &= (1 - \gamma^n) P_a + \gamma^n p_a(0), \\ p_b(n) &= (1 - \gamma^n) P_b + \gamma^n p_b(0). \end{aligned} \quad (90)$$