

Lección 3: Procesos de Markov

En este capítulo vamos a describir un subconjunto de procesos estocásticos que tienen la propiedad de Markov. Estos procesos son los más importantes en física y química.

Definición

- Un proceso de Markov se define como un proceso estocástico con la propiedad de que para cualquier conjunto sucesivo de n tiempos $t_1 < \dots < t_n$ se tiene

$$P_{1|n-1}(y_n, t_n | y_1, t_1; \dots; y_{n-1}, t_{n-1}) = P_{1|1}(y_n, t_n | y_{n-1}, t_{n-1}) \quad (132)$$

Esto es la densidad de probabilidad condicionada a t_n dada el valor y_{n-1} a t_{n-1} está únicamente determinada y no está afectada por lo que ocurre a tiempos anteriores. A $P_{1|1}$ se le llama *probabilidad de transición*.

- El proceso de Markov está completamente determinado por dos funciones $P_1(y_1, t_1)$ y $P_{1|1}(y_2, t_2 | y_1, t_1)$. A partir de estas funciones se puede obtener toda la jerarquía de probabilidades, por ejemplo para $t_1 < t_2 < t_3$ se tiene

$$\begin{aligned} P_3(y_1, t_1; y_2, t_2; y_3, t_3) &= P_2(y_1, t_1; y_2, t_2) P_{1|2}(y_3, t_3 | y_1, t_1; y_2, t_2) \\ &= P_1(y_1, t_1) P_{1|1}(y_2, t_2 | y_1, t_1) P_{1|1}(y_3, t_3 | y_2, t_2) \end{aligned} \quad (133)$$

y este algoritmo se puede seguir para calcular todas las P_n .

Ejercicio: Mostrar que para cualquier sistema determinista tal que $\dot{x} = f(x)$, con condición inicial x_0 en t_0 , la solución $x = \varphi(x_0, t - t_0)$ es un proceso de Markov con $P_{1|1}(x, t; x_0, t_0) = \delta[x - \varphi(x_0, t - t_0)]$.

- Un ejemplo típico de proceso de Markov en física es el movimiento Browniano. El movimiento browniano muestra de forma muy evidente como la naturaleza discreta de la materia en la escala microscópica se manifiesta a nivel macroscópico. La discretización de la materia causa fluctuaciones en la densidad de materia que a la larga causa efectos observables en el movimiento de la partícula Browniana. Éste se produce cuando una partícula pesada está inmersa en un fluido de partículas ligeras que colisionan de forma aleatoria. La velocidad de la partícula pesada varía como consecuencia de muchas de estas colisiones (supuestamente descorrelacionadas). Consideremos el caso unidimensional y

que la partícula tiene una velocidad V en t , entonces habrá más colisiones por delante que por detrás en promedio, de forma que la probabilidad de que en el siguiente paso Δt haya un cambio δV dependerá de V pero no de valores anteriores, y por lo tanto V constituye un proceso de Markov ($\delta t_c \ll \Delta t \ll \tau_V$). En equilibrio el proceso es estacionario y su tiempo de autocorrelación τ_V es el tiempo en el que una velocidad inicial se hace cero (veremos más adelante que este proceso dinámico corresponde a una partícula de Rayleigh).

Einstein y Smoluchowski razonaron que este no es el comportamiento que se observa experimentalmente. Entre dos observaciones consecutivas de la posición de la partícula pesada, su velocidad crece y decrece muchas veces: el intervalo de observación es mucho mayor que τ_V ($\tau_V \ll \Delta t_X$). Lo que se observa es el desplazamiento neto después de muchas variaciones de la velocidad. Supongamos una serie de observaciones de la partícula que da una secuencia de posiciones X_1, X_2, \dots . Cada desplazamiento $\Delta_{k+1} = X_{k+1} - X_k$ es estocástico, pero su distribución no depende de la historia $(X_{k-1}, X_{k-2}, \dots)$. Es en sí un proceso de Markov en un escala de tiempo promediada (*coarse-grained*) impuesta por las condiciones experimentales.

Hay dos propiedades a destacar en los procesos de Markov:

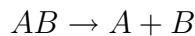
- La propiedad de Markov sólo se cumple de forma aproximada. Así, si $\Delta_k = X_k - X_{k+1}$ es grande hay más probabilidad de que V sea grande en el momento de observación de X_k , velocidad que permanecerá no nula durante un tiempo comparable a τ_V , lo que a su vez favorecerá que el próximo Δ_{k+1} sea también grande. Luego la existencia de $\tau_V > 0$ producirá cierta correlación entre dos desplazamientos consecutivos. Este efecto es pequeño dado que el tiempo de observación es mucho más grande que τ_V .

De igual forma la velocidad es ella misma aproximadamente un proceso de Markov, debido a que las colisiones no son instantáneas. Durante el tiempo en que tiene lugar una colisión el cambio en V en el pasado inmediato dice algo sobre el tipo de colisión y sobre el cambio en V en el futuro inmediato. Este efecto es pequeño pues el tiempo de colisión es pequeño (como si fueran esferas duras).

Además debemos de asumir que el movimiento de la partícula Browniana no genera un flujo ordenado en el fluido que la rodea (arrastre), lo que podría influir sobre la probabilidad de colisión en el futuro, actuando como memoria que violaría la hipótesis marcoviana. Así simulaciones en el ordenador de una gas de moléculas

han mostrado que la velocidad de una de ellas no es un proceso de Markov.

- Un mismo sistema físico puede ser descrito por dos procesos de Markov diferentes dependiendo del grado de *coarseness* (promediado) de la descripción. Esto ocurre en el caso de un gas de moléculas binarias que se disocia



En la escala temporal promedio (coarse-grained) cada molécula AB tiene una cierta probabilidad por unidad de tiempo de romperse por alguna colisión. Entonces el cambio en la concentración entre t y $t + \Delta t$ tiene una cierta distribución de probabilidad que depende de la concentración en t pero no en tiempos anteriores. Así, la concentración es un proceso de Markov. En una descripción más detallada (microscópica) del mecanismo de ruptura uno distingue los estados vibracionales de la molécula y estudia cómo las sucesivas colisiones hacen que la molécula cambie estocásticamente entre los distintos estados vibracionales hasta que se supera la barrera del potencial y la molécula se disocia. Este camino aleatorio entre los niveles vibracionales es otro proceso de Markov.

- El concepto de proceso de Markov se puede aplicar a procesos $Y(t)$ con r componentes: las tres componentes de la velocidad de una partícula Browniana; o las r componentes de una mezcla química reactante.

En un proceso con r componentes que es markoviano, el proceso constituido por $s < r$ componentes en general puede no serlo. Por ejemplo en el caso de la disociación de la molécula la distribución de probabilidad futura de una cantidad de cada componente químico está determinada por la presente cantidad de todos los componentes. Por el contrario, si un proceso estocástico físico es no Markoviano, es posible introducir nuevas componentes de forma que el conjunto sea Markoviano. Por ejemplo, un partícula Browniana bajo la acción de un campo externo in-homogéneo. El cambio en la velocidad en Δt está ahora determinado por las colisiones y el campo externo, y por lo tanto por la posición de la partícula. Esta depende de la velocidad en los tiempos anteriores, por lo que la velocidad no es un proceso de Markov. Sin embargo el proceso constituido por la velocidad y la posición de la partícula

Browniana si es un proceso de Markov (ver sección VIII.7 del van Kampen).

En principio, cualquier sistema físico aislado se puede describir como un proceso de Markov, introduciendo todas las variables microscópicas como componentes de Y . De hecho el movimiento microscópico en el espacio de las fases es determinista y por lo tanto marcoviano (ver ejercicio anterior). La cuestión es si existe un mínimo número de variables para dicho sistema cuyo comportamiento en el tiempo es un proceso de Markov multicomponente. El caso es que esto ocurre para la mayoría de los sistemas en la naturaleza, lo que ocurre es que tal descripción es aproximada y restringida a un nivel *coarse-grained* macroscópico.

La ecuación de Chapman-Kolmogorov

Si integramos la expresión (133) sobre y_2 para $t_1 < t_2 < t_3$ se tiene

$$P_2(y_1, t_1; y_3, t_3) = P_1(y_1, t_1) \int P_{1|1}(y_2, t_2|y_1, t_1) P_{1|1}(y_3, t_3|y_2, t_2) dy_2. \quad (134)$$

de donde dividiendo por $P_1(y_1, t_1)$ se tiene

$$P_{1|1}(y_3, t_3|y_1, t_1) = \int P_{1|1}(y_3, t_3|y_2, t_2) P_{1|1}(y_2, t_2|y_1, t_1) dy_2. \quad (135)$$

Esta ecuación se llama la *ecuación de Chapman-Kolmogorov*. Es una identidad que deben satisfacer las probabilidades de transición para cualquier proceso de markov. En esta identidad el orden de tiempos es esencial. Esta ecuación también se cumple cuando y es un vector con r componentes o cuando la y sólo toma valores discretos (en este caso la integral se sustituye por una suma).

Dado que P_1 y $P_{1|1}$ determinan completamente un proceso de Markov, éstas no pueden ser cogidas arbitrariamente sino que deben verificar las siguientes identidades:

- (i) La ecuación de Chapman-Kolmogorov
- (ii) La relación trivial (Bayes y propiedad jerárquica)

$$P_1(y_2, t_2) = \int P_{1|1}(y_2, t_2|y_1, t_1) P_1(y_1, t_1) dy_1 (Ejercicio). \quad (136)$$

o al contrario, cualquiera dos funciones no negativas P_1 y $P_{1|1}$ que verifican estas relaciones definen univocamente un proceso de Markov.

Los dos siguientes ejemplos de procesos de Markov son de especial importancia.

- **Procesos de Wiener-Levy.** Consideremos la probabilidad de transición

$$P_{1|1}(y_2, t_2 | y_1, t_1) = \frac{1}{\sqrt{2\pi(t_2 - t_1)}} \exp \left[-\frac{(y_2 - y_1)^2}{2(t_2 - t_1)} \right] \quad (137)$$

is tomamos $P_1(y_1, 0) = \delta(y_1)$ tenemos un proceso de Markov no estacionario llamado proceso de Wiener-Levy. Está considerado sólo para $t > 0$ y fue propuesto para describir el comportamiento estocástico de la partícula Browniana (capítulo VIII del van Kampen). Usando la segunda condición de proceso de Markov se tiene

$$P_1(y, t) = \frac{1}{\sqrt{2\pi t}} \exp \left[-\frac{y^2}{2t} \right] \quad (138)$$

- **Procesos de Poisson.** Supongamos $Y(t)$ que toma valores $n = 0, 1, 2, \dots$ y $t \geq 0$. Definimos un proceso de Markov mediante $(t_2 \geq t_1 \geq 0)$

$$P_{1|1}(n_2, t_2 | n_1, t_1) = \frac{(t_2 - t_1)^{n_2 - n_1}}{(n_2 - n_1)!} e^{-(t_2 - t_1)}, \quad P_1(n, 0) = \delta_{n,0} \quad (139)$$

se entiende que $P_{1|1} = 0$ para $n_2 < n_1$. Cada función muestra $y(t)$ es una sucesión de pasos de altura unidad a instantes de tiempo aleatorios. Está únicamente determinado por los instantes temporales a los que tienen lugar los saltos y que constituye un conjunto aleatorio de puntos en el eje temporal. Su número entre dos tiempos t_1, t_2 está distribuido de acuerdo a la distribución (139). A $Y(t)$ se le llama proceso de Poisson.

Procesos de Markov estacionarios

Los procesos estocásticos que son a la vez Markovianos y estacionarios tienen un interés especial porque describen fluctuaciones de equilibrio.

Supongamos que un sistema físico cerrado y aislado tiene una cierta magnitud, o conjunto de magnitudes $Y(t)$ que es un proceso de Markov. Cuando el sistema está en equilibrio $Y(t)$ es un proceso de Markov estacionario. Así, P_1 es independiente del tiempo y no es sino la distribución de equilibrio de la cantidad Y como la determinada por la mecánica estadística. Por ejemplo las fluctuaciones del voltaje a través de un condensador en un circuito RC son markovianas. Cuando R está a temperatura fija T , el voltaje es un proceso

de Markov estacionario con

$$P_1(y_1) = \left(\frac{C}{2\pi kT} \right)^{1/2} \exp \left[-\frac{Cy_1^2}{2kT} \right]. \quad (140)$$

Incluso cuando un sistema está en un estado estacionario diferente del equilibrio ciertas magnitudes físicas pueden ser un proceso de Markov estacionario. Por ejemplo en una partícula browniana sometida a un campo gravitatorio homogéneo: su velocidad vertical es un proceso estacionario, pero su posición no.

Para un proceso de Markov estacionario la probabilidad de transición $P_{1|1}$ no depende en t_1 y t_2 sino en la diferencia $\tau = t_2 - t_1$. Entonces denotamos

$$P_{1|1}(y_2, t_2 | y_1, t_1) = T_\tau(y_2 | y_1). \quad (141)$$

y la ecuación de Chapman-Kolmogorov queda:

$$T_{\tau+\tau'}(y_3 | y_1) = \int T_{\tau'}(y_3 | y_2) T_\tau(y_2 | y_1) dy_2. \quad (142)$$

Si miramos la integral como un producto de dos matrices, o kernels integrales la anterior ecuación queda

$$T_{\tau+\tau'} = T_{\tau'} T_\tau \quad (\tau, \tau' > 0) \quad (143)$$

Aunque (142) y (143) se cumplen sólo para $\tau, \tau' > 0$, la definición (141) no se restringe a $\tau > 0$. Se tiene que

$$P_2(y_2, t_2; y_1, t_1) = T_\tau(y_2 | y_1) P(y_1) \quad (144)$$

y como P_2 es simétrica

$$T_\tau(y_2 | y_1) P(y_1) = T_{-\tau}(y_1 | y_2) P(y_2), \quad (145)$$

que es una identidad para todos los procesos de Markov estacionarios y se puede aplicar a todos los sistemas físicos en equilibrio sin ninguna derivación adicional de las ecuaciones del movimiento (no confundir con la condición de balance detallado que es exactamente la misma pero cambiando $-\tau$ por τ en la derecha de la identidad y que requiere una derivación física).

Ejercicio: *Demostrar que la función de autocorrelación de un proceso de Markov estacionario con media cero es*

$$\kappa(\tau) = \iint y_1 y_2 T_\tau(y_2 | y_1) P_1(y_1) dy_1 dy_2 \quad (\tau \geq 0) \quad (146)$$

Ejercicio: *Demostrar las identidades*

$$\int T_\tau(y_2|y_1)dy_2 = 1 \quad (147)$$

$$\int T_\tau(y_2|y_1)P_1(y_1)dy_1 = P_1(y_2) \quad (148)$$

Según estas propiedades T_τ se puede ver como un operador con autovalor 1, con autovector por la izquierda $\psi(y) \equiv 1$ y autovector por la derecha P_1 .

- El ejemplo más conocido de proceso de Markov estacionario es el proceso de *Ornstein-Uhlenbeck* definido por

$$P_1(y_1) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y_1^2} \quad (149)$$

$$T_\tau(y_2|y_1) = \frac{1}{\sqrt{2\pi(1-e^{-2\tau})}} \exp\left[-\frac{(y_2 - y_1 e^{-\tau})^2}{2(1-e^{-2\tau})}\right] \quad (150)$$

Ejercicio: *Verificar que es un proceso de Markov*

Ejercicio: *Probar que $T(y_2|y_1)$ verifica:*

$$\frac{\partial T}{\partial \tau} = \frac{\partial}{\partial y_2}y_2T + \frac{\partial^2 T}{\partial y_2^2} \quad (151)$$

$$\frac{\partial T}{\partial \tau} = -y_1 \frac{\partial T}{\partial y_1} + \frac{\partial^2 T}{\partial y_1^2} \quad (152)$$

Claramente tiene media cero y autocorrelación

$$\kappa(\tau) = e^{-\tau} \quad (153)$$

El proceso de OU es estacionario, Gaussiano y Markoviano. El teorema de Doob muestra que salvo transformación lineal es el único proceso con estas tres propiedades. Vamos a verlo:

- Sea $Y(t)$ un proceso estacionario Gaussiano de Markov. Por ser Gaussiano $P_1(y)$ es Gaussiana y podemos suponerla igual a (149) (mediante reescalamientos y traslaciones de la variable). Por ser Gaussiano la probabilidad de transición tiene la forma general

$$T_\tau(y_2|y_1) = De^{-\frac{1}{2}(Ay_2^2 + 2By_2y_1 + Cy_1^2)}, \quad (154)$$

donde A, B, C, D son funciones de τ . La normalización (147) da

$$D = \sqrt{A/2\pi} \quad C = B^2/A \quad (155)$$

mientras que (148) da $B^2 = A(A - 1)$. El parámetro A se puede poner en términos de la función de autocorrelación (146) que da $A = (1 - \kappa^2)^{-1}$ e implica

$$T_\tau(y_2|y_1) = \frac{1}{\sqrt{2\pi(1 - \kappa^2)}} \exp\left[-\frac{(y_2 - \kappa y_1)^2}{2(1 - \kappa^2)}\right]. \quad (156)$$

Introduciendo $t_3 = t_2 + \tau'$ y la ecuación Chapman-Kolmogorov para un proceso de Markov estacionario se tiene

$$\begin{aligned} \kappa(t_3 - t_1) &= \int y_3 dy_3 \iint T_{\tau'}(y_3|y_2) T_\tau(y_2|y_1) y_1 P_1(y_1) dy_1 dy_2 \\ &= \kappa(t_3 - t_2) \kappa(t_2 - t_1) \end{aligned} \quad (157)$$

Ejercicio: Demostrarlo

que sugiere utilizar

$$\kappa(\tau) = e^{-\gamma\tau} \quad (158)$$

con lo que está demostrado el teorema de Doob.

- Si $Y(t)$ es estacionario, Gaussiano, y tiene una función de autocorrelación exponencial $\kappa(\tau) = \kappa(0)e^{-\gamma\tau}$ entonces $Y(t)$ es un proceso de OU y por tanto markoviano. Suponiendo media cero y varianza unidad su funcional generador es

$$G[k] = \exp\left[-\frac{1}{2} \iint k(t_1)k(t_2)e^{-\gamma|t_1-t_2|} dt_1 dt_2\right] \quad (159)$$

- Consideremos un proceso de Markov estacionario $Y(t)$ dado por $P_1(y_1)$ y $T_\tau(y_2|y_1)$. Consideremos un tiempo t_0 fijo y un valor y_0 fijo (esto es, $p(y) = \delta(y - y_0)$). Definimos un nuevo proceso no estacionario de Markov $Y^*(t)$ para $t \geq t_0$, de la siguiente forma

$$P_1^*(y_1, t_1) = T_{t_1 - t_0}(y_1|y_0) = \int T_{t_1 - t_0}(y_1|y) p(y) dy \quad (160)$$

$$P_{1|1}^*(y_2, t_2|y_1, t_1) = T_{t_2 - t_1}(y_2|y_1) \quad (161)$$

Podemos extraer un *subconjunto* en el que a un tiempo t_0 los valores de $Y(t_0)$ están distribuidos de acuerdo con la distribución $p(y_0)$. Entonces

tomamos

$$P_1^*(y_1, t_1) = \int T_{t_1-t_0}(y_1|y_0)p(y_0)dy_0 \quad (162)$$

$$P_{1|1}^*(y_2, t_2|y_1, t_1) = T_{t_2-t_1}(y_2|y_1) \quad (163)$$

estos procesos son no estacionarios pues imponemos una condición a un tiempo fijo t_0 (implica una condición inicial). Aun así la probabilidad de transición depende del intervalo temporal como la del proceso de Markov estacionario del que provienen. A este tipo de procesos Markov no estacionarios con probabilidad de transición no dependiente del tiempo se le llaman *procesos homogéneos*. El proceso de Wiener es un ejemplo de proceso homogéneo que no está embebido en un proceso de Markov estacionario.

Físicamente la extracción de este subproceso no estacionario significa que uno prepara el sistema en un estado de no-equilibrio. Uno espera que tras un tiempo el sistema vuelve al equilibrio

$$P_1^*(y_1, t_1) \rightarrow P_1(y_1) \text{ cuando } t \rightarrow \infty \quad (164)$$

como esto tiene que ser cierto para cualquier estado preparado arbitrario $p(y_0)$ se debe cumplir

$$T_{t_1-t_0}(y_1|y_0) \rightarrow P_1(y_1) \quad (165)$$

Por ejemplo, vamos a ver el caso de la teoría de la respuesta lineal. Consideremos un material paramagnético bajo la acción de un campo magnético uniforme externo F . La magnetización Y en la dirección del campo es un proceso estocástico estacionario con promedio macroscópico y fluctuaciones pequeñas entorno a este promedio. Supongamos que es un proceso de Markov. La función $P_1(y)$ viene por la distribución canónica

$$P_1(y) = \frac{\text{Tr}[\delta(y - Y)e^{-(H-BY)/kT}]}{\text{Tr}e^{-(H-BY)/kT}} \quad (166)$$

donde H es el operador hamiltoniano sin campo magnético externo y Y es el operador magnetización. La traza aparece porque nuestro sistema es cuántico. Supongamos que para $-\infty < t < t_0$ el campo vale $B + \Delta B$ y en t_0 cambia instantáneamente a B . entonces la distribución de probabilidad en t_0 es

$$p(y) = \frac{\text{Tr}[\delta(y - Y)e^{-\{H-(B+\Delta B)Y\}/kT}]}{\text{Tr}e^{-\{H-(B+\Delta B)Y\}/kT}} \quad (167)$$

Para $t > t_0$ la magnetización Y es un proceso homogéneo cuya distribución inicial es p y cuya probabilidad de transición es la misma que la de equilibrio con campo B . Entonces:

$$\langle Y(t) \rangle^* = \iint y T_{t-t_0}(y|y_0) p(y_0) dy_0 dy \quad (168)$$

esto es, la forma en que el promedio macroscópico se ajusta al nuevo campo está determinada por la probabilidad de transición de las fluctuaciones de equilibrio en el nuevo campo B .

Si sólo estamos interesados en la respuesta lineal expandimos $p(y)$ hasta primer orden en ΔB dando

$$p(y) = P_1(y) + \frac{\Delta B}{kT} \{y - \langle Y \rangle_B\} P_1(y) + \mathcal{O}(\Delta B^2) \quad (169)$$

de donde

$$\langle Y(t) \rangle^* = \langle Y \rangle_B + \frac{\Delta B}{kT} \kappa(t - t_0). \quad (170)$$

es decir la relajación irreversible de la magnetización promedio está determinada en aproximación lineal por la función de autocorrelación de las fluctuaciones de equilibrio. Esto es un ejemplo del teorema de fluctuación-disipación.

- Para derivar el resultado anterior hemos supuesto que $Y(t)$ es de Markov. Usualmente se está interesado en materiales en los que hay un efecto de memoria, pues ésta da más información sobre los momentos magnéticos microscópicos y su interacción. Aunque formalmente los resultados no cambian y que $p(y_0)$ es la distribución de $Y(t)$ en t_0 en el que el campo B cambia, no es cierto ahora que $p(y_0)$ especifica unívocamente el subconjunto y por tanto el futuro de $Y(t)$.
- Lo que es esencial es que el sistema haya alcanzado cierto estado estacionario en presencia de $B + \Delta B$ de forma que su densidad en el espacio de las fases sea canónica (colectividad) no sólo con respecto a Y si con las otras cantidades que determinan su futuro. El resultado obtenido no se puede aplicar por tanto para el caso de campos externos dependientes del tiempo, al menos que la variación del campo sea tan lenta que el sistema puede mantener en todo momento la distribución de equilibrio correspondiente a $B(t)$ instantáneo.

Cadenas de Markov

Es una clase especialmente sencilla de procesos de Markov que cumple:

- (i) El rango de Y es un conjunto discreto de estados
- (ii) La variable temporal es discreta y toma sólo valores enteros $t = \dots, -2, -1, 0, 1, \dots$
- (iii) El proceso es estacionario o al menos homogéneo, de forma que la probabilidad de transición sólo depende de la diferencia de tiempos.

Una cadena de Markov finita es aquella que tiene como rango un número finito de estados N . Su distribución primera de probabilidad $P_1(y, t)$ es un vector de N componentes $p_n(t)$ ($n=1, \dots, N$). La probabilidad de transición $T_\tau(y_2|y_1)$ es una matriz $N \times N$. La propiedad de Markov implica:

$$T_\tau = (T_1)^\tau \quad (171)$$

La distribución de probabilidad $p(t)$ que se obtiene de una distribución inicial $p(0)$ es

$$p(t) = T^t p(0). \quad (172)$$

Entonces el estudio de una cadena de Markov finita se reduce a estudiar las potencias de una matriz $N \times N$, esto es T , de la que conocemos que

- (i) sus elementos son no negativos y
- (ii) la suma de los elementos de cada columna es la unidad.

Estas matrices se llaman *matrices estocásticas* (no confundir con matrices aleatorias, cuyos elementos son variables estocásticas).

T tiene un autovector por la izquierda $(1, \dots, 1)$ con autovalor 1, y un autovector por la derecha p^s tal que $Tp^s = p^s$, que no es sino la $p_1(y)$ de un proceso estacionario (ver ejercicio anterior). No es necesario que sea un estado físico de equilibrio, puede ser un estado estacionario en el cual, por ejemplo, se mantiene un flujo constante.

Lo primero que hay que demostrar es que dada una distribución inicial se tiene

$$\lim_{t \rightarrow \infty} p(t) = \lim_{t \rightarrow \infty} T^t p(0) = p^s. \quad (173)$$

El teorema de Perron y Frobenius muestra que esto es siempre cierto salvo casos muy excepcionales.

Nota: hay procesos estocásticos que pueden hacerse markovianos introduciendo una variable adicional. Estos procesos se llaman *procesos markovianos de segundo grado*. Por ejemplo el caminante aleatorio con persistencia.

En este proceso la probabilidad α de un paso en la misma dirección que el paso anterior difiere de la probabilidad β en sentido contrario. El proceso es no markoviano pues su distribución de probabilidad en el tiempo $r + 1$ depende no sólo de su valor en r sino también de su valor en $r - 1$. El proceso puede hacerse de Markov considerando una variable adicional, esto es la posición en $r - 1$. El proceso bicomponente resultante es markoviano con probabilidad de transición:

$$\begin{aligned} T(n_2, m_2 | n_1, m_1) = & [\delta_{n_2, n_1+1}(\alpha \delta_{n_1, m_1+1} + \beta \delta_{n_1, m_1-1}) \\ & + \delta_{n_2, n_1-1}(\alpha \delta_{n_1, m_1-1} + \beta \delta_{n_1, m_1+1})] \delta_{m_2, m_1} \end{aligned} \quad (174)$$

Procesos de decaimiento

Para finalizar y como ejemplo ilustrativo vamos a estudiar el proceso de decaimiento de un material radioactivo. Consideremos que en $t = 0$ hay n_0 núcleos activos. El número $N(t)$ de núcleos activos que sobreviven a tiempo t es un proceso estocástico no estacionario. Es de Markov pues la distribución de probabilidad de $N(t_2)$ con $t_2 > t_1$ condicionada a que $N(t_1) = n_1$ no depende de la historia (depende sólo de la probabilidad individual de sobrevivir de cada núcleo). El proceso es entonces una combinación de procesos de decaimientos mútuamente independientes de los núcleos individuales. Sea w la probabilidad de un núcleo de sobrevivir en t_1 entonces la probabilidad de sobrevivir de n_1 núcleos en t_1 es

$$P_1(n_1, t_1) = \binom{n_0}{n_1} w^{n_1} (1-w)^{n_0-n_1} \quad (175)$$

esta fórmula es válida para todos los valores enteros de n_1 si hacemos $P_1(n_1, t_1) = 0$ para $n_1 < 0$ y $n_1 > n_0$.

De esta fórmula se tiene

$$\begin{aligned} \langle N(t_1) \rangle &= \left[\sum_{n_1} n_1 \binom{n_0}{n_1} w^{n_1} (1-w)^{n_0-n_1} \right]_{v=1-w} \\ &= \left[w \frac{\partial}{\partial w} (w + v)^{n_0} \right]_{v=1-w} \\ &= w n_0 \end{aligned} \quad (176)$$

Nos queda calcular w , pero sabemos que la probabilidad por unidad de tiempo dt de un núcleo en decaer es constante, es decir γ , de donde la variación en la población de núcleos en la unidad de tiempo es $dw = -\gamma w dt$, de donde

$$w(t) = e^{-\gamma t} \quad (177)$$

que da para P_1

$$P_1(n_1, t_1) = \binom{n_0}{n_1} e^{-\gamma t n_1} (1 - e^{-\gamma t})^{n_0 - n_1}. \quad (178)$$

y para la probabilidad de transición para $t_2 > t_1$

$$P_{1|1}(n_2, t_2 | n_1, t_1) = \binom{n_1}{n_2} e^{-\gamma(t_2 - t_1)n_2} (1 - e^{-\gamma(t_2 - t_1)})^{n_1 - n_2}. \quad (179)$$

que determina completamente el proceso.

Este tipo de cálculos se aplica también a la emisión de luz por átomos excitados, el escape de moléculas de un gas de Knudsen a través de un pequeño filtro, la destrucción de tropas enemigas mediante disparos aleatorios o la destrucción de células por radiación.