



aula - Processos Markovianos

①

Conforme veremos ao longo deste curso, uma série de processos, dentre eles dinâmicas de partículas e reações químicas, são processos aproximadamente markovianos, onde a probabilidade de ~~uma~~ dada ocorrência de um dado processo estar ~~traz~~ só depende do valor no estado determinado instantaneamente e não do valor que as variáveis ~~estavam~~ ~~traz~~ tinham em instantes anteriores.

O próprio movimento Browniano pode ser tratado como um processo markoviano, sendo que no caso de uma partícula "pesada" colidindo em um fluido "leve" a probabilidade de um valor ΔV ~~na~~ na velocidade no próximo instante de tempo Δt depende apenas de V , e não dos valores anteriores de velocidade.

Em termos mais concretos; vamos considerar

seja X_t uma variável estocástica discretizada ^{o caso das Caídas (variável de Markov (tempo discretos))}
(por simplicidade) que assume ~~os~~ valores inteiros e o tempo $0, 1, 2, 3, \dots$ (tem valores discretos)

O processo estocástico fica inteiramente definido até o instante t pela distribuição de probabilidades conjunta

$$P_2(n_0, n_1, \dots, n_t) \rightarrow$$



NOME: _____

PROFESSOR: _____

DATA: _____

de x_t tomar o valor n_0 em $t=0$,
o valor n_1 em $t=1$,
o valor n_2 em $t=2$,
:
o valor n_l em $t=l$.

O processo estocástico é dito markoviano se a probabilidade condicional

$$P_{l+1}(n_{l+1} / n_0, n_1, \dots, n_l) \text{ for}$$

igual a $\underbrace{P_{l+1}(n_{l+1} / n_l)}$

de x_t tomar o valor n_{l+1} em $t=l+1$

dado que x_t tomou o valor n_l em $t=l$.

Uma consequência imediata da propriedade acima é obtida considerando a probabilidade conjunta

$$P_l(n_0, n_1, n_2, \dots, n_l).$$

Para tal, devemos lembrar que

(2)

$$P_e(n_0, n_1, \dots, n_e) = P_e(n_e / n_0, n_1, \dots, n_{e-1}) P_{e-1}(n_0, \dots, n_{e-1})$$

Como o processo é Markoviano temos que

$$= P_e(n_e / n_{e-1}) P_{e-1}(n_0, \dots, n_{e-1})$$

Usando mais uma vez a ~~definição~~ definição de processos markovianos temos

$$P_e(n_0, n_1, \dots, n_e) = P_e(n_e / n_{e-1}) P_{e-1}(n_{e-1} / n_{e-2}) \cdot P_{e-2}(n_0, \dots, n_{e-2})$$

e sucessivamente obtemos a expressão

$$P_e(n_0, n_1, \dots, n_e) = P_e(n_e / n_{e-1}) P_{e-1}(n_{e-1} / n_{e-2}) P_{e-2}(n_{e-2} / n_{e-3}) \cdot \dots P_0(n_0)$$

↑ válida para processos markovianos

A partir da expressão ^{acima} para a probabilidade conjunta, podemos obter a probabilidade marginal $P_e(n_e)$, que é a probabilidade de que a variável aleatória x_t tome o valor n_e em $t = e$

independentemente de quais valores ela tenha tomado nos instantes anteriores:

$$P_e(n_e) = \sum_{n_0} \sum_{n_1} \dots \sum_{n_{e-1}} P_e(n_0, n_1, \dots, n_e)$$

$$P_e(n_e) = \sum_{n_0} \sum_{n_1} \dots \sum_{n_{e-1}} P_e(n_e | n_{e-1}) \dots \underbrace{P_1(n_1 | n_0) P_0(n_0)}_{P_1(n_1, n_0)}$$

$$= \sum_{n_1} \sum_{n_2} \dots \sum_{n_{e-1}} P_e(n_e | n_{e-1}) \dots P_1(n_1)$$

... Efetuando os cálculos de forma análoga

$$P_e(n_e) = P_e(n_e | n_{e-1}) P_{e-1}(n_{e-1})$$

Portanto, conhecendo $P_0(n_0)$ e $P_e(n_e | n_{e-1})$ podemos obter $P_e(n_e)$ em qualquer instante.

A probabilidade condicional $P_e(n_e | n_{e-1})$ pode ser interpretada como a probabilidade de transição do estado n_{e-1} para o estado n_e .

Em princípio ela pode depender do tempo (estudaremos aqui futuramente este caso).

Definindo $P_e(n_e | n_{e-1}) \equiv T(n_e, n_{e-1})$,

de forma que a equação acima se torna

$$P_e(n_e) = \sum_{n_{e-1}} T(n_e, n_{e-1}) P_{e-1}(n_{e-1}),$$

reforcando mais uma vez que um processo estocástico markoviano fica completamente definido pela probabilidade de transição e pela probabilidade inicial. A matriz de transição $T(n_e, n_{e-1})$ é denominada matriz estocástica.

Reescrevendo ainda a equação acima como

$$P_e(n) = \sum_m T(n, m) P_{e-1}(m),$$

onde $T(n, m) \geq 0$ (probabilidade)

e $\sum_n T(n, m) = 1$ de forma

que a soma das probabilidades de transição de um estado m para todos os estados n 's possíveis deve ser normalizada



$$\sum_n T(n, m) = 1$$

Temos ainda que $\rightarrow$

$$P_1(n) = \sum_m T(n, m) P_0(m)$$

$$P_2(n) = \sum_m T(n, m) P_1(m)$$

$$= \sum_m \sum_{m'} T(n, m) T(n, m') P_0(m')$$

ou sob a forma compacta

$$P_2 = T^2 P_0$$

Procedendo analogamente:

$$P_e = T P_{e-1} = T^e P_0,$$

e o problema de determinar $P_e(n)$ se reduz ao cálculo da e -ésima

potência da matriz estocástica

Ex: $T = \begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix}$

$$\begin{matrix} p_1 + q_1 = 1 \\ p_2 + q_2 = 1 \end{matrix}$$

$$T = \begin{pmatrix} 1-b & q \\ b & 1-q \end{pmatrix}$$

Teorema de Perron-Frobenius

4

O problema fundamental em processos markovianos é determinar quais são as propriedades da matriz estocástica T tal que

$$\lim_{t \rightarrow \infty} P_t \rightarrow P_{est.} \quad (\text{probabilidade estacionária})$$

que satisfaz a condição

$$T P_{est} = P_{est}.$$

Dois consequências são imediatas.

a 1ª é que T possui

auto vetor à esquerda dado por $(1 \ 1 \ 1 \dots 1)$ com autovalor 1 [tendo em vista que a soma dos elementos de uma coluna é 1 devido à normalização].

$$(1 \ 1) \begin{pmatrix} 1-b & q \\ b & 1-q \end{pmatrix} = \underline{1} (1 \ 1)$$

a 2ª consequência é que o auto vetor à

direita P_{est} satisfazendo $T P_{est} = P_{est}$ é um estado estacionário (ele não é necessariamente um estado de equilíbrio, podendo representar um estado estacionário com um fluxo constante).

O Teorema de Perron - Frobenius estabelece que o autovalor $\lambda = 1$ de uma matriz irredutível e não degenerado com autovetor correspondente possuindo todos os componentes estritamente positivos. Portanto, o Teorema de Perron - Frobenius afirma que a solução estacionária $T P_{est} = P_{est}$ é única e que $P_{est}(n) > 0$.

Além disso, com exceção do autovalor $\lambda = 1$, todos os autovalores de uma matriz regular são em módulo estritamente menores que a unidade $|\lambda| < 1$. Quando $l \rightarrow \infty$

T^l converge para uma matriz cujas colunas são todas iguais a P e portanto

$$\lim_{l \rightarrow \infty} T^l P_0 \rightarrow P_{est} \text{ qualquer que seja } P_0.$$

Pode haver autovalores tais que $|\lambda| = 1$, além do autovalor $\lambda = 1$. Isso ocorre com as matrizes cíclicas. Neste caso, a probabilidade não cresce entre os valores de λ que são iguais a 1.

Existem várias provas do Teorema de Perron - Frobenius, dentre elas algumas baseadas no teorema espectral

Conforme mencionamos anteriormente,
 Dada a matriz de transição T e a probabilidade inicial P_0 , podemos obter P_t a partir de $P_t = T^t P_0$. Uma forma conveniente de calcularmos T^t é a partir dos autovalores e autovetores de T .

Considerando o caso em que os autovalores de T sejam não degenerados, devemos ter em mente que os autovetores à direita e esquerda podem ser diferentes (matriz T é em geral não simétrica)

sejam $|\phi_k\rangle$ e $\langle\phi_k|$ os autovetores à direita e esquerda e λ_k o correspondente autovalor, de forma que

$$|\phi_j\rangle\langle\phi_k| = \delta_{jk} \quad e$$

$$\sum_k |\phi_k\rangle\langle\phi_k| = \underline{I}$$

matriz identidade.

Conforme mencionamos anteriormente, a probabilidade estacionária P_{est} corresponde ao autovetor $|\phi_0\rangle$ com autovalor $\lambda_0 = 1$ pois $T|\phi_0\rangle = |\phi_0\rangle$.

Além disso $\langle \phi_0 | T = \langle \phi_0 |$ devido a

normalização $\sum_n T(n, m) = 1$ e portanto

$\langle \phi_0 | = (1 \dots 1)$. Finalmente a normalização

$P_{est} = |\phi_0\rangle$ nos dá $\langle \phi_0 | \phi_0 \rangle = 1$

v^a. Multiplicando ambos os membros da expansão $T^l |\phi_k\rangle \langle \phi_k| = T^l I$.

Uma vez que $T |\phi_k\rangle = \lambda_k |\phi_k\rangle$,

segue que $T^l |\phi_k\rangle = \lambda_k^l |\phi_k\rangle$ e portanto

$$T^l = \sum_k \lambda_k^l |\phi_k\rangle \langle \phi_k|$$

$$T^l = \lambda_0^l |\phi_0\rangle \langle \phi_0| + \sum_{k \neq 0} \lambda_k^l |\phi_k\rangle \langle \phi_k|$$

$$T^l P_0 = \lambda_0^l |\phi_0\rangle \langle \phi_0| P_0 + \sum_{k \neq 0} \lambda_k^l |\phi_k\rangle \langle \phi_k| P_0$$

Uma vez que $|\lambda_k| < 1$, temos que $|\lambda_k|^l \rightarrow 0$ quando $l \rightarrow \infty$, de forma que

$$T^e P_0 \rightarrow \lambda_0^e |\phi_0\rangle \langle \phi_0| P_0 \dots e$$

(6)

por tanto $P_e \rightarrow \lambda_0^e \propto |\phi_0\rangle$ e $P_{est.}$

assim $P_e \rightarrow P_{est.}$, qualquer que seja a probabilidade inicial P_0 .

Aplicação:

$$T = \begin{pmatrix} 1-b & q \\ b & 1-q \end{pmatrix}$$

$$T = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \Leftrightarrow \begin{matrix} T_{1 \rightarrow 1} = 1-b \\ T_{1 \rightarrow 2} = b \\ T_{2 \rightarrow 1} = q \\ T_{2 \rightarrow 2} = 1-q \end{matrix}$$

$$T_{2 \rightarrow 1} = q$$

$$T_{2 \rightarrow 2} = 1-q$$

• Note que $T_{11} + T_{21} = 1$

$$T_{12} + T_{22} = 1$$

> Normalização da probabilidade

Cálculo dos autovalores

$$\begin{vmatrix} 1-b-\lambda & q \\ b & 1-q-\lambda \end{vmatrix} = 0$$

$$(1-b-\lambda)(1-q-\lambda) - qb = 0$$

$$\lambda^2 - \lambda(1-b+1-q) - qb + (1-q)(1-b) = 0$$

$$\lambda^2 - \lambda(2-q-b) - qb + 1-b-q+qb = 0$$

(8)

after obtaining the sets of eigenvalues and eigenvectors, we have that

$$\begin{aligned} T^l P_0 &= \frac{1}{q+b} \begin{pmatrix} q & \\ & b \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} + \frac{(1-q-b)^l}{(q+b)^l} \begin{pmatrix} 1 \\ -1 \end{pmatrix} (b-q) \\ &= \frac{1}{q+b} \begin{pmatrix} q & q \\ b & b \end{pmatrix} + \frac{(1-q-b)^l}{(q+b)^l} \begin{pmatrix} b & -q \\ -b & q \end{pmatrix} \end{aligned}$$

For an ^{initial} probability $P_0 = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$

$$\begin{aligned} P_l &= \frac{1}{q+b} \begin{pmatrix} q & q \\ b & b \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + \frac{(1-q-b)^l}{q+b} \begin{pmatrix} b & -q \\ -b & q \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \\ &= \frac{1}{q+b} \begin{pmatrix} q(p_1 + p_2) \\ b(p_1 + p_2) \end{pmatrix} + \frac{(1-q-b)^l}{q+b} \begin{pmatrix} b p_1 - q p_2 \\ -b p_1 + q p_2 \end{pmatrix} \end{aligned}$$

For $q+b \neq 0$: $P_l \rightarrow \frac{1}{q+b} \begin{pmatrix} q \\ b \end{pmatrix}$ irrespective
the P_0 for $l \rightarrow \infty$.

Reversibilidade Microscópica

9

Um conceito importante em processos markovianos é o da reversibilidade microscópica. A reversibilidade microscópica é capaz de distinguir processos em equilíbrio termodinâmico (descreto pela distribuição de probabilidade de Gibbs $P_{eq} = \frac{e^{-\beta H(o)}}{Z}$, $Z = \sum_{\sigma} e^{-\beta H(\sigma)}$) e sistemas fora do equilíbrio termodinâmico, cuja distribuição de probabilidades é desconhecida a priori na maioria dos casos. H .

Para uma cadeia de Markov cujas probabilidades de transição são independentes do tempo (p/ taxas dependentes do tempo, geralmente a condição de reversibilidade microscópica pode ser localmente verificada, a cada instante de tempo, mas não globalmente), a probabilidade de uma trajetória $n_0 \rightarrow n_1 \rightarrow \dots \rightarrow n_L$ é dada por

$$P(n_0, n_1, \dots, n_L) = T(n_L, n_{L-1}) T(n_{L-1}, n_{L-2}) \dots T(n_1, n_0) P(n_0)$$

(*)

Por simplicidade, consideramos o estado estacionário de modo que a probabilidade estacionária $P(n)$ satisfaz a relação

$$P(n) = \sum_m T(n, m) P(m)$$

Consideramos agora a trajetória reversa $n_e \rightarrow \dots \rightarrow n_1 \rightarrow n_0$. A probabilidade dessa trajetória é dada por

$$P(n_e, n_{e-1}, \dots, n_1, n_0) = T(n_0, n_1) T(n_1, n_2) \dots T(n_{e-1}, n_e) P(n_e)$$

Se definirmos $\tilde{T}(n, m) = T(m, n) \frac{P(n)}{P(m)}$,

a relação acima pode ser escrita como

$$P(n_e) \left[\tilde{T}(n_e, n_{e-1}) \frac{P(n_{e-1})}{P(n_e)} \right] \left[\tilde{T}(n_{e-1}, n_{e-2}) \frac{P(n_{e-2})}{P(n_{e-1})} \right] \dots \left[\tilde{T}(n_1, n_0) \frac{P(n_0)}{P(n_1)} \right] \\ = \tilde{T}(n_e, n_{e-1}) \tilde{T}(n_{e-1}, n_{e-2}) \dots \tilde{T}(n_1, n_0) P(n_0)$$

Comparando com (*) temos que

$$\tilde{T}(n_e, n_{e-1}) = T(n_e, n_{e-1})$$

$$\tilde{T}(n_{e-1}, n_{e-2}) = T(n_{e-1}, n_{e-2}) \quad \vdots \quad \tilde{T}(n_1, n_0) = T(n_1, n_0)$$

$$T(n_{e-1}, n_e) P(n_e) = P(n_{e-1}) T(n_e, n_{e-1})$$

ou p/ quaisquer estados m e n

$$\tilde{T}(n, m) = T(n, m)$$

$$T(m, n) P(n) = T(n, m) P(m)$$

que é a condição de reversibilidade microscópica ou condição do balanço detalhado.

Calhe ressaltar que a reversibilidade microscópica é uma propriedade do processo estocástico que estamos interessado (por meio de Markov estocástico). Alguns processos estocásticos possuem reversibilidade microscópica, enquanto outros não.

Discutiremos com mais detalhes estes pontos e conceitos sobre reversibilidade microscópica futuramente, em termos de outras quantidades, dentre elas, a produção de entropia.

Em princípio é possível se há reversibilidade microscópica sem determinar a probabilidade estacionária $P(n)$. Considere três estados quaisquer n, n' e n'' . No regime estacionário, a probabilidade de ocorrer a trajetória

$n \rightarrow n' \rightarrow n'' \rightarrow n$ é dada por

$$T(n, n'') T(n'', n') T(n', n) P(n)$$

$$T(n, n'') T(n'', n') T(n', n) P(n)$$

Por outro lado, a trajetória reversa

$n \rightarrow n'' \rightarrow n' \rightarrow n$ é dada por

$$T(n, n') T(n', n'') T(n'', n) P(n)$$

Caso haja reversibilidade microscópica teremos

$$T(n, n'') T(n'', n') T(n', n) P(n) = T(n, n') T(n', n'') T(n'', n) P(n)$$

O método de Monte Carlo

O método de Monte Carlo tem sido intensamente utilizado em diversas áreas como Física Estatística, Física do Estado Sólido, Física Nuclear, Teorias de Gauge e em muitas outras.

O grande mérito do método de Monte Carlo é que nenhuma aproximação é feita no modelo que vai ser simulado; daí a sua importância.

Neste capítulo faremos uma breve descrição sobre o método de Monte Carlo, apresentando suas principais características.

3.1 Descrição geral do método

Um dos problemas básicos em Física Estatística é o cálculo de valores esperados de grandezas da forma

$$\langle O \rangle = \frac{\sum_{\{\sigma\}} O(\sigma) e^{-\beta \mathcal{H}(\sigma)}}{\sum_{\{\sigma\}} e^{-\beta \mathcal{H}(\sigma)}}. \quad (3.1)$$

onde $O(\sigma)$ é um operador, $\sum_{\{\sigma\}}$ denota a soma sobre configurações da variável σ . Por exemplo, para o modelo de Ising com N sítios, σ denota os valores da variável de spin assumida para cada sítio i da rede numa certa configuração, ou $\sigma \equiv \{\sigma_1, \sigma_2, \dots, \sigma_N\}$.

3.1 Descrição geral do método

118

Com raras exceções, não é possível obter esse valor esperado explicitamente. Uma solução alternativa seria gerar no computador todas as configurações possíveis para o modelo e a partir delas, obter o valor esperado. No modelo de Ising, por exemplo, para calcularmos explicitamente o valor esperado acima, precisaríamos gerar as 2^N configurações, o que torna o problema impraticável mesmo para N não muito grande. Uma outra alternativa seria gerar aleatoriamente um certo número M de configurações e calcular o valor esperado usando somente esse subespaço de configurações. Isso poderia dar uma estimativa ruim, uma vez que uma certa configuração σ contribui com o fator $e^{-\beta\mathcal{H}(\sigma)}$ e para a maioria das configurações este fator é pequeno.

O cálculo de valores esperados é geralmente realizado através da amostragem por importância. Ao invés de se escolher M configurações aleatoriamente, estas são escolhidas com probabilidade $e^{-\beta\mathcal{H}(\sigma)}$. Ou seja, configurações mais prováveis são mais geradas, enquanto configurações menos prováveis são menos geradas. Como a probabilidade de cada configuração está sendo levada em conta no processo de geração,

$$\langle O \rangle \approx \frac{1}{M} \sum_{i=1}^M O(\sigma_i), \quad (3.2)$$

onde o valor esperado se reduz a uma média aritmética. Métodos que utilizam números aleatórios para o cálculo das grandezas são chamados de método de Monte Carlo.

Para gerarmos configurações de acordo com a distribuição de Boltzmann, fazemos uso de um passeio aleatório no espaço de configurações, através de uma cadeia de Markov. Este processo de Markov é definido pela probabilidade de transição $W_{\sigma_n \rightarrow \sigma_m}$, onde σ_n e σ_m são pontos no espaço de configurações. Para que a cadeia de Markov que produz este passeio aleatório nos leve até a distribuição de Boltzmann, é necessário e suficiente impor as seguintes condições sobre a probabilidade de transição $W_{\sigma_n \rightarrow \sigma_m} > 0$.

- Normalização:

3.1 Descrição geral do método

$$T_{mn} = 1$$

(3.3)

- Balanço:

$$\sum_m T_{nm} e^{-\beta H_m} = e^{-\beta H_n}$$

(3.4)

- Ergodicidade: Dadas duas configurações σ_m e σ_n quaisquer que satisfazem $e^{-\beta H_m} > 0$ e $e^{-\beta H_n} > 0$, existe um número k tal que $T_{\sigma_n \rightarrow \sigma_m}^k > 0$, onde

$$T_{mn}^k = \underbrace{(T \dots T)}_{k \text{ vezes}}_{mn} \quad T_{mn}^k > 0$$

(3.5)

Dessa forma, qualquer configuração pode ser atingida a partir de uma configuração inicial qualquer. Basta aplicar o algoritmo um número suficiente de vezes.

Na construção de algoritmos, o balanço é geralmente substituído pela condição de balanço detalhado, dada abaixo

$$T_{nm} e^{-\beta H_m} = T_{mn} e^{-\beta H_n}$$

(3.6)

onde $W_{\sigma_m \rightarrow \sigma_n}$ é a probabilidade de transição do sistema ir da configuração σ_m para σ_n e $W_{\sigma_n \rightarrow \sigma_m}$ é análogo.

Balanço detalhado implica em balanço. Para vermos isto, somemos sobre m ambos os lados de (3.6)

$$\sum_m T_{nm} e^{-\beta H_m} = \sum_m T_{mn} e^{-\beta H_n}$$

(3.7)

No lado direito da equação acima, o fator $e^{-\beta H_n}$, independe de m . Usando a normalização definida em (3.3) obtemos a expressão que caracteriza o balanço.

Precisamos escolher um algoritmo que nos leve até a distribuição de Boltzmann. Em outras palavras, precisamos de um algoritmo que seja ergódico

e satisfaça o balanço detalhado.

↑

Vamos mostrar que um algoritmo que satisfaz as condições de normalização, ergodicidade e balanço detalhado nos leva a uma distribuição de Boltzmann.

Começamos pelo definindo novamente a distribuição de Boltzmann dada por

$$P_n^B = c e^{-\beta H_n} \quad \text{onde} \quad \sum_n P_n^B = 1 \quad \text{e} \quad c = \frac{1}{\sum_n e^{-\beta H_n}}$$

de forma que a distribuição acima corresponde ao "ensemble" gerado pela distribuição de Boltzmann.

A ideia aqui para demonstrarmos que aplicando o algoritmo definido pelas propriedades (3.3), (3.4) e (3.5) nos levará ao "ensemble" de Boltzmann.

Faremos isto através da expressão

$$\|C - C^B\| \equiv \sum_n |P_n^J - P_n^B|,$$

onde P_n^J corresponde ao número de vezes que a configuração n aparece J vezes em $\{x_i\}^J$ no J -ésimo passo

$$\{x_i\}^0 \xrightarrow{T} \{x_i\}^1 \xrightarrow{T} \{x_i\}^2 \rightarrow \dots \{x_i\}^{k-1} \xrightarrow{T} \{x_i\}^k$$

Dessa forma
$$P_n^J = \frac{N_n^J}{N}$$

Vamos

16

Por outro lado

$$P_n^{j+1} = \sum_m T_{nm} P_m^j.$$

Se aplicarmos o algoritmo k vezes em $\{K\}^0$, temos que

$$P_n^k = \sum_{m_1, m_2, \dots, m_{k-1}} T_{nm_{k-1}} T_{m_{k-1} m_{k-2}} \dots T_{m_1 m} P_m^0 = \sum_m T_{nm}^k P_m^0 \quad (1)$$

Com essas conceitos é possível mostrarmos que

$$\|C^k - C^B\| < \|C^0 - C^B\|.$$

$$\begin{aligned} \|C^k - C^B\| &\equiv \sum_n |P_n^k - P_n^B| = \sum_n \left| \sum_m \underbrace{T_{nm}^k}_{P_n^k} \underbrace{P_m^0}_{\text{Da propriedade 3.4}} - \sum_m \underbrace{T_{nm}}_{\sum_n P_n^B} P_m^B \right| \\ &= \sum_n \left| \sum_m T_{nm}^k (P_m^0 - P_m^B) \right|. \end{aligned}$$

Usando a desigualdade triangular

($\|u+v\| \leq \|u\| + \|v\|$) temos

$$\|C^k - C^B\| \leq \|C^k\| + \|C^B\| + \sum_n \sum_m |T_{nm}^k (P_m^0 - P_m^B)|$$

Assum

$$\therefore \sum_n \left| \sum_m T_{nm}^k (P_m^0 - P_m^B) \right| \leq \sum_n \sum_m |T_{nm}^k (P_m^0 - P_m^B)|$$

Usando a propriedade $\sum_n T_{nm}^k = 1$, temos $\sum_m \sum_n T_{nm}^k |p_m^0 - p_m^B| = \sum_m |p_m^0 - p_m^B| = |c^0 - c^B|$

~~Alguns comentários são necessários:~~

- O sinal $<$ vai valer se alguns termos da soma em $\sum_m$ forem negativos. Se todos os termos em $\sum_m$ forem positivos ou nulos então vale a igualdade $=$.

Como as probabilidades são normalizadas

$$\sum_m p_m^0 = \sum_m p_m^B = 1, \text{ nesse caso } \sum_m (p_m^0 - p_m^B) = 0.$$

↓ Isso seria correto apenas se p_m^0 for igual a p_m^B

Se o "ensemble" inicial não for como $T_{nm}^k > 0$ (ergodicidade) ensemble de Boltzmann, então alguns termos $p_m^0 - p_m^B$ serão positivos e outros serão negativos. Ou seja,

$$\|c^k - c^B\| < \|c^0 - c^B\|$$

ao aplicarmos o algoritmo um certo número de vezes. Portanto, realmente

nos aproximamos da distribuição de Boltzmann ao aplicarmos o algoritmo com as propriedades (3.9) - (3.5).

Aula 2.1 Equação Mestra

Considere uma variável estocástica x_t que assume valores inteiros e t seja o instante de tempo que assume valores também inteiros $0, 1, 2, \dots$. Um processo estocástico é dito markoviano quando o valor da variável estocástica num determinado instante de tempo só depende do seu valor no instante de tempo anterior. Seja $P_{\ell+1}(n_{\ell+1}|n_0, n_1, n_2, \dots, n_\ell)$ a probabilidade condicional de que a variável x_t assumo o valor $n_{\ell+1}$ no instante de tempo $t_{\ell+1}$ dado que ela tenha tomado o valor n_0 no instante t_0 , o valor n_1 no instante t_1 , o valor n_2 no instante $t_2, \dots$, o valor n_ℓ no instante t_ℓ . Num processo estocástico essa probabilidade é igual à probabilidade condicional $P_{\ell+1}(n_{\ell+1}|n_\ell)$ de que a variável x_t assumo o valor $n_{\ell+1}$ no instante de tempo $t_\ell + 1$ dado que ela tenha tomado o valor n_ℓ no instante t_ℓ . Pode-se mostrar [5], que num processo markoviano a probabilidade $P_{\ell+1}(n_{\ell+1})$ de que a variável x_t tome o valor $n \equiv n_{\ell+1}$ no instante $t_{\ell+1}$ ela é dada por

$$P_\ell(n) = \sum_m T(n, m) P_{\ell-1}(m), \quad (2.1)$$

(2)

onde $T(n, m) \equiv P_t(n_t | n_{t-1})$. Isto nos permite interpretar $T(n, m)$ como um elemento de uma matriz estocástica T , que deve possuir as propriedades

$$T(n, m) \geq 0, \quad (2.2)$$

pois $T(n, m)$ é uma probabilidade condicional e

$$\sum_n T(n, m) = 1, \quad (2.3)$$

pois a soma das probabilidades de transição de um estado m qualquer para qualquer estado possível n deve ser 1. Ou seja, os elementos da matriz T devem ser não negativos e a soma dos elementos de uma coluna qualquer deve ser 1. Vamos supor agora que as transições ocorram a cada intervalo de tempo τ e que a matriz estocástica seja dada por

$$T(n, m) = \tau W(n, m) \quad n \neq m, \quad (2.4) \quad \checkmark$$

e

$$T(n, n) = 1 - \tau \Omega(n). \quad (2.5) \quad \checkmark$$

A segunda equação pode ser justificada pelo fato de que num intervalo de tempo τ muito pequeno a probabilidade de permanecer no mesmo estado $m = n$ deve ser próxima de 1. Utilizando a propriedade (2.3) e as equações acima, temos que

$$\Omega(n) = \sum_{m(\neq n)} W(m, n).$$

$$\sum_{m \neq n} T(m, n) + T(n, n) = 1$$

$$\Leftrightarrow \sum_{m \neq n} \tau W(m, n) + 1 - \tau \Omega(n) = 1$$

$$\Leftrightarrow \tau \sum_{m \neq n} W(m, n) = \tau \Omega(n)$$

Substituindo as equações (2.4) e (2.5) na equação (2.1), temos

$$P_\ell(n) = \tau \sum_{m \neq n} W(n, m) P_{\ell-1}(m) + T(n, n) P_{\ell-1}(n) \quad (2.6)$$

$$P_\ell(n) = \tau \sum_{m(\neq n)} W(n, m) P_{\ell-1}(m) + P_{\ell-1}(n) - \tau \sum_{m(\neq n)} W(m, n) P_{\ell-1}(n). \quad (2.7)$$

$$\Omega(n) = \sum_{m \neq n} W(m, n)$$

Escrevendo o instante de tempo da seguinte forma $(\ell - 1)\tau = t$ de modo que a probabilidade $P_\ell(n)$ possa ser escrita como $P(n, t)$, temos que

$$\frac{P(n, t + \tau) - P(n, t)}{\tau} = \sum_{m(\neq n)} \{W(n, m)P(m, t) - W(m, n)P(n, t)\}. \quad (2.8)$$

$$= \tau \sum_{m \neq n} W(n, m) P_{\ell-1}(m) + \left[1 - \tau \sum_{m \neq n} W(m, n) \right] P_{\ell-1}(n)$$

$$P_\ell(n) = \tau \sum_{m \neq n} W(n, m) P_{\ell-1}(m) - W(m, n) P_{\ell-1}(n) + P_{\ell-1}(n)$$



Ao tomarmos o limite $\tau \rightarrow 0$, temos a seguinte equação

$$\frac{d}{dt}P(n, t) = \sum_{m(\neq n)} \{W(n, m)P(m, t) - W(m, n)P(n, t)\}, \quad (2.9)$$

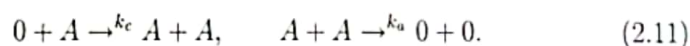
que chamada de equação mestra. A equação mestra descreve a evolução temporal da probabilidade do sistema estar num estado qualquer num determinado instante. O primeiro termo descreve todas as configurações m que dão origem as configurações n e daí o sinal positivo ("termo de ganho"), enquanto o segundo termo descreve as novas configurações m que serão originadas de n , daí o sinal negativo ("termo de perda"). Podemos utilizar uma outra simbologia para as taxas de transição acima, que evidencia melhor o caráter de "ganho" e de "perda" de cada um dos termos da equação mestra e é dada por

$$\frac{d}{dt}P(n, t) = \sum_{m(\neq n)} \{W_{m \rightarrow n}P(m, t) - W_{n \rightarrow m}P(n, t)\}, \quad (2.10)$$

2.1.1 Dois exemplos

Vamos considerar dois exemplos de sistemas para exemplificar a aplicabilidade da equação mestra.

Como primeiro exemplo, consideramos um processo zero-dimensional definido pelas reações químicas



onde k_c e k_a são as taxas de criação e aniquilação de partículas, respectivamente. Nesse exemplo há apenas um tipo de partícula, que chamamos de A . A evolução temporal da probabilidade é dada por

$$\frac{d}{dt}P(n, t) = W_{n-1 \rightarrow n}P(n-1, t) - W_{n \rightarrow n+1}P(n, t) + W_{n+2 \rightarrow n}P(n+2, t) - W_{n \rightarrow n-2}P(n, t), \quad (2.12)$$

onde os estados n e m correspondem ao número de partículas do sistema. As taxas de criação e aniquilação devem ser proporcionais ao número de

partículas nele existente de modo que

$$W_{n-1 \rightarrow n} = k_c(n-1), \quad (2.13)$$

e

$$W_{n \rightarrow n-2} = k_a n(n-1), \quad (2.14)$$

respectivamente, implicando a seguinte equação mestra

$$\begin{aligned} \frac{d}{dt}P(n, t) = & k_c(n-1)P(n-1, t) + k_a(n+2)(n+1)P(n+2, t) + \\ & -(k_c n + k_a n(n-1))P(n, t), \end{aligned} \quad (2.15)$$

que é o ponto de partida de uma formulação contínua para a percolação direcionada [24, 25, 26].

O segundo exemplo consiste em escrever a equação mestra para o processo de contato [9]. O processo de contato usual, introduzido por Harris [9], é um modelo em que as partículas são criadas cataliticamente e aniquiladas espontaneamente. A cada sítio i da rede é atribuída uma variável de ocupação η_i que toma os valores $\eta_i = 0, 1$ quando o sítio estiver vazio ou ocupado por uma partícula, respectivamente e múltiplas ocupações são proibidas. Conforme deduzimos na seção anterior, a evolução temporal da probabilidade de ocorrência da configuração $\eta \equiv (\eta_1, \eta_2, \dots, \eta_i, \dots, \eta_V)$ no instante de tempo t é governada pela equação mestra

$$\frac{d}{dt}P(\eta, t) = \sum_{\eta' (\neq \eta)} \{W_{\eta' \rightarrow \eta} P(\eta', t) - W_{\eta \rightarrow \eta'} P(\eta, t)\}, \quad (2.16)$$

onde V é o número total de sítios da rede. Consideraremos aqui o caso em que as transições ocorram apenas entre configurações que sejam diferentes apenas pelo estado de um único sítio. Escrevemos portanto

$$W_{\eta \rightarrow \eta'} = \sum_{i=1}^V \delta_{\eta_1, \eta'_1} \delta_{\eta_2, \eta'_2} \dots \delta_{\eta_i, 1-\eta'_i} \dots \delta_{\eta_V, \eta'_V} w_i(\eta), \quad (2.17)$$

onde $\delta_{\eta_i, \eta'_i} = 0, 1$ conforme η_i e η'_i sejam iguais ou diferentes, respectivamente e $w_i(\eta)$ é a taxa de transição do estado do i -ésimo sítio de η_i para $1 - \eta_i$.

5

Substituindo a expressão acima na equação (2.16) temos que

$$\frac{d}{dt}P(\eta, t) = \sum_{i=1}^V \{w_i(\eta^i)P(\eta^i, t) - w_i(\eta)P(\eta, t)\}, \quad (2.18)$$

onde $\eta^i \equiv (\eta_1, \eta_2, \dots, 1 - \eta_i, \dots, \eta_V)$.

Vamos considerar a evolução temporal da média $\langle f(\eta) \rangle$ de uma função $f(\eta)$ de estado definida por

$$\langle f(\eta) \rangle = \sum_{\eta} f(\eta)P(\eta, t). \quad (2.19)$$

Multiplicando ambos os lados da equação acima por $f(\eta)$ e somando sobre η temos que

$$\frac{d}{dt} \sum_{\eta} f(\eta)P(\eta, t) = \sum_{\eta} \sum_{i=1}^V f(\eta)w_i(\eta^i)P(\eta^i, t) - \sum_{\eta} \sum_{i=1}^V f(\eta)w_i(\eta)P(\eta, t), \quad (2.20)$$

que resulta na seguinte equação

$$\frac{d}{dt} \langle f(\eta) \rangle = \langle \{f(\eta^i) - f(\eta)\}w_i(\eta) \rangle, \quad (2.21)$$

onde fizemos a transformação $\eta \rightarrow \eta^i$ na primeira parcela depois da igualdade. Por simplicidade, vamos considerar o processo de contato unidimensional. A taxa de transição nesse caso é dada por

$$w_i(\eta) = \frac{1}{2}(1 - \eta_i)(\eta_{i+1} + \eta_{i-1}) + \alpha\eta_i. \quad (2.22)$$

A equação de evolução para $f(\eta) = \eta_i$ é obtida a partir da expressão (2.21) e é dada por

$$\frac{d}{dt} \langle \eta_i \rangle = \langle (1 - \eta_i)\eta_{i+1} \rangle - \alpha \langle \eta_i \rangle. \quad (2.23)$$

Podemos utilizar ainda uma outra formulação para a equação acima. A densidade de partículas ρ , dada pela média acima pode ser escrita da seguinte forma $\langle \eta_i \rangle = \sum_{\eta} \eta_i P(\eta) = P(1)$. Da mesma maneira temos que $\langle \eta_i \eta_{i+1} \rangle = \sum_{\eta} \sum_{\eta_{i+1}} \eta_i \eta_{i+1} P(\eta_i, \eta_{i+1}) = P(11)$. Temos então a seguinte expressão para a equação (2.23)

$$\frac{d}{dt}P(1) = P(01) - \alpha P(1). \quad (2.24)$$

.

de modo que as soluções para a equação (2.24) são

$$P(1) = 0, \quad (2.26)$$

correspondendo ao estado ausente de partículas (estado absorvente) e

$$P(0) = \alpha, \quad (2.27)$$

que corresponde à um estado em que partículas são criadas e aniquiladas, denominado de estado ativo. Dessa forma, a densidade de partículas $\rho \equiv \langle \eta_i \rangle$ no estado estacionário satisfaz a equação

$$\rho = 1 - \alpha. \quad (2.28)$$

A equação acima descreve uma transição de fase entre um estado ativo e absorvente em $\alpha_c = 1$, cujo expoente crítico associado β é dado por

$$\rho = (\alpha_c - \alpha)^\beta, \quad (2.29)$$

sendo $\beta = 1$. Resultados provenientes de simulações numéricas [4] e expansões em série perturbativa [13] mostram que uma transição de fases contínua entre um estado ativo e absorvente ocorre em $\alpha_c = 0.303228(2)$ cujo expoente crítico β é dado por $\beta = 0.27649\dots$, o que discorda com os valores obtidos aqui.

(8)

Aplicação 1 $\Rightarrow$ Reação química (simplex)



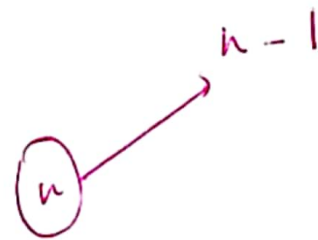
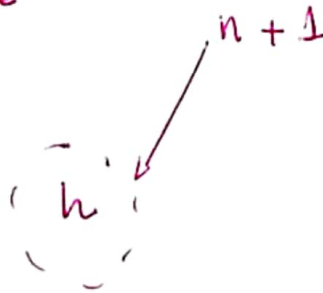
X se transforma na molécula Y com uma taxa k . Condys Inicial $P_n(0) = \delta_{n, N_0}$

↑
número inicial
de moléculas
de tipo X

Equação Mestre

$P_n(t) \rightarrow$ probabilidade de encontrar n moléculas de tipo X no instante t . Além disso, vamos supor

$(W(n_1, n_2) = kn_2) \rightarrow$



$$\frac{d}{dt} P(n, t) = k(n+1) P(n+1, t) - kn P(n, t)$$

(Truque) Multiplicando por n ambos os lados e somando $\sum_{n=0}^{\infty}$, temos

$$\begin{aligned} \frac{d}{dt} \sum_{n=0}^{\infty} n P(n, t) &= k \sum_{n=0}^{\infty} n(n+1) P(n+1, t) - k \sum_{n=0}^{\infty} n^2 P(n, t) \\ &= \sum_{n=0}^{\infty} n(n+1) P(n+1, t) - k \sum_{n=0}^{\infty} n^2 P(n, t) \\ &= k \sum_{n'=1}^{\infty} (n'-1)n' P(n', t) - k \sum_{n=0}^{\infty} n^2 P(n, t) \end{aligned}$$

$$= K \sum_{n'=1}^{\infty} n' (n'-1) P(n', t) - K \sum_{n=1}^{\infty} n^2 P(n, t)$$

(9)

$$= -K \sum_{n=1}^{\infty} n' P(n', t) = -K \sum_{n=0}^{\infty} n P(n, t)$$

$$\frac{d}{dt} \underbrace{\sum_{n=0}^{\infty} n P(n, t)}_{\langle N(t) \rangle} = -K \underbrace{\sum_{n=0}^{\infty} n' P(n, t)}_{\langle N(t) \rangle}$$

$$\begin{aligned} \langle N(t) \rangle &= N(0) e^{-kt} \\ &= N_0 e^{-kt} \end{aligned}$$

Por fim, a equação (9) é:

Em outras palavras, dada a

equação, $\frac{d}{dt} P_n = \sum_m W_{nm} P_m$ ou

ainda $\frac{dP}{dt} = W P(t)$, matriz coluna cujos elementos são $P_n(t)$,

a solução da equação acima dada a
condição inicial $P_n(0)$ é dada por

$$P(t) = e^{tW} P(0)$$

O vetor P_e , correspondente à distribuição de
probabilidades estacionárias $P_e(n)$ é solução da
equação $W P_e = 0$.

Dada as taxas de transição, isto é, dada
a matriz de evolução W , desejamos determinar
quais as condições que devemos impor sobre
 W para que a solução estacionária seja
única e para que

$$\lim_{t \rightarrow \infty} P(t) \rightarrow P_e,$$

isto é, para que $P(t)$ se aproxime da solução
estacionária para tempos longos. Com esta
finalidade, consideramos novamente a relação

$P(t) = T^t P(0)$ discutida no capítulo anterior
no caso de cadeias de Markov (tempo discreto)

e discretizamos o tempo em intervalos
iguais a Δt , de forma que $t = l \cdot \Delta t$

Assum: $P(t) = T^{t/\Delta t} P(0)$.

Definimos ainda a matriz $T = I + \Delta t \cdot W$.

~~Notar~~ cp

$$P(t) = (I + \Delta t W)^{t/\Delta t} P(0)$$

Note que
 $\Delta t \rightarrow 0 \quad T \rightarrow I$,

Lembrando que $\lim_{\Delta t \rightarrow 0} (1 + k\Delta t)^{t/\Delta t} = e^{kt}$,

$$P(t) = e^{Wt} P(0)$$

para $\Delta t \rightarrow 0$.
//

Conforme vimos, se a matriz T satisfizer os
requisitos do teorema de Perron-Frobenius,
a solução estacionária será única e ela atin-
girá no limite $t \rightarrow \infty$ para qualquer condição
inicial. No caso da matriz W vale as afirmações
abaixo:

- 1) a matriz de evolução W possui um autovalor
nulo
- 2) a parte real de qualquer autovalor de W é
negativa ou nula.

Questão
discussão

1) O autovalor nulo corresponde um autovetor com componentes não negativos

Se qualquer estado puder ser atingido a partir de qualquer outro, então T é uma matriz **irreduzível**. Sendo **irreduzível** é também regular, ou seja que sempre possui elementos na diagonal ($T = I + \Delta t W$ e estamos considerando Δt pequeno). Nessa forma, podemos também fazer as seguintes afirmações sobre a matriz W

4) O autovalor nulo é não degenerado e o autovetor correspondente possui todos os componentes estritamente positivos. Em outras palavras, o estado estacionário é único e $P_{ei} > 0$.

5) Os demais autovetores possuem parte real negativa.

É de forma análoga ao que fizemos no caso de processos markovianos, cuja matriz de transição é regular e irreductível, ~~obtemos~~ aqui a solução da equação mestre por meio dos autovalores e autovetores da matriz W

seja

$$W | \psi_k \rangle = \lambda_k | \psi_k \rangle \text{ e}$$

$$\langle \phi_k | W = \lambda_k \langle \phi_k | \text{ onde}$$

$\{ \langle \phi_k | \}$ e $\{ | \psi_k \rangle \}$ devem formar um conjunto ortogonalizado de autovetores,

segue $| \psi_k \rangle \langle \phi_j | = \delta_{jk} \text{ e}$

$$\sum_j | \psi_j \rangle \langle \phi_j | = I.$$

Novamente, o vetor probabilidade estacionária P_e é um autovetor com autovalor nulo

$$W P_e = 0 \text{ onde}$$

$$W | \psi_0 \rangle = 0 | \psi_0 \rangle, \text{ ou seja } P_e \rightarrow | \psi_0 \rangle.$$

Além disso, o vetor $\langle \phi_0 |$ é uma linha em que todos os componentes são iguais a unidade.

$$\text{senão } \langle \phi_0 | = (1 \dots 1)$$

$$\text{e } \langle \phi_0 | W = (1 \dots 1) \begin{pmatrix} - \\ - \\ - \\ - \\ - \end{pmatrix}$$

Como a soma ao longo de uma coluna da matriz W é sempre 0, segue que $\langle \phi_0 | W = 0$.

$$\text{se } \langle \phi_0 | = (1 \dots 1).$$

Dessa forma

$$\begin{aligned} e^{tW} &= e^{tW} \sum_k |\psi_k\rangle \langle \phi_k| \\ &= \sum_k e^{t\lambda_k} |\psi_k\rangle \langle \phi_k|. \end{aligned}$$

$$P(t) = e^{tW} P(0)$$

$$= \sum_k e^{t\lambda_k} |\psi_k\rangle \langle \phi_k | \psi(0) \rangle$$

$\uparrow$
 $P(0)$

$$\begin{aligned} &= e^{t\lambda_0} |\psi_0\rangle \langle \phi_0 | \psi(0) \rangle + \sum_{k \neq 0} e^{t\lambda_k} |\psi_k\rangle \langle \phi_k | \psi(0) \rangle \\ &= A |\psi_0\rangle + \sum_{k \neq 0} e^{t\lambda_k} |\psi_k\rangle \langle \phi_k | \psi(0) \rangle \end{aligned}$$

Como $\lambda_k < 0$ para todo $k \neq 0$.

16-2

$P(t) \longrightarrow A P_e$ qdo $t \rightarrow \infty$.

Exercício $\Rightarrow$ Encontre o vetor ~~esta~~ probabilidade estacionária da matriz de evolução

$$W = \begin{pmatrix} -a & b \\ a & -b \end{pmatrix}.$$