

Łańcuchy Markowa i procesy stochastyczne w Monte Carlo

plan wykładu

- łańcuch Markowa dla zmiennej dyskretnej
- błądzenie przypadkowe
- proces stochastyczny
- stacjonarność, jednorodność i ergodyczność procesu Markowa
- dynamika w MC: równanie typu Master
- algorytm Gillespie dla równania typu Master
- ruchy Browna
- równanie dyfuzji
- procesy Wienera i Orsteina-Uhlenbecka

literatura:

- D. Gamerman, H.F. Lopes, „Markov chain Monte Carlo – stochastic simulation for bayesian inference”
- M.H. Kalos, P.A. Whitlock, „Monte Carlo methods”
- C.W. Gardiner, „Handbook of stochastic methods for physics, chemistry and natural sciences”
- D. S. Lemons „An introduction to stochastic processes in physics”
- F. Petruccione, P. Biechele, „Stochastic methods for physics using Java: an introduction”
- D. T. Gillespie, J. Phys. Chem. 81, 2350 (1977)

Łańcuchy Markowa

W 1906 roku Andrei Markow badał częstość pojawiania się samogłosek i spółgłosek w poemacie „Oniegin” Puszkina. Opracował model stochastyczny, w którym pojawienie się kolejnej samogłoski było zależne wyłącznie od litery ją poprzedzającej. Model poprawnie odtwarzał częstość występowania samogłosek – nazwa pochodzi oczywiście od niego. Równocześnie Henri Poincare badał własności ciągów zmiennych losowych, które później okazały się łańcuchami Markowa. W MC łańcuchy Markowa zastosował Metropolis (1953) a od czasu publikacji Hastingsa (1970) liczba zastosowań gwałtownie rośnie.

łańcuch Markowa to rodzaj **procesu stochastycznego**, w którym ciąg zmiennych losowych generowany jest według pewnego „wzorca”, cechą charakterystyczną jest dążenie do uzyskania stacjonarnego rozkładu granicznego.

proces stochastyczny najogólniej definiujemy metodę generowania ciągu wielkości (np. liczb) losowych

$$\{x_t : t \in R\}$$

$$x_t \in S \subset R^d \quad \begin{array}{l} \text{- } S \text{ to przestrzeń dostępnych stanów układu, } d\text{-wymiarowa} \\ t \quad \text{- parametr (przeliczalny) numerujący stany} \end{array}$$

Podstawowa własność łańcucha Markowa – aktualny stan układu zależy wyłącznie od stanu poprzedniego, i nie zależy od wcześniejszej historii (stanów wcześniejszych).

Warunek ten można sformułować posługując się pojęciem prawdopodobieństwa warunkowego

$$P\{x_{n+1}|x_n, x_{n-1}, \dots, x_0\} = P\{x_{n+1}|x_n\}$$

$$x_n \in A^n \subset S \quad \begin{array}{l} \text{- zmienne losowe generowane są w różnych podzbiorach zawartych w } S \\ \text{np. bezpośrednie przejście z } x_0 \text{ do } x_n \text{ może nie być dozwolone, ale w } n\text{-krokach już tak,} \\ \text{przejście pomiędzy odległymi stanami musi trwać – jest więc pewnym procesem (losowym)} \end{array}$$

$$x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow \dots \rightarrow \underbrace{x_n \rightarrow x_{n+1}}_{\text{korelacja}}$$

W ogólnym przypadku prawdopodobieństwo wylosowania nowego stanu może zależeć od: $\{x, A, n\}$,
Jeśli nie zależy od n to **łańcuch jest jednorodny** – możemy określić
tzw. **prawdopodobieństwo przejścia** $T(x,y)$ o własnościach

(uwaga: rozważamy stany dyskretne, łańcuch dla rozkładu ciągłego podany będzie później)

$$T(x, y) = P\{y|x\} \geq 0, \quad \forall x, y \in S$$

$$\sum_y T(x, y) = 1, \quad x, y \in S$$

- z pierwszej własności wynika, że przejście pomiędzy pewnymi stanami może nie być realizowane w sposób bezpośredni tj. w pojedynczym kroku, ale z wykorzystaniem stanów pośrednich
- druga własność to warunek normalizacji, przejście z x do dowolnego stanu w S realizujemy z prawdopodobieństwem równym 1.

Błądzenie przypadkowe jako przykład łańcucha Markowa

Definiujemy przestrzeń stanów np. w postaci zbioru liczb całkowitych

$$S = \{\dots, -2, -1, 0, 1, 2, \dots\}$$

Błądzenie przypadkowe polega na losowym przechodzeniu układu z aktualnego stanu x_n do jednego ze stanów w jego **najbliższym sąsiedztwie** z określonym prawdopodobieństwem

$$\tilde{x} \in \{x_{n-1}, x_n, x_{n+1}\}$$

W najprostszej wersji prawdopodobieństwo przejścia możemy zdefiniować jak poniżej

$$T(x_n, x_{n+1}) = \begin{cases} p & \iff x_{n+1} = x_n + 1 \\ q & \iff x_{n+1} = x_n - 1 \\ z & \iff x_{n+1} = x_n \\ 0 & \iff x_{n+1} \notin \{x_n - 1, x_n, x_n + 1\} \end{cases}$$

uwzględniamy warunek normalizacji

$$\sum_{i=-1}^1 T(x_n, x_{n+i}) = p + q + z = 1$$

przesunięcie lewo-prawo (Δ) to zmienna losowa, więc położenie w n -tym kroku też jest zmienną losową

$$x_n = x_{n-1} + \Delta_n = \Delta_n + \Delta_{n-1} + \Delta_{n-2} + \dots + \Delta_1 + x_0$$

zatem ciąg wartości x_n jest łańcuchem Markowa

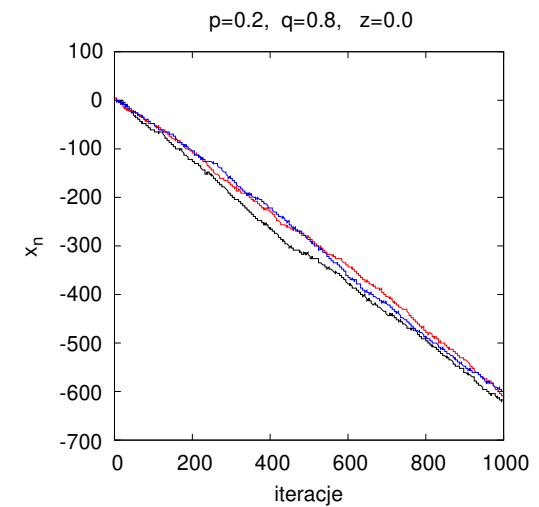
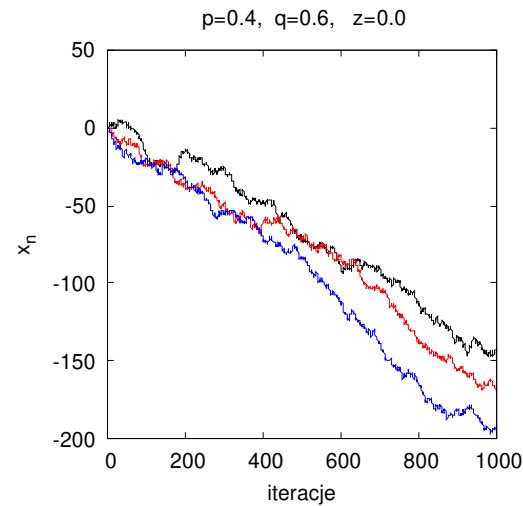
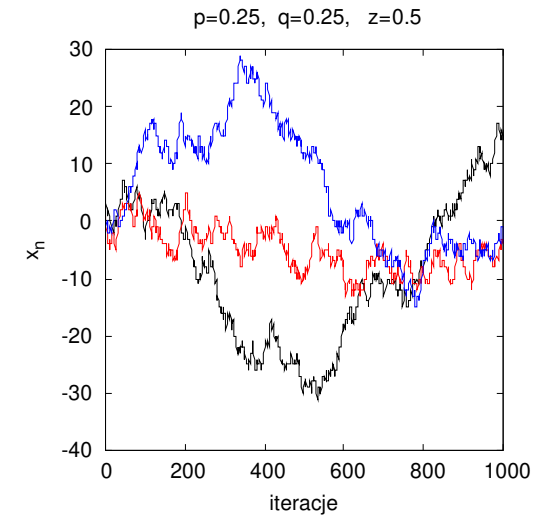
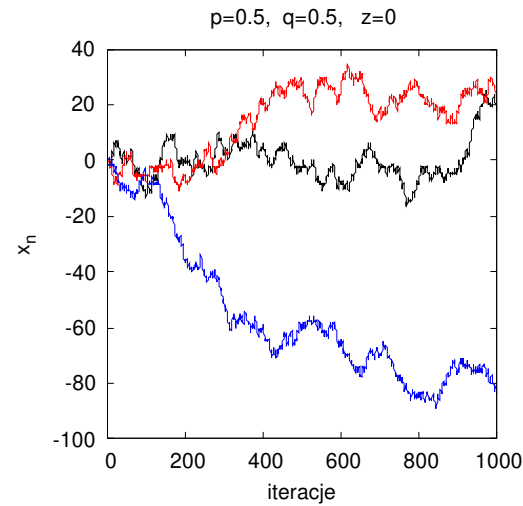
$$\{x_n : n \in N\} \in S$$

Błądzenie przypadkowe nazywane jest też ruchem pijanego marynarza.

pseudokod błędzenia przypadkowego

```

inicjalizacja:  $x_0$ 
do i=1 to n
   $U_1 \sim U(0,1)$ 
  if ( $U_1 < p$ ) then
     $x = x + 1$ 
  elseif ( $U_1 < (p + q)$ ) then
     $x = x - 1$ 
  else
     $x = x$ 
  end if
end do
    
```



W przypadku, gdy S jest przestrzenią skończoną stanów dyskretnych możemy prawdopodobieństwo przejścia sformułować w postaci **macierzy losowej P**

$$S = \{x_1, x_2, \dots, x_r\}$$

$$\mathbf{T} = \begin{bmatrix} T(x_1, x_1) & T(x_1, x_2) & \dots & T(x_1, x_r) \\ T(x_2, x_1) & T(x_2, x_2) & \dots & T(x_2, x_r) \\ \vdots & \vdots & \ddots & \vdots \\ T(x_r, x_1) & T(x_r, x_2) & \dots & T(x_r, x_r) \end{bmatrix}$$

warunek normalizacji

$$\sum_{j=1}^r T(x_i, x_j) = 1$$

warto zapamiętać własności macierzy losowych:

- suma elementów w wierszu jest równa 1 (warunek normalizacji)
- iloczyn macierzy losowych też jest macierzą losową
- przynajmniej jedna wartość własna macierzy równa jest 1

Jak możemy wykorzystać formę macierzową?

Policzmy prawdopodobieństwo przejścia pomiędzy stanami x_A i x_B dokładnie w m -krokach (tj. poprzez m stanów pośrednich)

X_k to zmienna losowa (stan):

$$X_k \in \{x_1, x_2, \dots, x_r\}$$

$$\begin{aligned} T^m(x_A, x_B) &= P\{X_{m+1} = x_B | X_0 = x_A\} \\ &= \sum_{X_1} \dots \sum_{X_m} P\{X_{m+1} = x_B, X_m, \dots, X_1 | X_0 = x_A\} \\ &= \sum_{X_1} \dots \sum_{X_m} P\{X_{m+1} = x_B | X_m\} \dots P\{X_1 | X_0 = x_A\} \\ &= \sum_{X_1} \dots \sum_{X_m} T(x_A, \mathbf{X}_1) T(\mathbf{X}_1, \mathbf{X}_2) T(\mathbf{X}_2, \mathbf{X}_3) \dots T(\mathbf{X}_m, x_B) \end{aligned}$$

co możemy zapisać w postaci macierzowej

$$T^m(x_A, x_B) = [\mathbf{T}^m]_{x_A, x_B}$$

Wynik można uogólnić (równanie **Chapmana-Kołmogorowa**)

$$T^{n+m}(x_a, x_B) = \sum_{x_i} P_m\{x_B|x_i\}P_n\{x_i|x_A\} = \sum_{x_i} T^n(x_A, x_i)T^m(x_i, x_B)$$

$$\mathbf{T}^{n+m} = \mathbf{T}^n \mathbf{T}^m$$

Co nam to daje?

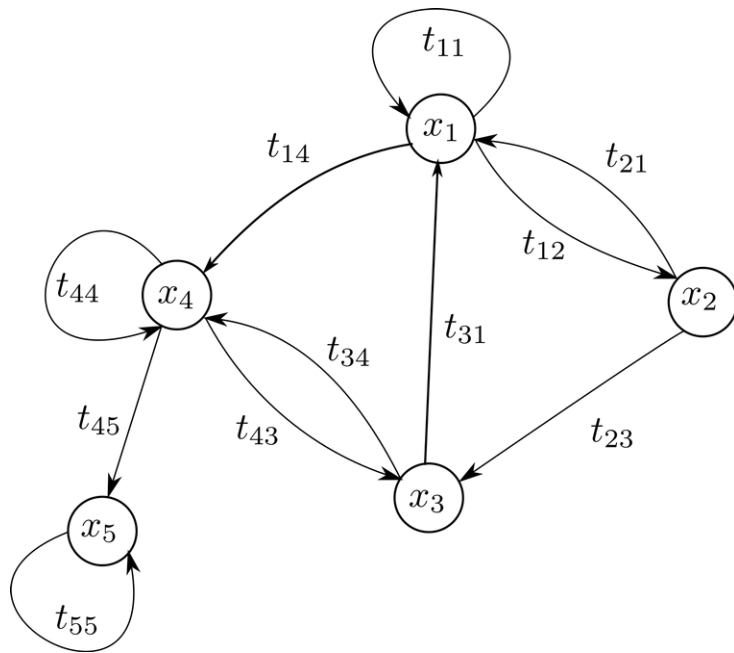
Powiedzmy, że chcemy wykonać $m=100$ kroków – ile razy musimy dokonać mnożenia macierzy?

$$\mathbf{T}^2 = \mathbf{T}\mathbf{T}, \quad \mathbf{T}^4 = \mathbf{T}^2\mathbf{T}^2, \quad \mathbf{T}^8 = \mathbf{T}^4\mathbf{T}^4, \quad \mathbf{T}^{16} = \mathbf{T}^8\mathbf{T}^8, \quad \mathbf{T}^{32} = \mathbf{T}^{16}\mathbf{T}^{16}, \quad \mathbf{T}^{64} = \mathbf{T}^{32}\mathbf{T}^{32}$$

$$\mathbf{T}^{100} = \mathbf{T}^{64}\mathbf{T}^{32}\mathbf{T}^4$$

wystarczy 8 operacji zamiast 99, więc sformułowanie macierzowe jest opłacalne – ułatwia rachunki

Błądzenie przypadkowe na grafie – macierz przejścia



Strzałki na grafie pokazują, dozwolone przejścia pomiędzy stanami – możemy skonstruować macierz pojedynczych przejść ($n=1$).

$$T = \begin{bmatrix} t_{11} & t_{12} & 0 & t_{14} & 0 \\ t_{21} & 0 & t_{23} & 0 & 0 \\ t_{31} & 0 & 0 & t_{34} & 0 \\ 0 & 0 & t_{43} & t_{44} & t_{45} \\ 0 & 0 & 0 & 0 & t_{55} \end{bmatrix}$$

Wnioski dotyczące pojedynczych przejść ($n=1$):

- z jednego stanu możemy przejść do kilku innych, ale nie do wszystkich
- w pewnych przypadkach istnieje skończone prawdopodobieństwo, że możemy pozostać w danym stanie
- jeśli dotrzemy do stanu x_5 to już w nim pozostaniemy (stan absorpcyjny)

Przejścia n -krokowe

- z każdego stanu x_1 - x_4 możemy dostać się do stanu x_5 , ze stanu x_5 nie przejdziemy do innych
- stany x_1 i x_3 nie są połączone bezpośrednio, ale możemy zrealizować przejść w postaci: x_1 - x_4 - x_3 , lub x_1 - x_2 - x_3 itd.
- dla stanów x_1, x_2, x_3, x_4 istnieje skończone prawdopodobieństwo, że startując z każdego z nich, powrócimy do danego stanu

Klasyfikacja stanów

- stany komunikujące się:
startując z x_k dotrzemy do stanu x_m w skończonej liczbie n kroków
- stan rekurencyjny (ang. *recurrent state*):
startując z danego stanu, na pewno do niego powrócimy w przyszłości (po n krokach, $n=1,2,3..$)
- stan przejściowy (ang. *transient*):
prawdopodobieństwo powrotu jest mniejsze od 1

$$X_i \in \{x_1, x_2, x_3, x_4, x_5\}$$

$$t_{km}^{(n)} = P\{X_n = x_m, n \geq 1 | X_0 = x_k\} > 0$$

$$t_{kk}^{(n)} = P\{X_n = x_k, n \geq 1 | X_0 = x_k\} = 1$$

$$0 < t_{kk}^{(n)} = P\{X_n = x_k, n \geq 1 | X_0 = x_k\} < 1$$

Klasyfikacja łańcuchów Markowa

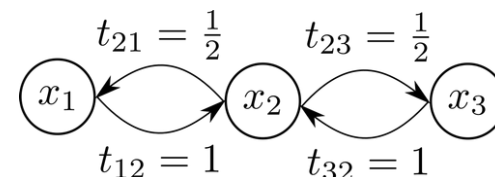
- łańcuch jest **nieredukowalny** jeśli z dowolnego stanu x_k możemy przejść do każdego stanu x_m

$$0 < P\{X_n = x_k, n \geq 1 | X_0 = x_k\} < 1$$

- łańcuch jest **periodyczny** jeśli startując z x_k powracamy do tego stanu zawsze po tej samej liczbie kroków

$$\bigwedge_{x_k \in S} P\{X_n = x_k, n \geq 1 | X_0 = x_k\} = 1$$

przykład
- łańcuch periodyczny



Jak sprawdzić czy łańcuch jest redukowalny? – najprościej jest policzyć T^n , elementy macierzowe odpowiadające przejściom ze stanów rekurencyjnych do pozostałych są zerami.

Jeśli łańcuch Markowa natrafi na podciąg stanów rekurencyjnych już nigdy go nie opuści (**absorpcja** – co oczywiście wykorzystujemy)

Przykład - łańcuch nieredukowalny

$$T = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \implies T^{(n)} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

Przykład – łańcuch redukowalny

- stany przejściowe:
 $\{x_2, x_3\}$
- dwa podciągi stanów rekurencyjnych:
 $\{x_1\}$, $\{x_4, x_5, x_6\}$
- wejście do podciągu rekurencyjnego, zatrzymuje w nim łańcuch, już się stamtąd nie wydostanie

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & 0 & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{3}{4} \end{bmatrix} \implies T^{(n)} = \begin{bmatrix} * & 0 & 0 & 0 & 0 & 0 \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * & * \end{bmatrix}$$

Rozkład prawdopodobieństwa realizacji stanów dyskretnych w przestrzeni S w n -tym kroku określa **wektor** π^n

$$\vec{\pi}^{(n)} = [p_1^{(n)}, p_2^{(n)}, \dots, p_r^{(n)}]$$

warunek normalizacji

$$\sum_{i=1}^n p_i^{(n)} = 1$$

Dla $n=0$ definiuje on **rozkład początkowy (startowy)**

$$\vec{\pi}^{(0)} = [p_1^{(0)}, p_2^{(0)}, \dots, p_r^{(0)}]$$

przykład:

$$\vec{\pi}^{(0)} = [1, 0, 0, 0, 0] \quad \vec{\pi}^{(0)} = \left[\frac{1}{2}, 0, 0, \frac{1}{2}, 0 \right]$$

Możemy obliczyć prawdopodobieństwo realizacji stanu x_B w n -tym kroku

$$\pi^{(n)}(x_B) = P\{X_n = x_B\} = \sum_{x_k \in S} P\{X_n = x_B | X_0 = x_k\} P\{X_0 = x_k\} = \sum_{x_k \in S} \pi^{(0)}(x_k) T^n(x_k, x_B)$$

Możemy też posłużyć się relacją macierzową

$$\vec{\pi}^{(1)} = \vec{\pi}^{(0)} \mathbf{T}$$

$$\vec{\pi}^{(2)} = \vec{\pi}^{(1)} \mathbf{T} = \vec{\pi}^{(0)} \mathbf{T} \mathbf{T} = \vec{\pi}^{(0)} \mathbf{T}^2$$

$$\vec{\pi}^{(n)} = \vec{\pi}^{(0)} \mathbf{T}^n$$

Rozkład niezmienniczy/stacjonarny łańcucha Markowa

Aperiodyczny i nieredukowalny łańcuch Markowa to łańcuch regularny.

Posiada on rozkład graniczny, który staje się **niezmienniczy** względem kolejnych operacji

$$\bigwedge_{\vec{\pi}^{(0)}} \vec{\pi}^{n+1} = \vec{\pi}^{(n)} \mathbf{T} \quad \longrightarrow \quad \vec{\pi}^{(n+1)} = \vec{\pi}^{(n)}$$

Ten wynik jest dla nas bardzo ważny:

- w zaawansowanych symulacjach MC, zazwyczaj nie znamy rozkładu końcowego opisującego stan dla układu, ale ta własność gwarantuje że niezależnie od wyboru punktu startowego, rozkład stacjonarny uzyskamy jeśli wygenerujemy dostatecznie długi łańcuch Markowa
- jeśli wystartujemy z rozkładu bliskiego rozkładowi stacjonarnemu wówczas do rozkładu stacjonarnego dotrzemy wykonując mniejszą liczbę kroków (krótszy łańcuch Markowa), będzie to równoważnik metody losowania ważonego

przykład – łańcuch jednorodny → rozkład niezmienniczy

$\mathbf{T} = \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.4 & 0.3 & 0.3 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$	k	π_1	π_2	π_3		k	π_1	π_2	π_3
	0	0.000000	0.500000	0.500000		0	1.000000	0.000000	0.000000
	1	0.350000	0.350000	0.300000		1	0.300000	0.200000	0.500000
	2	0.335000	0.295000	0.370000		2	0.320000	0.320000	0.360000
	3	0.329500	0.303500	0.367000		3	0.332000	0.304000	0.364000
	4	0.330350	0.303750	0.365900		4	0.330400	0.303200	0.366400
	5	0.330375	0.303555	0.366070		5	0.330320	0.303600	0.366080
	6	0.330355	0.303569	0.366075		6	0.330360	0.303576	0.366064
	7	0.330357	0.303572	0.366071		7	0.330358	0.303570	0.366072
	8	0.330357	0.303571	0.366071		8	0.330357	0.303571	0.366072
	9	0.330357	0.303571	0.366071		9	0.330357	0.303571	0.366071

przykład – łańcuch redukowalny → brak jednorodnego stanu stacjonarnego (wpływ stanów rekurencyjnych)

- kolor czerwony – aktualny wektor startowy
- ostatnia linijka – stan końcowy

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & 0 & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{3}{4} \end{bmatrix}$$

- pierwszy wektor startował w stanie rekurencyjnym
- drugi wektor w stanie przejściowym, ale ostatecznie końcowy rozkład zlokalizował się we wszystkich stanach rekurencyjnych
- trzeci wektor startuje w 2 podciągu stanów rekurencyjnych i widzimy, które separują się od pozostałych stanów
- nie ma jednego globalnego stanu stacjonarnego, różne wektory startowe dają różne stany końcowe

0	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
4	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
5	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
6	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
8	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000

0	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000
1	0.250000	0.500000	0.250000	0.000000	0.000000	0.000000
2	0.375000	0.300000	0.225000	0.050000	0.000000	0.050000
3	0.450000	0.195000	0.165000	0.065833	0.016667	0.107500
4	0.498750	0.130500	0.114750	0.079181	0.021944	0.154875
5	0.531375	0.088200	0.078525	0.085838	0.026394	0.189669
10	0.590096	0.012712	0.011385	0.098013	0.032357	0.255437
20	0.599793	0.000265	0.000238	0.099959	0.033313	0.266432
30	0.599996	0.000006	0.000005	0.099999	0.033333	0.266662
40	0.600000	0.000000	0.000000	0.100000	0.033333	0.266667
50	0.600000	0.000000	0.000000	0.100000	0.033333	0.266667

0	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000
1	0.000000	0.000000	0.000000	0.250000	0.000000	0.750000
2	0.000000	0.000000	0.000000	0.229167	0.083333	0.687500
3	0.000000	0.000000	0.000000	0.251736	0.076389	0.671875
4	0.000000	0.000000	0.000000	0.248119	0.083912	0.667969
5	0.000000	0.000000	0.000000	0.250301	0.082706	0.666992
10	0.000000	0.000000	0.000000	0.249998	0.083335	0.666667
20	0.000000	0.000000	0.000000	0.250000	0.083333	0.666667
30	0.000000	0.000000	0.000000	0.250000	0.083333	0.666667

Jaki jest związek pomiędzy pojedynczym łańcuchem Markowa a wektorem rozkładu prawdopodobieństwa π ?

Związek ten, jak to w Monte Carlo bywa, ujawnia się dopiero wówczas, gdy zamiast jednego łańcucha wygenerujemy ich dużo więcej np. $N \sim 10^6$

$$\vec{\pi}^{(0)} = [p_1^{(0)}, p_2^{(0)}, \dots, p_r^{(0)}]$$

- ustalamy liczbę łańcuchów N oraz określamy długość pojedynczego łańcucha n
- losujemy punkt startowy, jeden ze stanów w S , z rozkładu prawdopodobieństwa zdefiniowanego w $\pi^{(0)}$
- generujemy kolejne łańcuchy i zachowujemy informację, w jakim stanie zakończył się dany łańcuch
- po wygenerowaniu wszystkich łańcuchów określamy częstości stanów końcowych, częstości te są estymatorami rozkładu prawdopodobieństwa zapisanego w wektorze $\pi^{(n)}$

Po co więc generować łańcuchy Markowa?

Zapis macierzowy jest wygodny, ale możemy go stosować jedynie w przypadku dyskretnym.

Dla rozkładów ciągłych z którymi będziemy mieć zazwyczaj do czynienia, musimy tworzyć łańcuchy Markowa.

Algorytm generowania pojedynczego łańcucha Markowa

```

inicjalizacja:  $k = k_0$ ,  $x = x_k$ 
for i=1 to n do
   $U_1 \sim U(0,1)$ 
  p=0
  for m=1 to r do
     $p = p + T_{k,m}$ 
    if (  $U_1 < p$  ) then
       $k = m$ 
       $x = x_m$ 
      break
    end if
  end do
end do

```

- wybieramy numer stanu początkowego (k) oraz jego wartość (jeśli jest liczbą – może być wektorem)
- k określa numer wiersza w którym się poruszamy, sumujemy prawdopodobieństwa w wierszu do momentu aż $U < p$, wtedy zmieniamy stan $k \rightarrow m$

Proces stochastyczny

Definiujemy nową zmienną losową, zależną od 2 argumentów:

- zmiennej losowej X
- parametru t (czas, nr iteracji, niekoniecznie losowy)

$$Y = g(X, t), \quad X, Y \in R$$

Zmienna losowa Y (stan układu) określa proces stochastyczny, parametr t określa kolejność realizacji zmiennej w ciągu

$$t_1 < t_2 < t_3 < \dots < t_n \quad \longrightarrow \quad \{Y(X, t_1), Y(X, t_2), \dots, Y(X, t_n)\}$$

Zmienna losowa ma swoją fgp, jej wartość w punkcie y moglibyśmy policzyć tak

$$f_y(y) = \int \delta(y - y') f_y(y') dy' \quad (\delta - \text{delta Diraca})$$

- y stanowi parametr pod całką,
 y' jest zmienną ciągłą

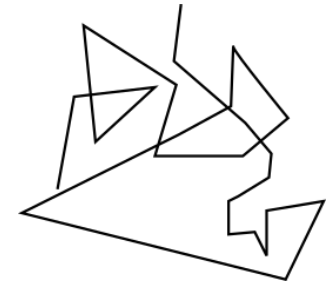
wiemy, że rzadko znamy fgp f_y więc dokonajmy transformacji i wyrażmy ją za pomocą f_x

$$f_y(y') = f_x(x) \left| \frac{dy'}{dx} \right|^{-1}$$

$$f_y(y) = \int \delta(y - g(x, t)) f_x(x) dx$$

Określmy teraz prawdopodobieństwo łączne wystąpienia zdarzeń y_1 w chwili t_1 , y_2 w chwili t_2 itd.

$$f_{n,y}(y_1, t_1; y_2, t_2; \dots; y_n, t_n) = \int \delta(y_1 - g(x_1, t_1), y_2 - g(x_2, t_2), \dots, y_n - g(x_n, t_n)) f_x(x_1) \dots f_x(x_n) d^n x$$



Jeśli proces jest całkowicie losowy (brak zależności pomiędzy y-mi) to fgp moglibyśmy zapisać w postaci iloczynu

$$f_{n,y}(y_1, t_1; y_2, t_2; \dots; y_n, t_n) = f_{1,y}(y_1, t_1) f_{1,y}(y_2, t_2) \dots f_{1,y}(y_n, t_n)$$

w przyrodzie jednak nie występują takie **procesy fizyczne** – poprzedni stan układu ma wpływ mniejszy lub większy na to co stanie się w przyszłości. Występuje pewien stopień zależności między Y-mi → korelacja.

Mając fgp możemy policzyć np. n-ty moment (lub jakiś inny)

$$\langle Y(t_1)Y(t_2) \dots Y(t_n) \rangle = \int y_1 y_2 \dots y_n f_{n,y}(y_1, t_1; \dots) dy_1 \dots dy_n$$

proces jest stacjonarny (przesunięcie w czasie nie zmienia wartości momentu) jeśli spełniony jest warunek

$$\bigwedge_{\tau > 0} \langle Y(t_1 + \tau)Y(t_2 + \tau) \dots Y(t_n + \tau) \rangle = \langle Y(t_1)Y(t_2) \dots Y(t_n) \rangle$$

Funkcja autokorelacji

$$\kappa(t_1, t_2) = \langle (Y(t_1) - \langle Y(t_1) \rangle)(Y(t_2) - \langle Y(t_2) \rangle) \rangle = \langle Y(t_1)Y(t_2) \rangle - \langle Y(t_1) \rangle \langle Y(t_2) \rangle$$

- dla $t_1=t_2$ dostajemy **wariancję** zależną od czasu

$$\kappa(t_1, t_1) = \sigma^2(t_1)$$

- czas (zaniku)/zasięg korelacji**

$$\kappa(t_1, t_2) \approx 0 \quad \Longleftrightarrow \quad |t_2 - t_1| \geq \tau_c$$

- korelacje możemy liczyć także pomiędzy zmiennymi różnego typu (np. położenie-prędkość, oddziaływanie-objętość)

$$Y_i(t) = g_i(X, t)$$

$$K_{ij}(t_1, t_2) = \langle (Y_i(t_1) - \langle Y_i(t_1) \rangle)(Y_j(t_2) - \langle Y_j(t_2) \rangle) \rangle$$

i skonstruować **macierz kowariancji**

$$\mathbf{K} = \begin{bmatrix} K_{11} & K_{12} & \dots & K_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ K_{n1} & K_{n2} & \dots & K_{nn} \end{bmatrix}$$

- K_{mm} to wartości funkcji autokorelacji

Łańcuch (proces) Markowa możemy traktować jako realizację procesu stochastycznego – proces Markowa.

Definiujemy warunkowe fgp korzystając z własności łańcucha Markowa

$$f_{1|n-1}(y, t | y_{n-1}, t_{n-1}; y_{n-2}, t_{n-2}; \dots; y_1, t_1) = f_{1|1}(y, t | y_{n-1}, t_{n-1})$$

i to fgp potraktować jako **prawdopodobieństwo przejścia** pomiędzy starym a nowym stanem

$$T(y, t | y_{n-1}, t_{n-1}) = f_{1|1}(y, t | y_{n-1}, t_{n-1})$$

Analogicznie jak w przypadku dyskretnym, łańcuch Markowa jest jednoznacznie określony przez podanie relacji wiążącej f_y i T

$$f_y(y, t) = f_{1,y}(y, t) \quad T(y_2, t_2 | y_1, t_1)$$

To czego szukamy tworząc łańcuch Markowa to najczęściej pewien wielowymiarowy rozkład stacjonarny, który generujemy sekwencyjnie

$$\begin{aligned} f_{3,y}(y_3, t_3; y_2, t_2; y_1, t_1) &= T(y_3, t_3 | y_2, t_2) f_{2,y}(y_1, t_1; y_2, t_2) \\ &= T(y_3, t_3 | y_2, t_2) T(y_2, t_2 | y_1, t_1) f_{1,y}(y_1, t_1) \end{aligned}$$

Widzimy że przejście $1 \rightarrow 3$ odbywa się z wykorzystaniem stanu pośredniego 2.
Jeśli interesuje nas tylko rozkład początkowy i końcowy, informację o 2 redukujemy

$$\begin{aligned} f_{2,y}(y_1, t_1; y_3, t_3) &= \int f_{3,y}(y_1, t_1; y_2, t_2; y_1, t_1) dy_2 \\ &= f_{1,y}(y_1, t_1) \int T(y_3, t_3 | y_2, t_2) T(y_2, t_2 | y_1, t_1) dy_2 \end{aligned}$$

- równanie Chapmana-Kolmogorova dla zmiennej losowej ciągłej

Stacjonarność procesu Markowa

Jeśli $X(t)$ definiuje proces stochastyczny i dla dowolnego $\varepsilon > 0$ proces $X(t+\varepsilon)$ ma te same statystyki, to jest to proces stacjonarny, np.

$$\langle X^n(t) \rangle = \langle X^n(t + \varepsilon) \rangle$$

Własność tę można sformułować w sposób alternatywny, bardziej użyteczny dla naszych przyszłych celów, mianowicie, proces stacjonarny to taki, dla którego łączny rozkład gęstości prawdopodobieństwa jest niezależny od czasu

$$f_{n,y}(x_n, t_n; x_{n-1}, t_{n-1}; \dots; x_1, t_1) = f_{n,y}(x_n, t_n + \varepsilon; x_{n-1}, t_{n-1} + \varepsilon; \dots; x_1, t_1 + \varepsilon)$$

Ponieważ w procesie Markowa taki rozkład uzyskujemy w wyniku złożenia prawdopodobieństw warunkowych dla 2 następujących po sobie chwil czasowych

$$f_{n,y}(y_n, t_n; y_{n-1}, t_{n-1}; \dots; y_1, t_1) = T(y_n, t_n | y_{n-1}, t_{n-1}) \dots T(y_2, t_2 | y_1, t_1) f_{1,y}(y_1, t_1)$$

wystarczy jeśli rozkład zredukowany $f_{1,y}$ stanie się niezależny od czasu

$$(y_n, t_n) \rightarrow (y, t) \quad f_{1,y}(y, t) = \int f_{n,y}(y, t; y_{n-1}, t_{n-1}; \dots; y_1, t_1) dy_{n-1} dy_{n-2} \dots dy_1$$

$$f_{1,y}(y, t) = f_{1,y}(y, t + \varepsilon) = f_{1,y}(y)$$

a rozkład warunkowy będzie tylko od różnicy czasów realizacji tych stanów, a nie od ich wartości bezwzględnych

$$T_{2,y}(y_{k+1}, t_{k+1} | y_k, t_k) = T_{2,y}(y_{k+1}, t_{k+1} + \varepsilon | y_k, t_k + \varepsilon) = T_{2,y}(y_{k+1}, t_{k+1} - t_k | y_k, 0)$$

Uwaga:

- bardzo często w modelowaniu lub symulacjach MC pierwszym wymaganym etapem jest uzyskanie stanu stacjonarnego
- uzyskanie stanu stacjonarnego oznacza, że udało nam się znaleźć poszukiwaną funkcję rozkładu $f_{n,y}$

Jednorodność procesu Markowa

Rozkład stacjonarny możemy zdefiniować też w postaci asymptotycznej dla zredukowanego rozkładu warunkowego

$$\lim_{t \rightarrow \infty} f_{1|1}(y, t | y_0, t_0) = f_1(y)$$

proces jednorodny to taki, który staje się stacjonarny niezależnie od rozkładu początkowego (punktu startowego)

$$\bigwedge_{y' \neq y_0, t' > t_0} \lim_{t \rightarrow \infty} f_{1|1}(y, t | y', t') = f_1(y)$$

Przykład tego typu łańcucha już widzieliśmy w przypadku rozkładu dyskretnego – jeśli nie było stanów rekurencyjnych (absorpcyjnych) to końcowy rozkład prawdopodobieństwa jednoznacznie ustalony.

Uwaga:

jednorodność to ważna cecha procesu Markowa, jeśli proces jest jednorodny to w wyniku przeprowadzenia symulacji MC zawsze znajdziemy rozkład stacjonarny niezależnie od wyboru punktu startowego

Ergodyczność procesu Markowa

Zgodnie z prawem wielkich liczb, błąd pomiędzy wartością oczekiwaną a średnią z próby powinien maleć ze wzrostem liczby losowań. Na tym założeniu bazuje twierdzenie o ergodyczności procesu Markowa.

Definiujemy zmienną losową w postaci średniej dla wielkości losowej $x(t)$

$$\bar{X}(T) = \frac{1}{2T} \int_{-T}^T x(t) dt \quad - \text{całka oznacza „sumowanie” po stanach określonych parametrem ciągłym } t$$

Jeśli $x(t)$ opisuje proces stacjonarny to wartość oczekiwaną średniej możemy liczyć jako średnią zmiennej losowej $x(t)$ generowanej w części stacjonarnej łańcucha

$$\langle \bar{X}(T) \rangle = \langle x \rangle_s$$

Żeby się o tym przekonać należy sprawdzić jak zmienia się wariancja, potrzebny będzie drugi moment uśredniony

$$\langle \bar{X}^2(T) \rangle = \frac{1}{4T^2} \int_{-T}^T \int_{-T}^T dt_1 dt_2 \langle x(t_1) x(t_2) \rangle$$

$$\begin{aligned} \text{var}\{\bar{X}(T)\} &= \langle \bar{X}^2(T) \rangle - \langle x \rangle^2 = \frac{1}{4T^2} \int_{-2T}^{2T} d\tau R(\tau) \underbrace{\int_{|\tau|}^{2T} dt}_{>0} \\ &= \frac{1}{4T^2} \int_{-2T}^{2T} d\tau R(\tau) (2T - |\tau|) \end{aligned}$$

transformacja – obrót:

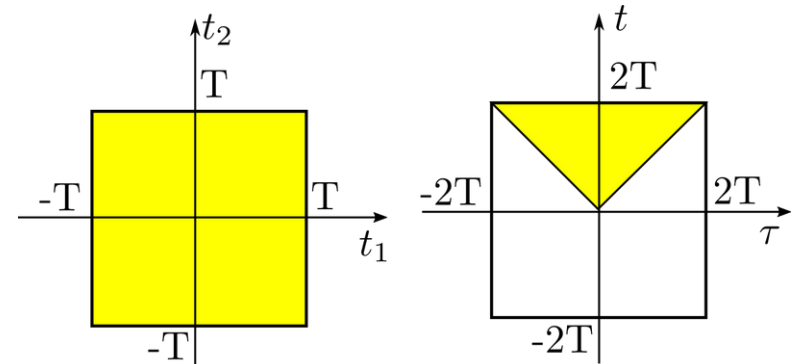
$$\tau' = (t_1 - t_2) / \sqrt{2}$$

$$t' = (t_1 + t_2) / \sqrt{2}$$

$R(t_1 - t_2)$ – funkcja autokorelacji

$$R(t_1 - t_2) = \langle x(t_1) x(t_2) \rangle - \langle x \rangle^2$$

- żółte pole dla obu par zmiennych jest identyczne
- całka po zmiennej t musi być nieujemna tj. nie może zmieniać znaku całki



Wariancja zależna jest od funkcji korelacji czasowej

$$\text{var}\{\bar{X}(T)\} = \frac{1}{2T} \int_{-2T}^{2T} d\tau R(\tau) \left(1 - \frac{|\tau|}{2T}\right)$$

Zazwyczaj korelacja w czasie zanika eksponencjalnie wówczas

$$R \sim C \exp\left(-\frac{|\tau|}{\tau_c}\right) \longrightarrow \int_0^\infty |R(\tau)| d\tau < \infty \longrightarrow \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-2T}^{2T} d\tau \left(1 - \frac{|\tau|}{2T}\right) R(\tau) = 0$$

τ_c - czas korelacji

Wariancja znika dla $T \rightarrow \infty$ więc średnia liczona w części stacjonarnej łańcucha Markowa, można interpretować jako pierwiastek z wartości średniej kwadratowej

$$\text{var}\{\bar{X}(T)\} = \langle \bar{X}^2(T) \rangle - \langle x \rangle^2 = 0 \longrightarrow \lim_{T \rightarrow \infty} \langle x \rangle = \sqrt{\langle \bar{X}^2(T) \rangle} \stackrel{\text{var}=0}{=} \sqrt{\bar{X}^2(T)} = \bar{X}(T)$$

$$\lim_{T \rightarrow \infty} \bar{X}(T) = \langle x \rangle_s$$

Uwagi:

- wynik ważny ze względu na sposób gromadzenia informacji w MC
- wartość oczekiwaną możemy szacować licząc średnią w stacjonarnej części łańcucha Markowa – jeden łańcuch, ale długi
- wariancję zmniejszamy wydłużając łańcuch, kolejne stany układu generowane są z rozkładu stacjonarnego

Dynamika procesu stochastycznego → równanie typu Master

$$T(y_2, t_2 | y_1, t_1) = T(y_2, t_1 + \tau | y_1, t_1) = T_\tau(y_2 | y_1), \quad \tau = t_2 - t_1$$

prawdopodobieństwo przejścia jest funkcją czasu

$$T_\tau(y_2 | y_1) = (1 - a\tau)\delta(y_2 - y_1) + \tau W(y_2 | y_1)$$

Parametr a to czynnik normalizacyjny

$$\tau \in (0, \infty)$$

$$1 - a\tau \geq 0 \quad \rightarrow \quad a \leq \frac{1}{\tau}$$

- wyraz $(1-a\tau)$ określa prawdopodobieństwo braku zmiany stanu

$$\tau = 0 \quad \rightarrow \quad T_\tau(y_2 | y_1) = \delta(y_2 - y_1) \rightarrow y_2 = y_1$$

- drugi wyraz zawiera tzw. częstość przejść $W(y_2 | y_1)$ - to gęstość prawdopodobieństwa na jednostkę czasu

Przejście do kolejnego stanu y_3 można wykonać wykorzystując równanie Chapmana

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

łączymy oba wyrażenia podstawiając $T_\tau(y_3 | y_2)$

$$\lim_{\tau' \rightarrow 0} \frac{T_{\tau+\tau'}(y_3 | y_1) - T_\tau(y_3 | y_1)}{\tau'} = \frac{\partial T_\tau(y_3 | y_1)}{\partial \tau} = \int [W(y_3 | y_2) T_\tau(y_2 | y_1) - W(y_2 | y_3) T_\tau(y_3 | y_1)] dy_2$$

Stan y_1 jest dowolny (załóżmy, że ustalony), więc liczymy szybkość zmian w stanie docelowym (y_3).

Zapiszmy równanie w bardziej intuicyjnej postaci

$$T_\tau(y_3|y_1) \rightarrow f_y(y, t)$$

równanie typu Master (ang. Master Equation)

rozkład ciągły:

$$\frac{\partial f_y(y, t)}{\partial t} = \int \left[\underbrace{W(y|y') f_y(y', t)}_{(1)} - \underbrace{W(y'|y) f_y(y, t)}_{(2)} \right] dy'$$

rozkład dyskretny:

$$\frac{dp_k(t)}{dt} = \sum_m \left[\underbrace{W_{k,m} p_m(t)}_{(1)} - \underbrace{W_{m,k} p_k(t)}_{(2)} \right]$$

- równanie typu Master opisuje szybkość zmian rozkładu gęstości/prawdopodobieństwa zmiennej losowej (ang. gain-lose equation)
- (1) - wyraz dodatni, określa tempo przejścia do stanu y/k (**zysk**)
- (2) – wyraz ujemny, określa tempo opuszczania danego stanu i przejście do innych (**strata**)
- po rozpoczęciu generowania łańcucha Markowa zaobserwujemy zmiany fgp,
ale co się stanie, gdy osiągniemy stan stacjonarny?

Stan stacjonarny łańcucha Markowa

W stanie stacjonarnym $f_y = \text{const}$, zatem pochodna znika

$$\frac{\partial f_y(y, t)}{\partial t} = \int \left[\underbrace{W(y|y') f_y(y', t)}_{(1)} - \underbrace{W(y'|y) f_y(y, t)}_{(2)} \right] dy' = 0$$

wyrażenie pod całką musi znikać co prowadzi do **warunku równowagi** (ang. *detailed balance*)

rozkład ciągły

$$W(y|y') f_y(y', t) = W(y'|y) f_y(y, t)$$

rozkład dyskretny

$$W_{k,m} p_m(t) = W_{m,k} p_k(t)$$

- przejście ze stanu y do stanu y' odbywa się z identyczną szybkością jak ze stanu y' do y
- warunek ten jest podstawą działania algorytmu Metropolis'a generowania łańcuchów Markowa
- równanie Master ma bezpośrednią interpretację fizyczną, iloczyn $W\Delta t$ określa prawdopodobieństwo przejścia w krótkim przedziale czasu

- równanie Master możemy użyć np. do opisu przejść w układzie kwantowym wywołanym jakimś czynnikiem zewnętrznym np. pole elektryczne może zmusić elektron do zmiany stanu orbitalnego

$$W_{mk}f(E_k) = \frac{2\pi}{\hbar} |H_{km}|^2 \rho(E_k)$$

$$H_{km} = \langle \psi_k(\vec{r}) | U(\vec{r}, t) | \psi_m(\vec{r}) \rangle \quad \bullet \text{ czynnik zaburzający-wymuszający przejście}$$

$$\rho(E) \quad \bullet \text{ gęstość stanów bez zaburzenia (jeśli stan jest nieobsadzony to nie generuje przejść)}$$

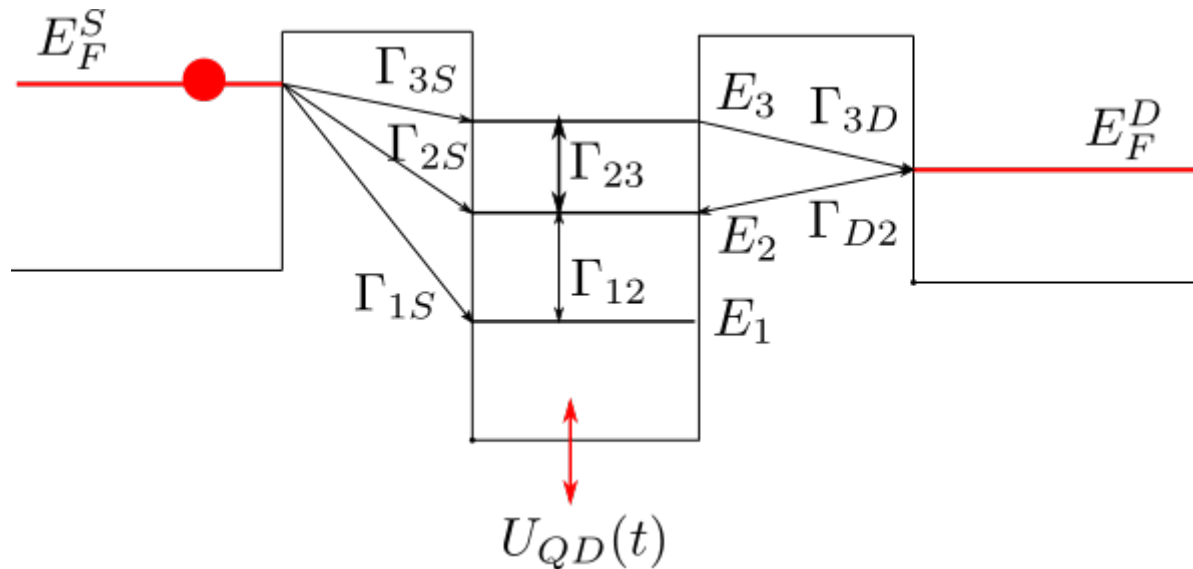
- z równaniem typu Master mamy do czynienia w problemie rozpadu promieniotwórczego (szeregi promieniotwórcze), podobnymi równaniami opisywana jest dynamika reakcji chemicznych

$$\left\{ \begin{array}{ll} \frac{dN_1}{dt} = -\lambda_1 N_1(t) & \text{- izotop matka} \\ \frac{dN_2}{dt} = \lambda_1 N_1(t) - \lambda_2 N_2(t) & \text{- potomek 1} \\ \frac{dN_3}{dt} = \lambda_2 N_2(t) - \lambda_3 N_3(t) & \text{- potomek 2} \end{array} \right.$$

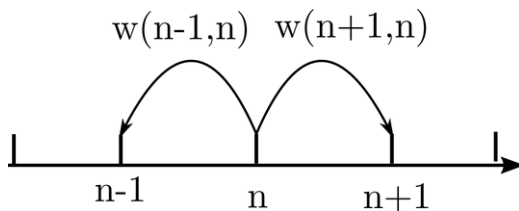
- izotop macierzysty tylko znika
- ilość atomów potomka 1 ubywa ze względu na rozpad, a przybywa ze względu na izotop macierzysty
- ta sama sytuacja

Przykład – transport przez kropkę kwantową.

- transport nierównowagowy wymuszony różnicą napięć pomiędzy drenem a źródłem (bramkami sterującymi)
- możemy określić częstości przejść Γ pomiędzy różnymi stanami kwantowymi
- celem może być wyznaczenie rozkładu stacjonarnego obsadzenia stanów kropki kwantowej dla ustalonych poziomów energii Fermiego źródła i drenu – **prawdopodobieństwa obsadzeń stanów**
- dodatkowy czynnik zaburzający to potencjał kropki, który może oscylować (stanu stacjonarnego raczej nie uzyskamy – ale to zależy od relacji między częstością przejść a częstością oscylacji zaburzenia)



Symulacja procesu jednokrokowego dla równania typu Master (birth-death process)



definiujemy częstość przejść tylko pomiędzy sąsiednimi stanami

$$\begin{cases} w(n', n) \neq 0, & n' = n \pm 1 \\ w(n', n) = 0, & |n' - n| > 1 \end{cases}$$

- równanie Master procesu

$$\frac{\partial f(n, t)}{\partial t} = \underbrace{w(n, n+1)f(n+1, t) + w(n, n-1)f(n-1, t)}_{\text{zysk}} - \underbrace{w(n+1, n)f(n, t) + w(n-1, n)f(n, t)}_{\text{strata}}$$

- wprowadzamy oznaczenia – krótszy zapis

$$r(n+1) = w(n, n+1)$$

$$r(n) = w(n-1, n)$$

$$g(n-1) = w(n, n-1)$$

$$g(n) = w(n+1, n)$$

$$\frac{\partial f(n, t)}{\partial t} = r(n+1)f(n+1, t) + g(n-1)f(n-1, t) - [g(n) + r(n)]f(n, t)$$

- zajmiemy się ostatnim wyrazem odpowiedzialnym za „opróżnianie/rozpad” stanu
- pierwsze dwa wyrazy (zysk) będą się generować automatycznie, gdy „opróżniać/rozpadać” się będą stany sąsiednie

- opróżnianie stanu n-tego będzie zachodziło z częstością określoną przez sumę po wszystkich kanałach rozpadu (tutaj są 2)

$$\lambda(n) = r(n) + g(n)$$

- definiujemy prawdopodobieństwo opróżnienia/rozpadu stanu n-tego w przedziale dt

$$q_1 = \lambda(n)dt \quad (1-q_1) - \text{prawdopodobieństwo, że nic się nie stanie}$$

- załóżmy, że rozpad nastąpi po czasie $t=(N+1)dt$

$$q_{N+1} = [1 - \lambda(n)dt]^N \lambda(n)dt = \left[1 - \frac{\lambda(N+1)dt}{N+1}\right]^N \lambda dt = \left[1 - \frac{\lambda t}{N+1}\right]^N \lambda dt$$

$$\lim_{N \rightarrow \infty} \left[1 - \frac{\lambda t}{N+1}\right]^N = \underbrace{e^{-\frac{\lambda t}{N+1}} \dots e^{-\frac{\lambda t}{N+1}}}_{\times N} = \exp\left(-\frac{\lambda t N}{N+1}\right) = \exp\left(-\frac{\lambda t}{1 + \frac{1}{N}}\right) \approx \exp(-\lambda t)$$

$$\lim_{\substack{dt \rightarrow 0 \\ N \rightarrow \infty}} t = (N+1)dt = \text{const} \quad \longrightarrow \quad q_t = \lambda e^{-\lambda t} dt = f(t)dt$$

- prawdopodobieństwo, że w chwili t nastąpi przejście do innego stanu

- mamy f(t) więc teraz możemy określić wartość oczekiwaną czasu po jakim nastąpi rozpad

$$\tau = -\frac{1}{\lambda} \ln(U_1), \quad U_1 \sim U(0, 1)$$

- wiemy już kiedy, teraz dokonujemy wyboru stanu docelowego (mamy dwie możliwości)

$$\lambda = g + r \quad p_r = \frac{r}{\lambda} \quad p_g = \frac{g}{\lambda} = 1 - p_r$$

$$U_2 \sim U(0, 1) \quad \longrightarrow \quad U_2 \leq p_r \rightarrow n = n - 1 \quad \vee \quad U_2 > p_r \rightarrow n = n + 1$$

Pseudokod generowania M łańcuchów dla procesu jednokrokowego (**birth-death process**)

```

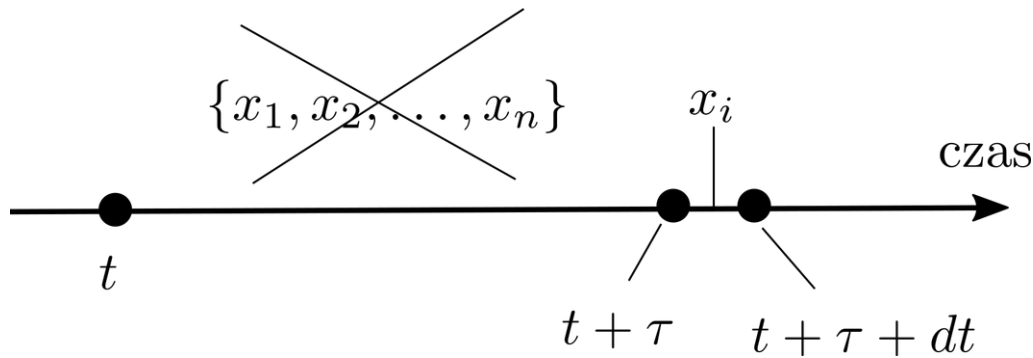
1  inicjalizacja:  $r, g, \lambda, t_{end}, t_0, n_0$ 
2
3  for  $m=1$  to  $M$  do
4       $t = t_0$ 
5       $n = n_0$ 
6      while  $t < t_{end}$  do
7           $U_1 \sim U(0, 1)$ 
8           $\tau = -\frac{1}{\lambda} \ln(U_1)$ 
9           $U_2 \sim U(0, 1)$ 
10         if  $(U_2 \leq \frac{r}{\lambda})$ 
11              $k = -1$ 
12         else
13              $k = 1$ 
14         endif
15          $n = n + k$ 
16          $t = t + \tau$ 
17
18         zbieramy potrzebne dane
19
20     end do
21
22 end do

```

Rozwiązanie równania typu Master dla grupy zdarzeń losowych – algorytm Gillespie

Rozważmy ogólny problem:

- dany jest ciąg zdarzeń losowych $\{x_1, x_2, \dots, x_n\}$
- jakie jest prawdopodobieństwo, że w przedziale czasu $(t, t+\tau)$ nie zaobserwujemy realizacji żadnego stanu, ale w przedziale $(t+\tau, t+\tau+dt)$ zostanie zarejestrowane zdarzenie x_i ?



Przydatna będzie tu wielkość opisująca „dynamikę” zachodzących **zdarzeń/procesów losowych**

$$\Gamma_i = \frac{\text{liczba obserwowanych zdarzeń losowych } X_i}{\text{jednostka czasu}} = \left[\frac{1}{s} \right]$$

Wielkość Γ określa szybkość (tempo) zachodzących procesów (ang. **rate**).

Uwaga: wielkość ta nie jest częstością, ponieważ termin ten odnosi się do procesów periodycznych (np. w czasie).

$$p_i = \Gamma_i dt$$

- to prawdopodobieństwo zaobserwowania zdarzenia x_i w przedziale czasu $(t, t+dt)$

Korzystając ze związku prawdopodobieństwa z szybkością zachodzącego procesu możemy zapisać dwie relacje

$$p(X_i) = \Gamma_i \Delta t$$

$$p(\tilde{X}_i) = 1 - \Gamma_i \Delta t$$

- to prawdopodobieństwo że zdarzenie nie zostanie zaobserwowane

Drugą relację wykorzystamy do określenia, prawdopodobieństwa że w czasie Δt nie zajdzie jakiegokolwiek zdarzenie zawarte w grupie stanów $\{x_1, x_2, \dots, x_n\}$

$$p(\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n) = (1 - \Gamma_1 \Delta t) (1 - \Gamma_2 \Delta t) \cdot \dots \cdot (1 - \Gamma_n \Delta t) = \prod_{i=1}^n (1 - \Gamma_i \Delta t)$$

Powyższe wyrażenie jest prawdziwe dla małych wartości Δt . Ponieważ chcemy otrzymać podobny wynik dla skończonego (domyślnie: długiego) przedziału czasu τ , musimy problem rozważyć jako sekwencję N krótkich odcinków czasowych

$$\tau = N \cdot \Delta t \quad \rightarrow \quad \Delta t = \frac{\tau}{N}$$

Konstruujemy docelowe wyrażenie opisujące prawdopodobieństwo braku jakichkolwiek zdarzeń w przedziale o długości τ i zdarzenie X_i występujące po upływie tego odcinka czasu

$$p(\tau, i) = \lim_{N \rightarrow \infty} \underbrace{\left[\prod_{i=1}^n \left(1 - \Gamma_i \frac{\tau}{N} \right) \right]^N}_{\text{brak zdarzeń w } (t, t+\tau)} \cdot \underbrace{\Gamma_i dt}_{\text{zdarzenie } X_i}$$

Wyrażenie upraszczamy wymnażając wyrażenia w nawiasach okrągłych, a następnie zostawiając tylko wyrazy liniowe względem Γ_i

$$p(\tau, i) = \lim_{N \rightarrow \infty} \left[\prod_{i=1}^n \left(1 - \Gamma_i \frac{\tau}{N} \right) \right]^N \cdot \Gamma_i dt \approx \lim_{N \rightarrow \infty} \left[1 - \sum_{i=1}^n \frac{\tau}{N} \Gamma_i \right]^N \cdot \Gamma_i dt$$

podstawmy: $\Gamma_{max} = \sum_{i=1}^n \Gamma_i$

$$p(\tau, i) = \lim_{N \rightarrow \infty} \left[1 - \frac{\tau}{N} \Gamma_{max} \right]^N \cdot \Gamma_i dt = e^{-\tau \Gamma_{max}} \Gamma_i dt$$

$$p(\tau, i) = \Gamma_{max} e^{-\tau \Gamma_{max}} \cdot dt \cdot \frac{\Gamma_i}{\Gamma_{max}}$$

powyższe wyrażenie możemy zapisać w ogólnej postaci

$$p(\tau, i) = p(\tau) \cdot p(i|\tau)$$

$$p(\tau) = \Gamma_{max} e^{-\tau \Gamma_{max}} \cdot dt$$

$$p(i|\tau) = \frac{\Gamma_i}{\Gamma_{max}}$$

$p(\tau)$ – określa rozkład prawdopodobieństwa (**fgp**) dla zmiennej czasowej τ , czyli przedziału czasowego pomiędzy dwoma kolejnymi zdarzeniami

$p(i|\tau)$ – to prawdopodobieństwo warunkowe realizacji zdarzenia X_i po czasie τ (od poprzedniego)

Algorytm Gillespie (wersja podstawowa)

1) określamy ciąg stanów losowych oraz odpowiadających im szybkości realizacji i obliczamy ich sumę $\Gamma_{\max}$

$$\{x_1, x_2, \dots, x_s\} \xrightarrow{s \leq n} \{\Gamma_1, \Gamma_2, \dots, \Gamma_n\} \rightarrow \Gamma_{\max} = \sum_{i=1}^n \Gamma_i$$

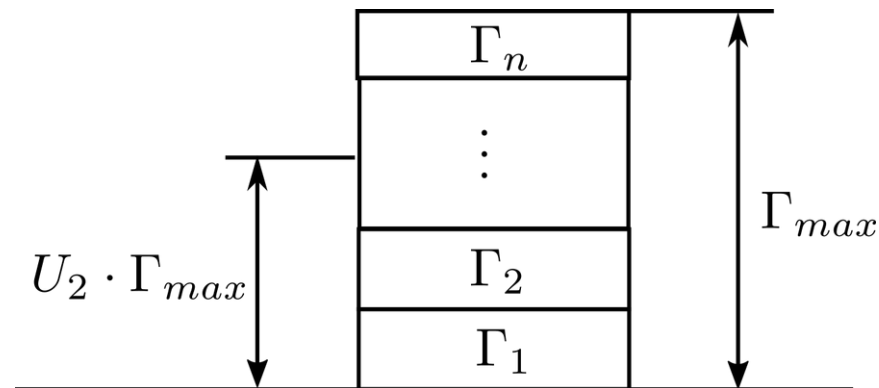
2) dzięki znajomości aktualnej wartości $\Gamma_{\max}$ (może się zmieniać w czasie) losujemy wartość τ

$$F(\tau) = \int_0^{\tau} d\tau' \Gamma_{\max} e^{-\tau' \Gamma_{\max}} = 1 - e^{-\tau \Gamma_{\max}} = U_1 \sim U(0, 1)$$

$$\tau = -\frac{1}{\Gamma_{\max}} \ln(U_1)$$

3) dla chwili $t+\tau$ określamy typ realizowanego zdarzenia x_i – losowanie z rozkładu dyskretnego

$$U_2 \sim U(0, 1)$$



4) po wylosowaniu określonego zdarzenia – modyfikujemy stan układu

Przykład zastosowania algorytmu Gillespie: symulacja dynamiki reakcji chemicznej

W układzie znajdują się trzy związki chemiczne $\{x_1, x_2, x_3\}$, x_3 powstaje w wyniku połączenia substratów x_1 i x_2 (k_3), związek x_3 jest usuwany z szybkością określoną przez (k_4). Aby reakcja zachodziła w sposób ciągły, do układu dodawane są składniki x_1 i x_2 z intensywnością określoną przez (k_1) i (k_2).

Zdefiniujmy problem matematycznie za pomocą układu równań różniczkowych

$$\begin{array}{l} x_1 + x_2 \xrightarrow{k_3} x_3 \\ x_3 \xrightarrow{k_4} 0 \end{array} \quad \frac{dx_3}{dt} = k_3 x_1 x_2 - k_4 x_3$$

$$\begin{array}{l} x_1 \xrightarrow{\Gamma_3} 0 \\ 0 \xrightarrow{k_1} x_1 \end{array} \quad \frac{dx_1}{dt} = -k_3 x_1 x_2 + k_1$$

$$\begin{array}{l} x_2 \xrightarrow{\Gamma_3} 0 \\ 0 \xrightarrow{k_2} x_2 \end{array} \quad \frac{dx_2}{dt} = -k_3 x_1 x_2 + k_2$$

Jeśli ilość składników x_1 , x_2 , x_3 jest duża ($\sim N_{\text{avogadro}}$) to układ równań różniczkowych możemy rozwiązywać przy użyciu standardowych metod numerycznych (np. metoda RK4) – mimo iż proces jest stochastyczny, to fluktuacje są na tyle małe że można je zaniedbać.

Ale, gdy ilość składników jest mała, wówczas realizacja jednego zdarzenia powoduje duże fluktuacje ilości składników, które zmieniają się w sposób dyskretny. Dyskretnych zmian ilości składników, standardowe metody rozwiązywania RRCz nam nie odtworzą.

Konieczne jest użycie MC.

W układzie możliwe jest wystąpienie 4 zdarzeń, w których **ilości składników zmieniają się o 1**:

$Lp.$	rodzaj procesu	zmiana stanu układu	szybkość procesu
1	dodanie składnika x_1	$x_1 \rightarrow x_1 + 1$	$\Gamma_1 = k_1$
2	dodanie składnika x_2	$x_2 \rightarrow x_2 + 1$	$\Gamma_2 = k_2$
3	reakcja syntezy składnika x_3	$x_3 \rightarrow x_3 + 1, x_1 \rightarrow x_1 - 1, x_2 \rightarrow x_2 - 1$	$\Gamma_3 = k_3 x_1 x_2$
4	usuwanie składnika x_3	$x_3 \rightarrow x_3 - 1$	$\Gamma_4 = k_4 x_3$

Pseudokod algorytmu Gillespie dla problemu reakcji chemicznej

`initialization :`

`$x_1, x_2, x_3$`

`$k_1, k_2, k_3, k_4$`

`$t_{max}$`

`set :  $t = 0$`

`while (  $t < t_{max}$  ) do`

`compute :  $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4 \rightarrow \Gamma_{max} = \sum_{i=1}^n \Gamma_i$`

`draw  $\tau$  :  $U_1 \sim U(0, 1) \rightarrow \tau = -\frac{1}{\Gamma_{max}} \ln(U)$`

`draw event :  $U_2 \sim U(0, 1) \rightarrow i = \min \left\{ s : \sum_{j=1}^s \Gamma_j > U_2 \cdot \Gamma_{max} \right\}$`

`for event "i" update states :  $x_1, x_2, x_3$`

`update time :  $t = t + \tau$`

`end do`

Równanie ruchu z czynnikiem stochastycznym

Najprostsze równanie ruchu ma postać

$$\frac{dx}{dt} = y(x, t)$$

rozdzielmy zmienne

$$dx = y(x, t)dt \quad \longrightarrow \quad x(t + dt) - x(t) = y(x, t)dt$$

taki proces jest ciągły

$$\lim_{dt \rightarrow 0} x(t + dt) = x(t)$$

a ponieważ pochodna istnieje to proces jest „gładki”

$$\lim_{dt \rightarrow 0} \frac{x(t + dt) - x(t)}{dt} < \infty$$

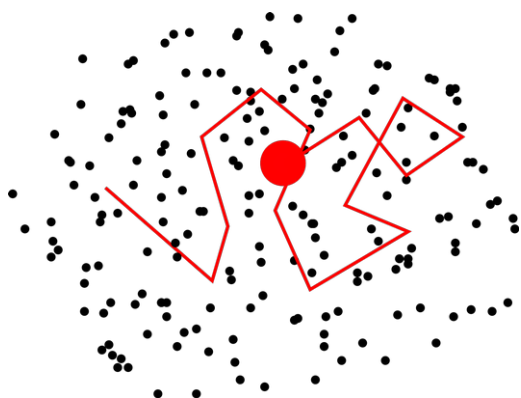
Prawa strona równania różniczkowego może mieć charakter losowy np. co wynika z oddziaływania z otoczeniem

$$x(t + dt) = x(t) + Y(x, t)dt$$

Jeśli na $Y(x, t)$ narzucimy warunek równowagi, to powyższe równanie będzie generatorem łańcucha Markowa zmiennej ciągłej x , ponieważ realizacja stanu dla $t+dt$ zależy tylko i wyłącznie od tego w jakim stanie był układ w chwili t .

Jak można zdefiniować $Y(x, t)$? - najprostszy przypadek to **ruchy Browna**

Ruchy Browna



- drobina kurzu na powierzchni wody nie spoczywa, ale porusza się chaotycznym ruchem pod wpływem zderzeń z cząsteczkami wody, efekt jest potwierdzeniem atomowej budowy ciał
- niemożliwe jest zarejestrowanie każdego zderzenia i jego wpływu, to co widzimy to uśredniony efekt makroskopowy
- po raz pierwszy efekt zaobserwował duński biolog Jan IngenHousz w 1785
- efekt spopularyzował Szkot Robert Brown w 1827 roku i od niego wzięła się też nazwa

Przemieszczenia ciała są wypadkową wielu zderzeń, ale w przybliżeniu możemy założyć że obiekt porusza się w krótkich odcinkach czasu po linii prostej

$$X = x_0 + X_1 + X_2 + \dots + X_n$$

załóżmy że start jest w 0

$$x_0 = 0$$

Proces zderzeń jest chaotyczny i anizotropowy – przesunięcia lewo/prawo, góra/dół są równieprawdopodobne, lub inaczej: ruch Browna jest realizacją błędzenia przypadkowego z prawdopodobieństwem niezależnym od kierunku

$$\langle X_1 \rangle = \langle X_2 \rangle = \dots = \langle X_n \rangle = 0 \quad \longrightarrow \quad \langle X \rangle = \left\langle \sum_{i=1}^n X_i \right\rangle = \sum_{i=1}^n \langle X_i \rangle = 0$$

Kolejne zderzenia i będące ich wynikiem przemieszczenia są realizacją tego samego procesu – wariancja jest identyczna

$$\text{var}\{X_1\} = \text{var}\{X_2\} = \dots = \text{var}\{X_n\} = \Delta x^2$$

Wariancja całego ruchu

$$\text{var}\{X\} = \langle X^2 \rangle - \langle X \rangle^2 = \langle X^2 \rangle = \sum_{i=1}^n \langle X_i^2 \rangle = \sum_{i=1}^n \Delta x^2 = n \Delta x^2$$

Zmiany obserwujemy w kolejnych chwilach czasowych, które możemy potraktować jako równoodległe

$$t = n \Delta t$$

$$\text{var}\{X\} = \frac{\Delta x^2}{2 \Delta t} 2t = 2Dt = \sigma^2$$

Wariancja rośnie z czasem, D to współczynnik dyfuzji – określa rodzaj oddziaływań cząstka-otoczenie.

Przemieszczanie się cząsteczki obserwujemy przez długi czas, zatem suma zmiennych losowych powinna dążyć do rozkładu normalnego, którego szerokość określać będzie zmienna w czasie wariancja.

$$\langle X(t) \rangle = \sqrt{2Dt} N(0, 1) \quad \longrightarrow \quad X_i = \sqrt{2D\Delta t} N(0, 1)$$

uwaga: zderzenia i przesunięcia to zdarzenia losowe niezależne, dlatego każde nowe położenie możemy potraktować jako punkt startowy, dlatego pod pierwiastkiem jest Δt zamiast czasu całkowitego t .

Generator łańcucha Markowa dla ruchów Browna ma postać procesu Wienera

$$x(t + \Delta t) = x(t) + \sqrt{2D\Delta t} N(0, 1)$$

Proces Wienera – wartość oczekiwana zmiennej losowej jest równa zero, a jej zmiany opisuje rozkład normalny (proces Gaussowski)

Modelowanie ruchów Browna łańcuchem Markowa

$$x(t + \Delta t) = x(t) + \sqrt{D\Delta t}N(0, 1)$$

```
inicjalizacja:  $x = x_0$ ,  $y = y_0$ ,  
do i=1 to n  
     $\Delta x \sim N(0, \sqrt{2D\Delta t})$   
     $\Delta y \sim N(0, \sqrt{2D\Delta t})$   
     $x = x + \Delta x$   
     $y = y + \Delta y$   
end do
```

Ruchy Browna → proces Wienera → równanie dyfuzji

X to zmienna losowa o rozkładzie normalnym więc możemy określić jej fgp (**funkcję rozkładu**): $f(x,t)$

$$X_i \sim N(0, \sqrt{2D\Delta t}) \quad \longrightarrow \quad f(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}$$

Skoro jest funkcją dwuargumentową, policzmy jej pochodne cząstkowe i porównajmy je ze sobą

$$\frac{\partial f(x, t)}{\partial t} = -\frac{f(x, t)}{t} \left[\frac{1}{2} - \frac{x^2}{4Dt} \right]$$

$$\frac{\partial f}{\partial x} = -\frac{x}{2Dt} f(x, t)$$

$$\frac{\partial^2 f}{\partial x^2} = -\frac{f(x, t)}{Dt} \left[\frac{1}{2} - \frac{x^2}{4Dt} \right]$$

z porównania (1) i (3) dostajemy **równanie dyfuzji**

$$\frac{\partial f(x, t)}{\partial t} = D \frac{\partial^2 f(x, t)}{\partial x^2}$$

proces Wienera – łańcuch Markowa

$$x(t + \Delta t) = x(t) + \sqrt{2D\Delta t} N(0, 1)$$

- równania te są matematycznie równoważne bo opisują ten sam proces stochastyczny – ruchy Browna
- proces Wienera opisuje zmiany zmiennej losowej X w czasie
- równanie dyfuzji określa związek pomiędzy pochodnymi fgp
- **rozkład fgp możemy określić generując wiele łańcuchów Markowa i uśredniając wyniki dla kolejnych chwil czasowych**

Równanie dyfuzji formalnie można wyprowadzić zaczynając rozważania od **prawa Ficka**, które wiąże gęstość prądu dla fgp z gradientem funkcji

$$J = -D \frac{\partial f(x, t)}{\partial x}, \quad D \geq 0$$

- prąd rozkładu gęstości prawdopodobieństwa płynie w kierunku przeciwnym do gradientu, aby go zmniejszyć

Prawo Ficka łączymy z równaniem ciągłości w postaci różniczkowej

$$\frac{\partial f(x, t)}{\partial t} + \frac{\partial J}{\partial x} = 0$$

co po wstawieniu jawnej postaci J daje równanie dyfuzji

$$\frac{\partial f(x, t)}{\partial t} = D \frac{\partial^2 f(x, t)}{\partial x^2}$$

Równanie Langevina → proces Ornsteina-Uhlenbecka

Punktem wyjścia będzie układ równań ruchu

$$\frac{d^2x}{dt^2} = \frac{F(t)}{m} \quad \longrightarrow \quad \begin{cases} \frac{dx}{dt} = V(t) \\ \frac{dV}{dt} = \frac{F(t)}{m} \end{cases} \quad \longrightarrow \quad \begin{cases} x(t+dt) - x(t) = V(t)dt \\ V(t+dt) - V(t) = \frac{F(t)}{m}dt \end{cases}$$

gdzie: m to masa, F(t) to siła

Langevin badając ruchy Browna zauważył, że zderzenia cząstki badanej z cząstkami płynu powinny prowadzić nie tylko do losowej zmiany kierunku prędkości, ale także do jej zmniejszania ze względu na **tarcie** (**lepkość**)

$$V(t+dt) - V(t) = \underbrace{-\gamma V(t)dt}_{\text{tarcie}} + \underbrace{\sqrt{\beta^2 dt} N(0,1)}_{\text{ruchy Browna}}$$

- γ to współczynnik tarcia
- równanie Langevina opisuje proces Ornsteina-Uhlenbecka z parametrem losowości β
- tarcie i fluktuacje prędkości to efekty komplementarne
– te same zderzenia z cząstkami płynu prowadzą do fluktuacji i dyssypacji energii

Mamy równanie w postaci różniczkowej, czy wystarczy ono do wygenerowania łańcucha Markowa?
Znajdźmy jego rozwiązanie ogólne (wartość oczekiwana i wariancja) i porównajmy oba podejścia.

Dla procesu Wienera nie było różnicy – każda zmienna miała identyczną wagę,
w procesie O-U zmienne mogą mieć inne wagi (tarcie) istotne jest dokładne uwzględnienie tego efektu.

Obliczmy wartość **oczekiwaną prędkości**

$$\langle V(t + dt) - V(t) \rangle = \langle -\gamma V(t)dt + \sqrt{\beta^2 dt} N(0, 1) \rangle$$

operacja jest liniowa

$$\langle V(t + dt) \rangle - \langle V(t) \rangle = \langle -\gamma V(t)dt \rangle + \langle \sqrt{\beta^2 dt} N(0, 1) \rangle$$

dzielmy przez dt i uśredniamy

$$\frac{d\langle V(t) \rangle}{dt} = -\gamma \langle V(t) \rangle \qquad \langle N(0, 1) \rangle = 0$$

rozwiązujemy RRZ

$$\langle V(t) \rangle = v_0 e^{-\gamma t}$$

- wartość oczekiwana prędkości

wariancja wymaga obliczenia

$$\text{var}\{V(t)\} = \langle V^2(t) \rangle - \langle V(t) \rangle^2$$

gdzie korzysta się z

$$d[V^2(t)] = [V(t + dt)]^2 - [V(t)]^2 = [V(t)(1 - \gamma dt) + \sqrt{\beta^2 dt} N(0, 1)]^2 - [V(t)]^2 = \dots$$

$$d\langle V^2(t) \rangle = -2\langle V^2(t) \rangle \gamma dt + 2 \underbrace{\langle V(t) N(0, 1) \rangle}_{\langle V(t) \rangle \langle N(0, 1) \rangle = 0} \sqrt{\beta^2 dt} + \beta^2 dt$$

$$\frac{d}{dt} \langle V^2(t) \rangle = -2\gamma \langle V^2(t) \rangle + \beta^2 \quad \implies \quad \langle V^2(t) \rangle = v_0^2 e^{-2\gamma t} + \frac{\beta^2}{2\gamma} (1 - e^{-2\gamma t})$$

Możemy teraz połączyć oba wyniki

$$\text{var}\{V(t)\} = \langle V^2(t) \rangle - \langle V(t) \rangle^2 = \frac{\beta^2}{2\gamma} (1 - e^{-2\gamma t})$$

$$\langle V(t) \rangle = v_0 e^{-\gamma t}$$

dla $t=0$ (początek procesu)

$$\text{var}\{V(t)\} = 0$$

Mamy zatem proces dla którego określiliśmy wartość oczekiwaną i wariancję, a fluktuacje mają rozkład normalny. Możemy zapisać wynik dla dowolnego t

$$V(t) = N \left(v_0 e^{-\gamma t}, \sqrt{\frac{\beta^2}{2\gamma} (1 - e^{-2\gamma t})} \right), \quad t \in [0, \infty)$$

i przez analogię do procesu Wienera zapiszemy fgp procesu Ornsteina-Uhlenbecka (funkcję rozkładu)

$$f(V, t) = \frac{1}{\sqrt{2\pi \frac{\beta^2}{2\gamma} (1 - e^{-2\gamma t})}} \exp \left[-\frac{(V - V_0 e^{-\gamma t})^2}{2 \frac{\beta^2}{2\gamma} (1 - e^{-2\gamma t})} \right]$$

- w równaniu występują oba parametry charakterystyczne dla procesu: β i γ
- dla $t=0$ $f(V, t)$ przechodzi w deltę Diraca
- w długim czasie wpływ czynnika γ zanika, dyfuzja dominuje – efekty są przeciwstawne (β jest skalowane eksponentą)

Skoro mamy tarcie w układzie, to możemy też dodać pewną siłę o stałej amplitudzie (elektryczną grawitacyjną), której istnienie w połączeniu ze zmienną siłą tarcia musi prowadzić do stałej prędkości – **prędkości dryfu** $V_d = \text{const}$

$$V_d = \text{const}$$

$$V(t + dt) - V(t) = -\gamma[V(t) - V_d]dt + \sqrt{\beta^2 dt}N(0, 1)$$

- gdy $V(t)$ zbliża się do prędkości dryfu, pierwszy wyraz po prawej stronie jest bliski 0 + fluktuacje
- fluktuacje są wynikiem ruchu termicznego cząsteczek ośrodka możemy je powiązać → **tw. dyssypacyjno-fluktuacyjne**

Twierdzenie fluktuacyjno-dyssypacyjne

Ciało poruszające się ze stałą prędkością (V_d) w płynie oddziałuje z cząsteczkami ośrodka. Przechodząc do układu środka masy ciała, fluktuacje te będą wpływać na energię kinetyczną

$$\delta E_k^{CM} = \lim_{t \rightarrow \infty} \frac{m \text{var}\{V(t)\}}{2} = \frac{m\beta^2}{4\gamma} \quad (\text{m – masa ciała})$$

te fluktuacje są miarą średniej energii termicznej cząstek, wobec tego z zasady ekwipartycji energii dostajemy relację

$$\frac{m\beta^2}{4\gamma} = \frac{kT}{2} \implies \frac{\beta^2}{2\gamma} = \frac{kT}{m}$$

m – masa obiektu
k – stała Boltzmannna
T - temperatura

wyrażenie wiąże fluktuacje (β) z dyssypacją (γ) w stanie równowagi termicznej ($t \rightarrow \infty$)

- **jeśli obserwujemy dyssypację energii to występują też fluktuacje**
- **natomiast fluktuacje mogą występować bez dyssypacji (brak prędkości dryfu)**

Przykład 1

Stalowa kulka opada ze „stałą” prędkością w rurze wypełnioną gliceryną (ciecz lepka) o współczynniku lepkości η

Stałą tarcia można wyznaczyć z prawa Stokes’a, a następnie parametr fluktuacyjny D

$$\gamma = \frac{6\eta r}{m} \quad \Rightarrow \quad \beta = \sqrt{\frac{12\pi\eta r kT}{m^2}}$$

- fluktuacje powodują, że nie możemy dokładnie określić położenia kulki w czasie

Przykład 2

Badana cząsteczka zderza się z cząsteczkami gazu – jak można określić parametr fluktuacyjny β ?

prędkość termiczna
cząsteczek gazu

$$v_{av} = \sqrt{\frac{kT}{m_0}}$$

gęstość gazu

n

przekrój czynny
na zderzenie

σ

częstość zderzeń

$$\nu = v_{av} \cdot n \cdot \sigma$$

$$\gamma = \frac{m_0}{m} \nu \quad \Rightarrow \quad \beta = \sqrt{\frac{2m_0 \nu kT}{m^2}}$$

- dostaliśmy prostą zależność opisującą współczynnik dyfuzji

Symulacja MC procesu O-U dla prędkości

Mamy dwa równania - które wybrać jako podstawę algorytmu MC?

$$(*) \quad V(t + dt) - V(t) = -\gamma V(t)dt + \frac{F(t)}{m}dt + \sqrt{\beta^2 dt} N(0, 1)$$

$$F(t) = \text{const} \quad \longrightarrow \quad V_d = \frac{F}{\gamma m}$$

$$(**) \quad V(t) = N \left(v_0 e^{-\gamma t} + V_d(1 - e^{-\gamma t}), \sqrt{\frac{\beta^2}{2\gamma} (1 - e^{-2\gamma t})} \right), \quad t \in [0, \infty)$$

Zwróćmy najpierw uwagę na istotną różnicę:

(*) jest prawdziwe dla $dt \rightarrow 0$,

(**) jest prawdziwe dla dowolnej chwili t

Równanie (*) możemy przekształcić do (**), ale powyższe założenie musi być spełnione

$$\lim_{dt \rightarrow 0} V(t + dt) = V(t)(1 - \gamma dt) + \dots = V(t) \exp(-\gamma t) + \dots$$

Z drugim równaniem tego problemu nie ma, t jest dowolnie duże. Możemy zatem utworzyć łańcuch Markowa dla procesu O-U

$$V(t_0) \rightarrow V(t_1) \rightarrow V(t_2) \rightarrow \dots$$

załóżmy jeszcze stały krok czasowy $t_{i+1} - t_i = \Delta t$

Zauważmy też że każda zmienna losowa $V(t_i)$ stanowi punkt startowy do wykonania kolejnego kroku.

$$V(t_{i+1}) = V(t_i)e^{-\gamma \Delta t} + \frac{F}{\gamma m}(1 - e^{-\gamma \Delta t}) + \sqrt{\frac{\beta^2}{2\gamma} (1 - e^{-2\gamma \Delta t})} N(0, 1)$$

- generator łańcucha Markowa dla prędkości w procesie O-U i dowolnego kroku czasowego

Pseudokod dla łańcucha Markowa w procesie O-U (prędkość)

```
inicjalizacja:  $V = V_0, \gamma, \beta, \Delta t$   
do i=1 to n  
     $U_1 \sim N(0, 1)$   
     $V = V \cdot e^{-\gamma \Delta t} + \frac{F}{\gamma m} (1 - e^{-\gamma \Delta t}) + \sqrt{\frac{\beta^2}{2\gamma} (1 - e^{-2\gamma \Delta t})} U_1$   
end do
```

Symulacja procesu z dwiema zmiennymi

Wiemy jak modelować proces O-U dla pojedynczej zmiennej (V).

Wróćmy do pierwotnego problemu, równania ruchu, które zawiera dwie zmienne: x i V .

Zmienne te są skorelowane.

równanie

$$V(t + dt) - V(t) = -\gamma V(t)dt + \sqrt{\beta^2 dt} N(0, 1)$$

rozwiązanie

$$V(t) = N \left(v_0 e^{-\gamma t}, \sqrt{\frac{\beta^2}{2\gamma} (1 - e^{-2\gamma t})} \right), \quad t \in [0, \infty)$$

$$V(t_{i+1}) = V(t_i) e^{-\gamma \Delta t} + \sqrt{\frac{\beta^2}{2\gamma} (1 - e^{-2\gamma \Delta t})} N(0, 1)$$

równanie

$$x(t + dt) - x(t) = V(t)dt$$

rozwiązanie

????????????????????

Gdzie leży problem?

Naiwnie moglibyśmy pomyśleć, że podczas wykonywania jednego kroku prędkość jest stała i próbować całkować drugie równanie. Ale prędkość ulega cały czas zmianie z powodu tarcia.

Jak należy postąpić?

- spodziewamy się że skoro zmienna $V(t)$ ma rozkład normalny to $x(t)$ również $X(t) = N(\langle X(t) \rangle, \text{var}\{X(t)\})$
- złożony rozkład prawdopodobieństwa dwóch zmiennych normalnych jest jednoznacznie zdefiniowany przez ich:
(1) wartości oczekiwane, (2) wariancje i (3) kowariancję
- należy obliczyć (1), (2) dla $x(t)$ oraz (3) dla x i V

- wartość oczekiwana $X(t)$

$$X(t + dt) - X(t) = V(t)dt \quad \longrightarrow \quad dX = Vdt \quad \longrightarrow \quad \frac{d\langle X \rangle}{dt} = \langle V \rangle$$

$$\langle X(t) \rangle = x_0 + \frac{V_0}{\gamma}(1 - e^{-\gamma t}) + \frac{V_d}{\gamma}(\gamma t + e^{-\gamma t} - 1)$$

- wariancja

$$\text{var}\{X\} = \langle X^2 \rangle - \langle X \rangle^2$$

$$dX = Vdt$$

$$dX^2(t) = X^2(t + dt) - X^2(t) = (X + dX)^2 - X^2 = 2XdX + (dX)^2 = 2XVdt + (Vdt)^2$$

$$\frac{d\langle X^2 \rangle}{dt} = 2\langle XV \rangle \quad \longrightarrow \quad \frac{d\text{var}\{X\}}{dt} = \frac{d}{dt}[\langle X^2 \rangle - \langle X \rangle^2]$$

$$\frac{d\text{var}\{X\}}{dt} = \frac{d}{dt}[\langle X^2 \rangle - \langle X \rangle^2] = \frac{d\langle X^2 \rangle}{dt} - 2\langle X \rangle \frac{d\langle X \rangle}{dt} = 2\langle XV \rangle - 2\langle X \rangle \langle V \rangle = 2\text{cov}\{X, V\}$$

- czyli wariancja zmiennej $X(t)$ zależy od jej kowariancji z $V(t)$

Aby obliczyć $\text{cov}(X, V)$ należy policzyć jej różniczkę, zachować wyrazy najniższego rzędu, po podzieleniu przez dt dostaniemy RRZ, które należy rozwiązać

$$d\text{cov}\{X, V\} = \dots \quad \Longrightarrow \quad \frac{d}{dt}\text{cov}\{X, V\} = -\gamma\text{cov}\{X, V\} + \text{var}\{V\}$$

kowariancja

$$\text{cov}\{X, V\} = \frac{\beta^2}{2\gamma^2}(1 - 2e^{-\gamma t} + e^{-2\gamma t})$$

Rozwiązując RRZ dla wariancji dostajemy

$$\frac{dvar\{X\}}{dt} = \frac{d}{dt}[\langle X^2 \rangle - \langle X \rangle^2]$$

$$var\{X\} = \frac{\beta^2}{\gamma^2} \left[t - \frac{2}{\gamma}(1 - e^{-\gamma t}) + \frac{1}{2\gamma}(1 - e^{-2\gamma t}) \right]$$

$$\langle X(t) \rangle = x_0 + \frac{V_0}{\gamma}(1 - e^{-\gamma t}) + \frac{V_d}{\gamma}(\gamma t + e^{-\gamma t} - 1)$$

$$X(t) = N(\langle X(t) \rangle, var\{X(t)\})$$

Możemy skonstruować łańcuch Markowa dla pary (X,V)

$$X(t + \Delta t) = a_0 + a_1 N(0, 1) + a_2 N(0, 1)$$

$$V(t + \Delta t) = b_0 + b_1 N(0, 1)$$

$$a_0 = \langle X(t + \Delta t) \rangle$$

$$b_0 = \langle V(t + \Delta t) \rangle$$

$$b_1 = \sqrt{var\{V(t + \Delta t)\}}$$

$$a_1 = \frac{cov\{X(t + \Delta t), V(t + \Delta t)\}}{\sqrt{var\{V(t + \Delta t)\}}}$$

$$a_2 = \sqrt{var\{X(t + \Delta t)\} - \frac{cov^2\{X(t + \Delta t), V(t + \Delta t)\}}{var\{V(t + \Delta t)\}}}$$

- oczywiście rozwiązania w kolejnej chwili czasowej $t+\Delta t$ (współczynniki a,b) znajdujemy biorąc jako punkty startowe rozwiązania z chwili poprzedniej t