

Rozwiązywanie układów równań liniowych metodami błędzenia przypadkowego

- **Wprowadzenie.**
- **Błądzenie przypadkowe.**
- **Łańcuch (proces) Markowa.**
- **Układy równań liniowych – sformułowanie zagadnienia.**
- **Metody von Neumanna–Ulama.**
 - ▷ **Metoda podstawowa.**
 - ▷ **Metoda dualna.**
 - ▷ **Uogólnienia.**
- **Metody Wasowa.**

▶ PRZYKŁAD 1:

Sumę: $S = A + B$ możemy przedstawić w postaci:

$$S = p \frac{A}{p} + (1 - p) \frac{B}{1 - p}, \quad 0 < p < 1,$$

i interpretować jako wartość oczekiwaną zmiennej losowej:

$$W = \begin{cases} \frac{A}{p} & \text{z prawdopodobieństwem } p, \\ \frac{B}{1-p} & \text{z prawdopodobieństwem } 1 - p. \end{cases}$$

▷ Schemat:

Zmienną losową W losujemy N -razy i obliczamy:

$$\hat{S} = \frac{1}{N} \sum_{i=1}^N W_i,$$

gdzie W_i – wynik i -tego losowania.

⇒ \hat{S} jest estymatorem zgodnym i nieobciążonym sumy S .

► PRZYKŁAD 2:

Niech: X i Y – nieznanne wielkości powiązane następującymi zależnościami:

$$\begin{cases} X = pY + (1 - p)A, \\ Y = qX + (1 - q)B, \end{cases} \quad A, B, p, q \text{ – ustalone oraz } 0 \leq p, q \leq 1.$$

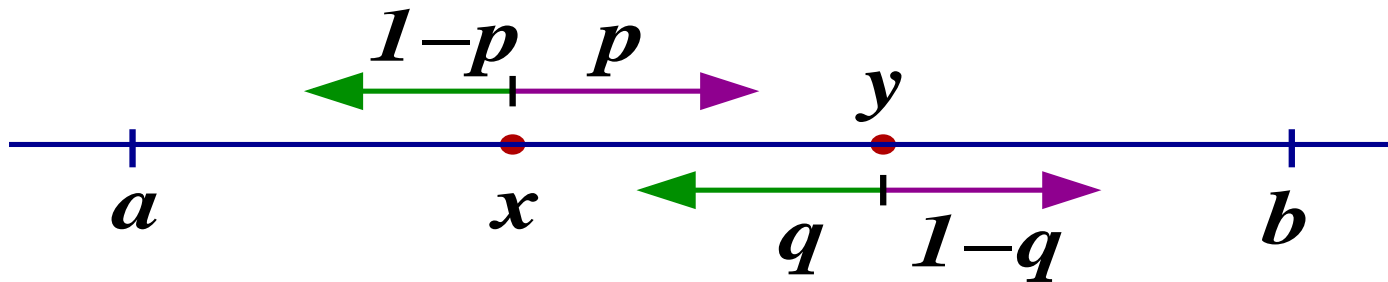
● **Zadanie:** Oszacować wartość zmiennej X .

▷ Schemat:

1. Losujemy pierwszy składnik pierwszego równania z prawdopodobieństwem p , a drugi z prawdopodobieństwem $(1 - p)$.
2. Jeżeli wylosujemy drugi składnik, to za wynik losowania W przyjmujemy wartość A .
3. Jeżeli wylosujemy pierwszy składnik, to przechodzimy do drugiego równania i z prawdopodob. q losujemy pierwszy składnik, a z prawdopodob. $(1 - q)$ drugi składnik tego równania.
4. Jeżeli wylosujemy drugi składnik, to za wynik losowania W przyjmujemy wartość B .
5. Jeżeli wylosujemy pierwszy składnik, to wracamy do kroku 1.
6. Losowanie kontynuujemy dopóty, dopóki nie wylosujemy jednej ze znanych wartości A lub B .

Powyższy schemat losowania powtarzamy N razy i jako **oszacowanie** wartości X przyjmujemy:

$$\hat{X} = \frac{1}{N} \sum_{i=1}^N W_i, \quad \hat{\sigma}_{\hat{X}} = \frac{1}{\sqrt{N-1}} \sqrt{\frac{1}{N} \sum_{i=1}^N W_i^2 - \hat{X}^2}.$$



Na początku znajdujemy się w punkcie x i możemy wędrować z punktu do punktu według następującego „regulaminu”:

- Z punktu x z prawdopodobieństwem p przechodzimy do punktu y lub z prawdopodobieństwem $(1 - p)$ do punktu a .
 - Z punktu y z prawdopodobieństwem q przechodzimy do punktu x lub z prawdopodobieństwem $(1 - q)$ do punktu b .
 - Wędrowkę kończymy wówczas, gdy znajdziemy się w punkcie a lub b .
 - Jeżeli wędrowkę zakończymy w punkcie a , to otrzymujemy nagrodę w wysokości A , natomiast jeżeli zakończymy ją w punkcie b , otrzymujemy nagrodę w wysokości B .
 - X jest oczekiwaną nagrodą, jaką otrzymamy startując z punktu x (analogicznie Y jest oczekiwaną nagrodą przy starcie z punktu y).
- Powyższą wędrowkę nazywamy **błądzeniem przypadkowym** po punktach zbioru $\{a, x, y, b\}$.
- ▷ Przedstawiony schemat możemy traktować jako metodę rozwiązania układu dwóch równań z dwiema niewiadomymi: X i Y .

- Weźmy układ o skończonej lub przeliczalnej liczbie możliwych stanów S_1, S_2, S_3, \dots
 - Niech X_t będzie stanem tego układu w chwili czasu t .
 - Rozważmy przypadek dyskretnego czasu i kolejne chwile oznaczmy przez: $1, 2, 3, \dots$
- $\Rightarrow X_t$ jest zmienną losową i możemy zdefiniować prawdopodobieństwa warunkowe:

$$\mathcal{P}(X_t = S_j | X_{t_1} = S_{i_1}, X_{t_2} = S_{i_2}, \dots, X_{t_n} = S_{i_n}).$$

- Układ jest **łańcuchem (procesem) Markowa** jeżeli rozkład X_t jest **niezależny od wszystkich poprzednich stanów z wyjątkiem ostatniego** X_{t-1} , tzn.

$$\mathcal{P}(X_t = S_j | X_{t-1} = S_{i_{t-1}}, \dots, X_2 = S_{i_2}, X_1 = S_{i_1}) = \mathcal{P}(X_t = S_j | X_{t-1} = S_{i_{t-1}}).$$

- ▷ Powyższe można łatwo rozszerzyć na stany ciągłe przez zastąpienie prawdopodobieństw funkcjami gęstości prawdopodobieństwa.

- **Przykłady łańcuchów (procesów) Markowa (błądzenie przypadkowe):**

Ruchy Browna, dyfuzja w gazach, „spacer pijanego marynarza ulicami miasta portowego” itd.

- Metody Monte Carlo oparte o procesy Markowa stosuje się do rozwiązywania układów równań liniowych, równań całkowych, równań różniczkowych cząstkowych, znajdowania wartości własnych, odwracania macierzy, optymalizacji itd.

Rozważmy układ równań liniowych:

$$\mathbf{A} \vec{x} = \vec{b}, \quad \det \mathbf{A} \neq 0, \quad (1)$$

gdzie: $\mathbf{A} = (a_{ij}), i, j = 1, 2, \dots, n$ – dana macierz, $\vec{b} = (b_1, b_2, \dots, b_n)$ – dany wektor,
 $\vec{x} = (x_1, x_2, \dots, x_n)$ – wektor niewiadomych.

► Ozn. $\vec{x}^0 = (x_1^0, x_2^0, \dots, x_n^0)$ – szukane rozwiązanie.

Powyższy układ równań przekształcamy do postaci iteracyjnej:

$$\vec{x} = \vec{a} + \mathbf{H} \vec{x}, \quad (2)$$

gdzie: $\mathbf{H} = (h_{ij}), i, j = 1, \dots, n$ jest daną macierzą, a $\vec{a} = (a_1, \dots, a_n)$ jest danym wektorem.

• Zakładamy, że norma macierzy \mathbf{H} :

$$\|\mathbf{H}\| \equiv \max_{1 \leq i \leq n} \sum_{j=1}^n |h_{ij}| < 1. \quad (3)$$

¶ **Dygresja:** Każdy układ równań postaci (1) można sprowadzić do układu (2) z warunkiem (3).

Przedstawiamy macierz $\mathbf{A} = \mathbf{V} - \mathbf{W}$ tak, że macierz \mathbf{V}^{-1} można łatwo obliczyć \rightarrow układ (1):

$$(\mathbf{V} - \mathbf{W}) \vec{x} = \vec{b}$$

Mnożąc obie strony równania lewostronnie przez macierz \mathbf{V}^{-1} dostajemy:

$$\vec{x} = \underbrace{\mathbf{V}^{-1} \vec{b}}_{\vec{a}} + \underbrace{\mathbf{V}^{-1} \mathbf{W}}_{\mathbf{H}} \vec{x}.$$

Układ równań (2) przedstawmy w postaci:

$$(\mathbf{I} - \mathbf{H}) \vec{x} = \vec{a}, \quad (4)$$

gdzie: $\mathbf{I} = (\delta_{ij})$ – macierz jednostkowa (δ_{ij} – delta Kroneckera, $= 1, 0$ dla $i = j, i \neq j$).

► Rozwiązanie powyższego układu można przedstawić w postaci szeregu Neumanna:

$$\vec{x}^0 = (\mathbf{I} - \mathbf{H})^{-1} \vec{a} = \vec{a} + \mathbf{H} \vec{a} + \mathbf{H}^2 \vec{a} + \mathbf{H}^3 \vec{a} + \dots \quad (5)$$

⇒ Dla i -tej składowej rozwiązania mamy:

$$\begin{aligned} x_i^0 = & a_i + \sum_{j=1}^n h_{ij} a_j + \sum_{j_1=1}^n \sum_{j_2=1}^n h_{ij_1} h_{j_1,j_2} a_{j_2} + \dots \\ & + \sum_{j_1=1}^n \dots \sum_{j_k=1}^n h_{ij_1} h_{j_1,j_2} \dots h_{j_{k-1},j_k} a_{j_k} + \dots \end{aligned} \quad (6)$$

▷ Sformułujemy zagadnienie probabilistyczne w postaci pewnego **łańcucha (procesu) Markowa (błądzenie przypadkowe)**, a następnie pokażemy, że jego rozwiązanie da się wyrazić w formie wzoru (6) – będzie to dowodem na równoważność tego zagadnienia z problemem rozwiązania układu równań (1).

I. Przypadek: $h_{ij} \geq 0, i, j = 1, \dots, n$

▷ Dodatkowo definiujemy:

$$h_{i,0} \equiv 1 - \sum_{j=1}^n h_{ij} > 0, \quad \text{bo} \quad \sum_{j=1}^n h_{ij} < 1.$$

Zbiór indeksów $\{0, 1, \dots, n\}$ nazwiemy **zbiorem stanów**, a każdy jego element – **stanem**.

Niech: X – punkt, który „wędruje” po zbiorze stanów,

a i_t ($i_t = 0, 1, \dots, n; t = 0, 1, \dots$) – stan, w którym punkt X znajduje się w chwili t .

► Regulamin wędrówki punktu X :

1. W chwili $t = 0$ punkt X znajduje się w stanie i_0 .
 2. Jeżeli w chwili t punkt X znajduje się w stanie i_t ($i_t \neq 0$), to w chwili $(t + 1)$ znajdzie się w stanie i_{t+1} z prawdopodobieństwem równym elementowi $h_{i_t, i_{t+1}}$ macierzy \mathbf{H} . Przejścia punktu X między kolejnymi stanami są **niezależne**.
 3. Jeżeli w pewnej chwili t punkt X znajdzie się w stanie 0, błędzenie punktu zostaje **zakończone**.
- ▷ Drogę: $\gamma \equiv (i_0, i_1, i_2, \dots, i_k, 0), i_j \neq 0, 0 \leq j \leq k$ nazwiemy **trajektorią** punktu X .
- ▷ Każdej trajektorii γ punktu X przypisujemy liczbę:

$$X(\gamma) = X(i_0, i_1, i_2, \dots, i_k, 0) = \frac{a_{i_k}}{h_{i_k, 0}}.$$

- ▶ $X(\gamma)$ – zmienna losowa o wartościach ze zbioru: $\{a_1/h_{1,0}, a_2/h_{2,0}, \dots, a_n/h_{n,0}\}$.
 → Prawdopodobieństwo, że $X(\gamma) = a_j/h_{j,0}$ jest równe prawdopodobieństwu, że **ostatnim niezerowym** stanem trajektorii γ jest stan j .

- ▶ Wartość oczekiwana zmiennej losowej $X(\gamma)$ pod warunkiem, że trajektoria rozpoczyna się w ustalonym punkcie $i_0 = s$:

$$E\{X(\gamma)|i_0 = s\} = \sum_{k=0}^{\infty} \sum_{\{\gamma_k\}} X(\gamma)P(\gamma),$$

gdzie γ_k oznacza trajektorię o długości k rozpoczynającą się w punkcie $i_0 = s$,
 a $P(\gamma)$ – prawdopodobieństwo pojawienia się takiej trajektorii.

Zatem:

$$E\{X(\gamma)|i_0 = s\} = \sum_{\{\gamma_0\}} X(\gamma)P(\gamma) + \sum_{\{\gamma_1\}} X(\gamma)P(\gamma) + \dots + \sum_{\{\gamma_k\}} X(\gamma)P(\gamma) + \dots$$

- ▷ Rozważymy poszczególne elementy powyższego szeregu.

$\{\gamma_0\}$: Jedyna trajektoria: $\gamma_0 = (i_0 = s, 0)$, a $P(\gamma_0) = h_{s,0}$ oraz $X(\gamma_0) = a_s/h_{s,0}$,

$$\Rightarrow \sum_{\{\gamma_0\}} X(\gamma)P(\gamma) = \frac{a_s}{h_{s,0}} h_{s,0} = a_s.$$

$\{\gamma_1\}$: Możliwe trajektorie: $\gamma_1 = (i_0 = s, i_1, 0)$, $i_1 \neq 0$, a $P(\gamma_1) = P(s, i_1, 0) = h_{s,i_1} h_{i_1,0}$ oraz $X(\gamma_1) = a_{i_1}/h_{i_1,0}$,

$$\Rightarrow \sum_{\{\gamma_1\}} X(\gamma)P(\gamma) = \sum_{i_1=1}^n \frac{a_{i_1}}{h_{i_1,0}} h_{s,i_1} h_{i_1,0} = \sum_{i_1=1}^n h_{s,i_1} a_{i_1}.$$

$\{\gamma_2\}$: Możliwe trajektorie: $\gamma_2 = (i_0 = s, i_1, i_2, 0)$, $i_1 \neq 0$, $i_2 \neq 0$,

a $P(\gamma_2) = P(s, i_1, i_2, 0) = h_{s,i_1} h_{i_1,i_2} h_{i_2,0}$ oraz $X(\gamma_2) = a_{i_2}/h_{i_2,0}$,

$$\Rightarrow \sum_{\{\gamma_2\}} X(\gamma)P(\gamma) = \sum_{i_1=1}^n \sum_{i_2=1}^n \frac{a_{i_2}}{h_{i_2,0}} h_{s,i_1} h_{i_1,i_2} h_{i_2,0} = \sum_{i_1=1}^n \sum_{i_2=1}^n h_{s,i_1} h_{i_1,i_2} a_{i_2}.$$

$\{\gamma_k\}$: Analogicznie dostajemy:

$$\sum_{\{\gamma_k\}} X(\gamma)P(\gamma) = \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_k=1}^n h_{s,i_1} h_{i_1,i_2} \dots h_{i_{k-1},i_k} a_{i_k}.$$

► Ostatecznie po zsumowaniu:

$$\begin{aligned}
 E\{X(\gamma) | i_0 = s\} = & a_s + \sum_{i_1=1}^n h_{s,i_1} a_{i_1} + \sum_{i_1=1}^n \sum_{i_2=1}^n h_{s,i_1} h_{i_1,i_2} a_{i_2} + \dots \\
 & + \sum_{i_1=1}^n \sum_{i_2=1}^n \dots \sum_{i_k=1}^n h_{s,i_1} h_{i_1,i_2} \dots h_{i_{k-1},i_k} a_{i_k} + \dots
 \end{aligned} \tag{7}$$

▷ Porównując powyższe wyrażenie z wyrażeniem (6) na rozwiązanie układu równań liniowych otrzymujemy:

$$x_i^0 = E\{X(\gamma) | i_0 = i\}$$

Podsumowanie:

Zatem rozwiązanie układu równań liniowych postaci (2) sprowadza się do obliczenia wartości oczekiwanej zmiennej losowej $X(\gamma)$ w procesie błądzenia przypadkowego. Błędy otrzymanych wyników szacujemy na podstawie centralnego twierdzenia granicznego, w oparciu o wyliczane praktyczne odchylenie standardowe (podobnie jak przy obliczaniu całek metodami Monte Carlo).

PRZYKŁAD:

Dany jest następujący układ równań:

$$\vec{x} = \begin{bmatrix} 1.5 \\ -1.0 \\ 0.7 \end{bmatrix} + \begin{bmatrix} 0.2 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.2 \\ 0.3 & 0.1 & 0.1 \end{bmatrix} \vec{x}. \quad (8)$$

Jego rozwiązaniem jest: $\vec{x}^0 = (2.154303, 0.237389, 1.522255)$.

Ponieważ

$$\|\mathbf{H}\| \equiv \max_{1 \leq i \leq 3} \sum_{j=1}^3 |h_{ij}| = \max\{0.6, 0.9, 0.5\} = 0.9 < 1 \quad \text{oraz} \quad h_{ij} > 0, \quad i, j = 1, 2, 3,$$

więc możemy zastosować opisaną metodę von Neumanna–Ulama.

▷ Macierz prawdopodobieństw h_{ij} ma postać:

$i \backslash j$	1	2	3	0
1	0.2	0.3	0.1	0.4
2	0.4	0.3	0.2	0.1
3	0.3	0.1	0.1	0.5

▶ Sposób rozwiązania:

Stosujemy metodę von Neumanna–Ulama startując kolejno z $i_0 = 1, 2, 3$.

▷ Przykładowe rozwiązanie dla 25 trajektorii:

$$x_1^0 = 2.086 \pm 0.744.$$

⇒ **Ćw. N10.1:** Rozwiązać powyższy układ równań metodą von Neumanna–Ulama.

▷ Metoda podstawowa von Neumanna–Ulama wymaga powtórzenia schematu obliczeń dla każdej współrzędnej wektora rozwiązań \vec{x}^0 .

► **Metoda dualna von Neumanna–Ulama** umożliwia oszacowanie od razu **całego wektora** \vec{x}^0 .

Schemat:

Na zbiorze indeksów $\{0, 1, \dots, n\}$ określamy (dowolnie) rozkład prawdopodobieństwa:

$$q_1, q_2, \dots, q_n, \text{ tzn. } q_i > 0, i = 1, 2, \dots, n \text{ oraz } \sum_{i=1}^n q_i = 1.$$

1. Stan początkowy wędrującego punktu X losujemy według rozkładu q_i .
2. Jeżeli w chwili t punkt X znajduje się w stanie i_t ($i_t \neq 0$), to z prawdopodobieństwem $p(i_{t+1}|i_t) = h_{i_{t+1}, i_t}$ w chwili $(t + 1)$ znajdzie się w stanie i_{t+1} ; dla $i_{t+1} = 0$ definiujemy $h_{0, i_t} = 1 - \sum_{j=1}^n h_{j, i_t}$ i zakładamy, że $h_{0, i_t} > 0$ (ogólny przypadek będzie później).
▷ Teraz macierz prawdopodobieństw przejść jest macierzą transponowaną: \mathbf{H}^T .
3. Błądzenie punktu X po zbiorze stanów $\{0, 1, \dots, n\}$ kończy się z chwilą osiągnięcia stanu 0.
4. Trajektorii $\gamma = (i_0, i_1, \dots, i_k, 0)$, $i_j \neq 0$ dla $0 \leq j \leq k$, przyporządkowujemy wektor:

$$\vec{Y}(\gamma) = \frac{a_{i_0}}{q_{i_0} p(0|i_k)} \hat{\mathbf{e}}_{i_k} \in \mathbb{R}^n,$$

gdzie $\hat{\mathbf{e}}_{i_k}$ jest i_k -tym wersorem, tzn. jego i_k -ta składowa jest równa 1, a pozostałe są równe 0.

Kroki 1–4 powtarzamy N razy. → Oszacowaniem wektora \vec{x}^0 jest średnia: $\frac{1}{N} \sum \vec{Y}(\gamma)$.

Dowód:

Niech $Y_{(j)}(\gamma)$ oznacza j -tą składową wektora $\vec{Y}(\gamma)$. Wystarczy dowieść, że dla każdego j :

$$E\{Y_{(j)}(\gamma)\} = x_j^0.$$

Z definicji wektora $\vec{Y}(\gamma)$ mamy:

$$Y_{(j)}(i_0, i_1, \dots, i_k, 0) = \begin{cases} \frac{a_{i_0}}{q_{i_0} p(0|j)}, & i_k = j, \\ 0, & i_k \neq j. \end{cases}$$

Zatem wartość oczekiwana:

$$E\{Y_{(j)}(\gamma)\} = \sum_{\text{trajektorie}} \frac{a_{i_0}}{q_{i_0} p(0|j)} P(i_0, i_1, \dots, i_{k-1}, j, 0),$$

gdzie $P(i_0, i_1, \dots, i_{k-1}, j, 0)$ – prawdopodobieństwo zaobserwowania trajektorii

$(i_0, i_1, \dots, i_{k-1}, j, 0)$, $i_l \neq 0$ dla $0 \leq l < k$, $j \neq 0$, a sumowanie przebiega po wszystkich trajektoriach.

▷ Prawdopodobieństwo zaobserwowania trajektorii (na podstawie punktu 1 i 2):

$$P(i_0, i_1, \dots, i_{k-1}, j, 0) = q_{i_0} h_{i_1, i_0} h_{i_2, i_1} \dots h_{j, i_{k-1}} p(0|j).$$

► Ostatecznie dla wartości oczekiwanej dostajemy:

$$E\{Y_{(j)}(\gamma)\} = \sum_{k=0}^{\infty} \sum_{i_{k-1}=1}^n \dots \sum_{i_1=1}^n \sum_{i_0=1}^n h_{j, i_{k-1}} \dots h_{i_2, i_1} h_{i_1, i_0} a_{i_0} \stackrel{(6)}{=} x_j^0, \quad j = 1, \dots, n. \quad \square$$

PRZYKŁAD:

Dany jest następujący układ równań:

$$\vec{x} = \begin{bmatrix} 1.5 \\ -1.0 \\ 0.7 \end{bmatrix} + \begin{bmatrix} 0.2 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.1 \end{bmatrix} \vec{x}. \quad (9)$$

Jego rozwiązaniem jest: $\vec{x}^0 = (2, 0, 1)$.

► Przyjmijmy początkowy rozkład w postaci:

$$q_1 = q_2 = q_3 = \frac{1}{3}.$$

▷ Prawdopodobieństwa przejść $p(i_{t+1}|i_t)$:

$i_t \backslash i_{t+1}$	1	2	3	0
1	0.2	0.4	0.1	0.3
2	0.3	0.3	0.1	0.3
3	0.1	0.2	0.1	0.6

▷ Przykładowe rozwiązania:

N	Oszacowanie rozwiązania		
25	3.16	-1.72	0.60
50	2.02	-1.18	0.79
75	2.31	-0.79	0.61
100	2.63	-0.63	0.60

→ Wolna zbieżność!

⇒ Ćw. N10.2: Rozwiązać powyższy układ równań używając metody dualnej von Neumanna–Ulama.

II. Elementy h_{ij} macierzy H mogą być ujemne

Weźmy macierz $P = (p_{ij})$ o następujących własnościach:

$$p_{ij} \geq 0, \quad p_{ij} = 0 \text{ gdy } h_{ij} = 0, \quad p_{i0} = 1 - \sum_{j=1}^n p_{ij} > 0.$$

W celu oszacowania składowej x_i^0 rozwiązania układu równań (2) konstruujemy łańcuch Markowa (błądzenie przypadkowe) na zbiorze indeksów $\{0, 1, \dots, n\}$ z macierzą przejść P , gdzie każda trajektoria rozpoczyna się w stanie $i_0 = i$.

► Prawdopodobieństwo zaobserwowania trajektorii o długości k wynosi:

$$P(\gamma_k) = p_{i,i_1} p_{i_1,i_2} \cdots p_{i_{k-1},i_k} p_{i_k,0}.$$

• Trajektorii $\gamma_k = (i_0 = i, i_1, \dots, i_k, 0)$ przypisujemy zmienną losową:

$$X(\gamma_k) = v_{i,i_1} v_{i_1,i_2} \cdots v_{i_{k-1},i_k} \frac{a_{i_k}}{p_{i_k,0}},$$

gdzie:

$$v_{ij} = \begin{cases} h_{ij}/p_{ij}, & p_{ij} \neq 0, \\ 1, & p_{ij} = 0. \end{cases}$$

► Wartość oczekiwana zmiennej losowej $X(\gamma)$:

$$E\{X(\gamma_k)\} = \sum_{k=0}^{\infty} \sum_{\{\gamma_k\}} X(\gamma_k) P\{X(\gamma_k)\},$$

gdzie rozkład zmiennej losowej $X(\gamma)$ jest dany:

$$\begin{aligned} P\{X(\gamma_k)\} &= \mathcal{P}\{X(\gamma_k) = v_{i,i_1} v_{i_1,i_2} \cdots v_{i_{k-1},i_k} a_{i_k} / p_{i_k,0}\} \\ &= p_{i,i_1} p_{i_1,i_2} \cdots p_{i_{k-1},i_k} p_{i_k,0}. \end{aligned}$$

Ponieważ:

$$X(\gamma_k) P\{X(\gamma_k)\} = h_{i,i_1} h_{i_1,i_2} \cdots h_{i_{k-1},i_k} a_{i_k},$$

to:

$$\sum_{\{\gamma_k\}} X(\gamma_k) P\{X(\gamma_k)\} = \sum_{i_1=1}^n \cdots \sum_{i_k=1}^n h_{i,i_1} h_{i_1,i_2} \cdots h_{i_{k-1},i_k} a_{i_k}.$$

Stąd na podstawie wzoru (6) mamy:

$$E\{X(\gamma_k)\} = x_i^0. \quad \boxtimes$$

Schemat:

Ustalamy (dowolnie) macierz P o podanych wyżej własnościach.

1. W chwili początkowej punkt X umieszczamy w stanie $i_0 = i$.
2. Jeżeli w chwili t punkt X znajduje się w stanie i_t ($i_t \neq 0$), to z prawdopodobieństwem $p_{i_t, i_{t+1}}$ w chwili $(t + 1)$ znajdzie się w stanie i_{t+1} .
3. Błądzenie punktu zostaje **zakończone**, gdy znajdzie się on w stanie 0.
4. Zaobserwowanej trajektorii przypisujemy wartość $X(\gamma_k)$.

Kroki 1–4 powtarzamy N razy.

- Oszacowaniem rozwiązania x_i^0 jest średnia arytmetyczna zaobserwowanych wartości $X(\gamma_k)$.
- ▷ Aby otrzymać wektor rozwiązań \vec{x}^0 , powyższe obliczenia powtarzamy dla wszystkich jego składowych: $i = 1, \dots, n$.

Schemat:

- Na zbiorze indeksów $\{0, 1, \dots, n\}$ określamy (dowolnie) rozkład prawdopodobieństwa:

$$q_1, q_2, \dots, q_n, \text{ tzn. } q_i > 0, i = 1, 2, \dots, n \text{ oraz } \sum_{i=1}^n q_i = 1.$$

- Ustalamy (dowolnie) macierz P o podanych wyżej własnościach.

1. Stan początkowy wędrującego punktu X losujemy według rozkładu q_i .
2. Jeżeli w chwili t punkt X znajduje się w stanie i_t ($i_t \neq 0$), to z prawdopodobieństwem $p_{i_t, i_{t+1}}$ w chwili $(t + 1)$ znajdzie się w stanie i_{t+1} .
3. Błądzenie punktu X po zbiorze stanów $\{0, 1, \dots, n\}$ kończy się z chwilą osiągnięcia stanu 0.
4. Trajektorii $\gamma = (i_0, i_1, \dots, i_k, 0)$, $i_j \neq 0$ dla $0 \leq j \leq k$, przyporządkowujemy wektor:

$$\vec{Y}(\gamma) = \frac{a_{i_0}}{q_{i_0} p_{i_k, 0}} w_{i_0, i_1} w_{i_1, i_2} \dots w_{i_{k-1}, i_k} \hat{e}_{i_k} \in \mathbb{R}^n,$$

gdzie:

$$w_{ij} = \begin{cases} h_{ji}/p_{ij}, & p_{ij} \neq 0, \\ 1, & p_{ij} = 0. \end{cases}$$

a \hat{e}_{i_k} jest i_k -tym wersorem.

Kroki 1–4 powtarzamy N razy.

► Oszacowaniem rozwiązania \bar{x}^0 jest średnia arytmetyczna zaobserwowanych wektorów $\vec{Y}(\gamma)$.

Dowód:

Ponieważ:

$$\begin{aligned} \mathcal{P}\left\{\vec{Y}(\gamma) = \frac{a_{i_0}}{q_{i_0} p_{i_k,0}} w_{i_0,i_1} w_{i_1,i_2} \cdots w_{i_{k-1},i_k} \hat{e}_{i_k}\right\} &= \mathcal{P}\{\gamma_k = (i_0, i_1, \dots, i_k, 0)\} \\ &= q_{i_0} p_{i_0,i_1} p_{i_1,i_2} \cdots p_{i_{k-1},i_k} p_{i_k,0}, \end{aligned}$$

to wartość oczekiwana:

$$E\{\vec{Y}(\gamma)\} = \sum_{k=0}^{\infty} \sum_{\{\gamma_k\}} h_{i_1,i_0} h_{i_2,i_1} \cdots h_{i_k,i_{k-1}} a_{i_0} \hat{e}_{i_k}.$$

Stąd wartość oczekiwana warunkowa, przy warunku $i_k = j$, wynosi:

$$\begin{aligned} E\{\vec{Y}(\gamma) | i_k = j\} &= \sum_{k=0}^{\infty} \sum_{\{\gamma_k\}} h_{i_1,i_0} h_{i_2,i_1} \cdots h_{j,i_{k-1}} a_{i_0} \\ &= \sum_{k=0}^{\infty} \sum_{i_{k-1}=1}^n \cdots \sum_{i_1=1}^n \sum_{i_0=1}^n h_{j,i_{k-1}} \cdots h_{i_2,i_1} h_{i_1,i_0} a_{i_0} \stackrel{(6)}{=} x_j^0. \end{aligned}$$

Zatem wartość oczekiwana zmiennej losowej $\vec{Y}(\gamma)$:

$$E\{\vec{Y}(\gamma)\} = \vec{x}^0. \quad \boxtimes$$

▷ W metodzie von Neumanna–Ulama za pomocą zmiennej losowej:

$$X(\gamma) = v_{i_0, i_1} v_{i_1, i_2} \cdots v_{i_{k-1}, i_k} a_{i_k}$$

otrzymanej dla trajektorii $\gamma = (i_0, i_1, \dots, i_k, 0)$ szacuje się jeden składnik sumy (6).

► **Metoda Wasowa** korzysta z każdej zaobserwowanej trajektorii do oszacowania całej sumy x_i^0 .

1. Metoda podstawowa:

Rozważmy wszystkie początkowe odcinki trajektorii $\gamma = (i_0, i_1, \dots, i_k, 0)$, tzn.

$$(i_0), (i_0, i_1), (i_0, i_1, i_2), \dots, (i_0, i_1, \dots, i_k)$$

i każdemu odcinkowi (i_0, i_1, \dots, i_m) , $0 \leq m \leq k$ przyporządkujemy liczbę:

$$v_{i_0, i_1} v_{i_1, i_2} \cdots v_{i_{m-1}, i_m} a_{i_m}.$$

Dla danej trajektorii $\gamma = (i_0, i_1, \dots, i_k, 0)$ definiujemy zmienną losową:

$$X^*(\gamma) = \sum_{m=0}^k v_{i_0, i_1} v_{i_1, i_2} \cdots v_{i_{m-1}, i_m} a_{i_m}.$$

Można udowodnić (analogicznie jak dla metody podstawowej von Neumanna–Ulama), że dla ustalonego $i_0 = i$ wartość oczekiwana zmiennej losowej $X^*(\gamma)$ jest estymatorem nieobciążonym i -tej składowej rozwiązania \vec{x}^0 :

$$E\{X^*(\gamma) | i_0 = i\} = x_i^0.$$

2. Metoda dualna:

Każdemu odcinkowi (i_0, i_1, \dots, i_m) , $0 \leq m \leq k$ trajektorii $\gamma = (i_0, i_1, \dots, i_k, 0)$ przypisujemy wektor:

$$\frac{a_{i_0}}{q_{i_0}} w_{i_0, i_1} w_{i_1, i_2} \cdots w_{i_{m-1}, i_m} \hat{e}_{i_m},$$

a całej trajektorii γ zmienną losową:

$$\vec{Y}^*(\gamma) = \sum_{m=0}^k \frac{a_{i_0}}{q_{i_0}} w_{i_0, i_1} w_{i_1, i_2} \cdots w_{i_{m-1}, i_m} \hat{e}_{i_m},$$

Można pokazać, że wartość oczekiwana powyższej zmiennej losowej jest estymatorem nieobciążonym rozwiązania \vec{x}^0 .

- ▶ Szybsza zbieżność metod Wasowa w porównaniu z metodami von Neumanna–Ulama została pokazana jedynie dla przypadków szczególnych – nie jest to ogólna reguła!
- ▷ Halton (1962) zaproponował kilka metod sekwencyjnych, w których w każdym kroku obliczeń korzysta się z wyników uzyskanych w poprzednich etapach. Może to prowadzić do poprawy wydajności przedstawionych metod rozwiązywania układów równań liniowych.
- ⇒ **Ćw. N10.3:** Używając metody podstawowej Wasowa rozwiązać układ równań (8). Porównać jej zbieżność ze zbieżnością metody podstawowej von Neumanna–Ulama.