Generatory liczb pseudolosowych - rozkłady skorelowane w 2D

11 marca 2024

Na zajęciach przećwiczymy generowanie rozkładów dwuwymiarowych: sferycznie konturowany - normalny (gaussowski), jednorodny w kole 2D oraz transformację afiniczną i wyznaczanie i użycie macierzy kowariancji dla rozkładów skorelowanych.

1 Rozkłady 2D

1.1 rozkład sferycznie konturowany - normalny

Liczby pseudolosowe o rozkładzie normalnym generujemy metodą Boxa-Mullera. Algorytm dla rozkładu 2D jest następujący

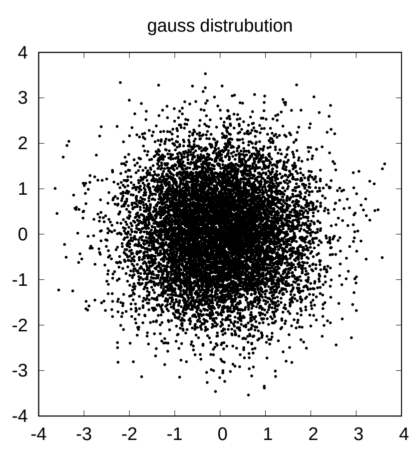
$$U_1 \sim U(0,1), \quad U_2 \sim U(0,1) \tag{1}$$

$$X = \sqrt{-2\,\ln(1-U_1)}\cos(2\pi U_2), \qquad X \sim N(0,1) \tag{2}$$

$$Y = \sqrt{-2\,\ln(1-U_1)}\sin(2\pi U_2), \qquad Y \sim N(0,1) \tag{3}$$

Wektory (X, Y) mają rozkład sferycznie konturowany ponieważ jego gęstość zależy tylko od odegłości od środka rozkładu

$$f(x,y) = f(x)f(y) = f(r), \quad r = \sqrt{x^2+y^2} \tag{4}$$



Rysunek 1: Rozkład normalny w 2D

1.2 Rozkład jednorodny w kole $K^2(0,1)$

Dysponując rozkładem sferycznie konturowanym możemy możemy teraz umieścić wylosowane punkty na obwodzie okręgu o promieniu jednostkowym normalizując zmienne

$$X' = \frac{X}{\sqrt{X^2+Y^2}} \tag{5}$$

$$Y' = \frac{Y}{\sqrt{X^2+Y^2}} \tag{6}$$

a następnie przesunąć je do środka okręgu. Aby rozkład w kole był jedorodny, skalujemy zmienne zmienną losową z rozkładu o fgp

$$h(r) = kr^{k-1}, \qquad r \in [0,1] \tag{7}$$

gdzie $k = 2$ to liczba wymiarów w naszym problemie

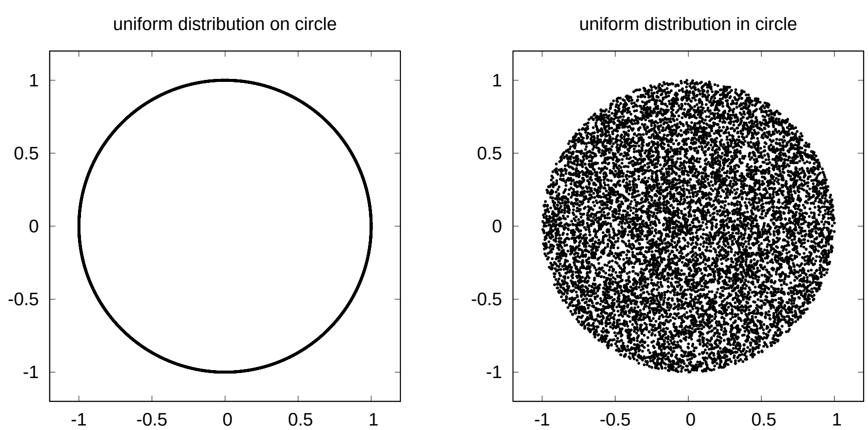
$$F(R) = R^2 = U_1 \sim U(0,1) \tag{8}$$

$$R = \sqrt{U_1} \tag{9}$$

$$X'' = RX' \tag{10}$$

$$Y'' = RY' \tag{11}$$

Wektory (X'', Y'') mają rozkład jednorodny w kole o promieniu jednostkowym.



Rysunek 2: Rozkład jednorodny na okręgu i w kole 2D.

1.3 Transformacja afiniczna: koło → elipsa

Rozkład dwywymiarowy możemy podać transformacji liniowej (afinicznej), która przekształci koło w elipsę. Docelowy kształt elipsy definiujemy podając wektory określające półosie główne

$$\vec{r}_1 = \begin{bmatrix} r_{11} \\ r_{21} \end{bmatrix} \qquad \vec{r}_2 = \begin{bmatrix} r_{12} \\ r_{22} \end{bmatrix} \tag{12}$$

Wektory te stanowią kolumny macierzy transformcji $A = [\vec{r}_1 | \vec{r}_2]$

$$A = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \tag{13}$$

Macierz A określa obrót oraz skalowanie wzdłuż półosi głównych, rozkład możemy też przesunąć o wektor $\vec{c}^T = [c_1, c_2]^T$. Transformację możemy zapisać w postaci macierzowej

$$\vec{X}' = A\vec{X} + \vec{c} \tag{14}$$

1.3.1 Wybór osi i skalowanie

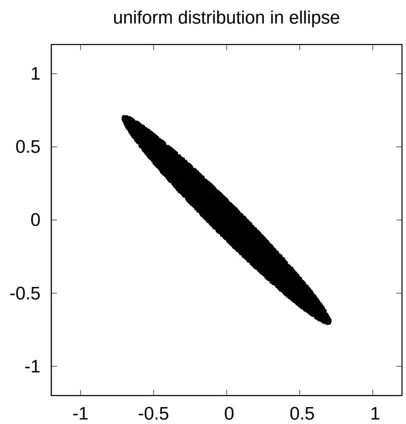
Ponieważ półosie główne elipsy muszą być ortogonalne tak jak wersory układu kartezjańskiego wystarczy więc tylko obrócić je o zadany kąt α przy użyciu macierzy obrotu R_α i przeskalować ich długości

$$\vec{r}_1 = b_1 R_\alpha \hat{e}_x, \qquad \hat{e}_x = [1,0]^T \tag{15}$$

$$\vec{r}_2 = b_2 R_\alpha \hat{e}_y, \qquad \hat{e}_y = [0,1]^T \tag{16}$$

$$R_\alpha = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}, \qquad \alpha \in [0, 2\pi] \tag{17}$$

gdzie: $b_1, b_2 > 0$ to współczynniki skalujące



Rysunek 3: Transformacja rozkładu jednorodnego w kole do skorelowanego w elipsie.

1.4 Wyznaczanie macierzy kowariancji

Macierz kowariancji ma postać

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix} \qquad \sigma_{xy} = \sigma_{yx} \tag{18}$$

Jej elementy liczymy następująco

$$\sigma_x^2 = \iint (x-\mu_x)^2 f(x,y)dxdy = \iint x^2 f(x,y)dxdy - \mu_x^2 = \langle x^2 \rangle - \langle x \rangle^2 \tag{19}$$

$$\begin{aligned} \sigma_{xy} = \sigma_{yx} &= \iint (x-\mu_x)(y-\mu_y) f(x,y)dxdy \\ &= \iint xy f(x,y)dxdy - \mu_x \mu_y \\ &= \langle xy \rangle - \langle x \rangle \langle y \rangle \end{aligned} \tag{20}$$

gdzie $\langle x \rangle$ oznacza wartość średnią z próby. Dla ciągu losowych wektorów

$$\{(x_1, y_1),\ (x_2, y_2), \ldots, (x_n, y_n)\} \tag{21}$$

możemy oszacować niezbędne wielkości do wyznaczenia elementów macierzy kowariancji

$$\begin{aligned}
\overline{x} &= \tfrac{1}{n} \sum_{i=1}^{n} x_i && (22)\\
\overline{x^2} &= \tfrac{1}{n} \sum_{i=1}^{n} x_i^2 && (23)\\
\overline{xy} &= \tfrac{1}{n} \sum_{i=1}^{n} x_i y_i && (24)\\
\sigma_x^2 &= \overline{x^2} - (\overline{x})^2 && (25)\\
\sigma_{xy} &= \overline{x\,y} - \overline{x} \cdot \overline{y} && (26)\\
& && (27)
\end{aligned}$$

Elementy macierzy kowariancji możemy wykorzystac do wyznaczenia współczynnika korelacji zmiennych

$$r_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2 \sigma_y^2}}, \qquad r \in [-1, 1] \tag{28}$$

1.5 Transformacja afiniczna a macierz kowariancji dla rozkładu gaussowskiego

Ogólna postać funkcji gęstości prawdopodobieństwa dla k-wymiarowego rozkładu normalnego $N^k(0,1)$

$$f(\vec{r}) = \frac{1}{(2\pi)^{k/2}\sqrt{\|D\|}} \exp\left(-\frac{\vec{r}^T D^{-1} \vec{r}}{2}\right) \tag{29}$$

Transformacja afiniczna zmiennych (bez przesunięcia $\vec{c} = \vec{0}$) ma postać

$$\vec{r}' = A\vec{r} \qquad \rightarrow \qquad \vec{r} = A^{-1}\vec{r}' \tag{30}$$

co po podstawieniu do fgp daje zmianę w wykładniku

$$\vec{r}D^{-1}\vec{r} = \vec{r}'^T (A^{-1})^T D^{-1} A^{-1} \vec{r}' = \vec{r}'^T \Sigma^{-1} \vec{r}' \tag{31}$$

Nowa macierz kowariancji

$$\Sigma = ADA^T \tag{32}$$

dla pierwotnego rozkładu $N^2(0,1)$ czyli

$$D = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{33}$$

ma prostą konstrukcję

$$\Sigma = AA^T \qquad \rightarrow \qquad \Sigma^{-1} = (A^T)^{-1} A^{-1} \tag{34}$$

Jeśli oznaczymy

$$A^{-1}\vec{r}' = \vec{z} \tag{35}$$

to losując wektory $\vec{z}$ z rozkładu $N^2(0,1)$ dostaniemy rozkład skorelowany

$$\vec{r}' = A\vec{z} \tag{36}$$

określony macierzą kowariancji $\Sigma = AA^T$.

2 Zadania do wykonania

1. Wylosować $n = 10^4$ punktów z dwuwymiarowego rozkładu normalnego $N^2(0,1)$ przy użyciu metody Boxa-Mullera. Sporządzić rysunek pokazujący położenia punktów.

2. Wygenerować $n = 10^4$ punktów wewnątrz koła o promieniu jednostkowym korzystając z rozkładu sferycznie konturowanego. Sporządzić rysunek pokazujący położenia punktów.

3. Dla kąta $\alpha = \pi/4$ oraz współczynników skalujących $b_1 = 1$ i $b_2 = 0.2$ przekształcić wersory układu kartezjańskiego w półosie główne elipsy $\vec{r}_1$ i $\vec{r}_2$. Skonstruować macierz transformacji A i wykonać transformację na próbce punktów o rozkładzie jednorodnym w kole jednostkowym. Sporządzić rysunek pokazujący położenia punktów dla nowego rozkładu. Wyznaczyć macierz kowariancji rozkładu po wykonaniu transformacji. Obliczyć wartość współczynnika korelacji r_{xy}.

4. Przy pomocy macierzy transformacji A z poprzedniego punktu wykonac transformację dla rozkładu normalnego $N^2(0,1)$. Sporządzić rysunek pokazujący położenia punktów dla nowego rozkładu. Wyznaczyć macierz kowariancji rozkładu po wykonaniu transformacji. Obliczyć wartość współczynnika korelacji r_{xy}.

5. W raporcie przedyskutować uzyskane wyniki.