2.
Powtórka z algebry i statystyki

1. Pojęcie normy, normy wektora [Kiełbasiński, Schwetlick]

wektor $\mathbf{x} \in R^d$

$$\mathbf{x} = (x_1, x_2, ..., x_d)^T \text{ – wektor, punkt w przestrzeni } d\text{-wymiarowej}$$

norma wektora – własności

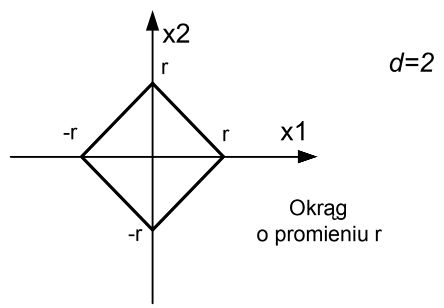
(1) $\|\mathbf{x}\| \geqslant 0$, $\|\mathbf{x}\| = 0$ tylko wtedy, gdy $\mathbf{x} = 0$

(2) $\|a\mathbf{x}\| = |a| \, \|\mathbf{x}\|$

(3) $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$

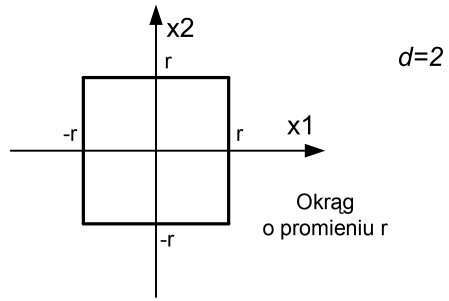
norma typu "suma"

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| + ... |x_d| = \sum_{i=1}^{d} |x_i| \text{ – nieujemna liczba rzeczywista (skalar)}$$



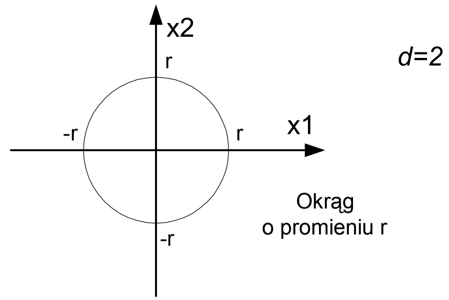
norma typu "maksimum"

$$\|\mathbf{x}\|_\infty = \max_{i=1,..,d} |x_i| \text{ – nieujemna liczba rzeczywista (skalar)}$$



norma euklidesowa

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T\mathbf{x}} = \left(x_1^2 + x_2^2 + ...x_d^2\right)^{1/2} = \sqrt{\sum_{i=1}^{d} x_i^2} \text{ – nieujemna liczba rzeczywista (skalar)}$$



2. Normy macierzy

norma macierzy $\mathbf{A} \in R^{m,d}$ indukowana przez normę wektora typu p

$$\|\mathbf{A}\|_p = \max\left\{\frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p} : \mathbf{x} \neq 0\right\} = \max\left\{\|\mathbf{Ax}\|_p : \|\mathbf{x}\|_p \leqslant 1\right\}$$

dla $p = 1$

$$\|\mathbf{A}\|_1 = \max_{j=1,\dots,d} \sum_{i=1}^{m} |a_{i,j}|$$

dla $p = 2$ **– tzw. norma spektralna macierzy**

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^T\mathbf{A})}$$

dla $p = \infty$

$$\|\mathbf{A}\|_\infty = \max_{i=1,\dots,m} \sum_{j=1}^{d} |a_{i,j}|$$

3. Statystyki opisowe zmiennych losowych [Klonecki]

ω – zdarzenie losowe, $\omega \in \Omega$ (Ω – tzw. zbiór zdarzeń elementarnych)
funkcja $X(\omega) \in R$ – zmienna losowa (tu skalar, może być wektorem)
dyskretne zmienne losowe – gdy zbiór wartości przyjmowanych przez $X(\omega)$ jest przeliczalny

$$\text{wartość oczekiwana} \quad : \quad EX = \sum_{\omega\in\Omega} X(\omega)P(\omega) \text{ – średnie } X \text{ ważone prawdopodobieństwami (wielkość nie losowa!)}$$

$$\text{wariancja } var X \;=\; E\left\{(X - EX)^2\right\} \text{ – średni kwadrat odchylenia } X \text{ od } EX \text{ ważony prawdopodobieństwami}$$

ciągłe zmienne losowe

$$f(x) \text{ – funkcja gęstości prawdopodobieństwa}$$

$$P(X \in (a,b)) = \int_a^b f(x)dx$$

$$EX = \int_{-\infty}^{\infty} xf(x)dx \qquad var X = \int_{-\infty}^{\infty} (x - EX)^2 f(x)dx$$

4. Popularne rozkłady

rozkład jednostajny (*ang. uniform distribution*) $\tilde{}U[a,b]$

$$f(x) = \begin{cases} \frac{1}{b-a}, \text{ gdy } x \in (a,b) \\ 0, \text{ w przeciwnym przypadku} \end{cases}$$

$$\begin{aligned} EX &= \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{a+b}{2} \text{ – środek przedziału} \\ varX &= \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a}dx = ... = \frac{(b-a)^2}{12} \end{aligned}$$

rozkład normalny (*ang. normal distribution*) $\tilde{}N(m, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}} \qquad EX = m \qquad varX = \sigma^2$$

rozkład wykładniczy

$$f(x) = \begin{cases} \alpha e^{-\alpha x}, \text{ gdy } x \geqslant 0 \\ 0, \text{ w przeciwnym przypadku} \end{cases} \qquad EX = \frac{1}{\alpha} \qquad varX = ? \text{ – zadanie domowe}$$

pakiet **STATISTICA** – hasło "dystrybuanty" w indeksie pomocy

5. Eksperyment

$x_1, x_2, ..., x_N$ – ciąg liczb losowych (np. realizacje pewnej zmiennej losowej X)

6. Typy zbieżności probabilistycznych

Fakt zbieżności ciągów deterministycznych (nie losowych)

$$\lim_{k\to\infty} a_k = g \Leftrightarrow \bigwedge_{\varepsilon>0} \bigvee_{k_0} \bigwedge_{k>k_0} |a_k - g| < \varepsilon$$

Szybkość zbieżności ciągów deterministycznych

symbol $o()$ – "rząd niższy"

$$a_k = o(b_k) \Leftrightarrow \lim_{k\to\infty} \frac{a_k}{b_k} = 0 \text{ (i oba ciągi dążą do zera)}$$

symbol $O()$ – "ta sama szybkość"

$$a_k = O(b_k) \Leftrightarrow \bigvee_{c<\infty} |a_k| \leqslant c\,|b_k|$$

Ciągi zmiennych losowych $\{\varkappa_k\}$

tutaj operator $\lim_{k\to\infty}$ nie wystarcza, gdyż warunek $|a_k - g| < \varepsilon$ określa pewne zdarzenie **losowe**

Definicja 1 *Ciąg zmiennych losowych $\{\varkappa_k\}$ jest przy $k \to \infty$ zbieżny* ***według prawdopodobieństwa*** *(słabo) do $\varkappa^\#$ jeśli dla każdego $\varepsilon > 0$ zachodzi*

$$\lim_{k\to\infty} P(\left|\varkappa_k - \varkappa^\#\right| > \varepsilon) = 0, \textit{ lub równoważnie } \lim_{k\to\infty} P(\left|\varkappa_k - \varkappa^\#\right| < \varepsilon) = 1$$

Wartość $\varkappa^\#$ nazywamy granicą stochastyczną ciągu $\{\varkappa_k\}$ i zapisujemy

$$P \lim_{k\to\infty} \varkappa_k = \varkappa^\# \tag{1}$$

Zapis $Plim_{N\to\infty}\mathbf{X}_N = \mathbf{X}$ dla sekwencji wektorów losowych $\{\mathbf{X}_N\}$, oznacza, że $\mathbf{X}_N \to \mathbf{X}$ według prawdopodobieństwa, gdy $N \to \infty$.

Definicja 2 *Ciąg zmiennych losowych $\{\varkappa_k\}$ jest przy $k \to \infty$ zbieżny **z prawdopodobieństwem 1** (mocno) do $\varkappa^*$ jeśli zachodzi*

$$P(\lim_{k\to\infty} \varkappa_k = \varkappa^*) = 1$$

Lemat 1 *Ze zbieżności z prawdopodobieństwem 1 wynika zbieżność według prawdopodobieństwa.*

Definicja 3 *Ciąg zmiennych losowych $\{\varkappa_k\}$ jest przy $k \to \infty$ zbieżny **według średniej z potęgą** r do $\varkappa^*$ jeśli zachodzi*

$$\lim_{k\to\infty} \mathbf{E}\left|\varkappa_k - \varkappa^*\right|^r = 0$$

*w szczególności jest zbieżny **według średniej z kwadratem (średniokwadratowo)**, gdy*

$$\lim_{k\to\infty} \mathbf{E}(\varkappa_k - \varkappa^*)^2 = 0$$

Definicja 4 *Ciąg zmiennych losowych $\{\varkappa_k\}$ ma **szybkość zbieżności rzędu** $O(e_k)$ **według prawdopodobieństwa** przy $k \to \infty$ (tj. asymptotycznie), gdzie $\{e_k\}$ jest ciągiem liczb dodatnich zbieżnym do zera, tzn.*

$$\varkappa_k = O(e_k) \text{ według prawdopodobieństwa}$$

wtedy i tylko wtedy, gdy $\left\{\frac{\varkappa_k}{e_k}\chi_k\right\}$ jest zbieżny według prawdopodobieństwa do zera dla każdego ciągu liczbowego $\{\chi_k\}$, takiego że $\lim_{k\to\infty} \chi_k = 0$.

Definicja 5 *Ciąg zmiennych losowych $\{\varkappa_k\}$ ma **szybkość zbieżności rzędu** $O(e_k)$ **według średniej z kwadratem** przy $k \to \infty$ jeżeli istnieje stała $0 \le c < \infty$, taka, że*

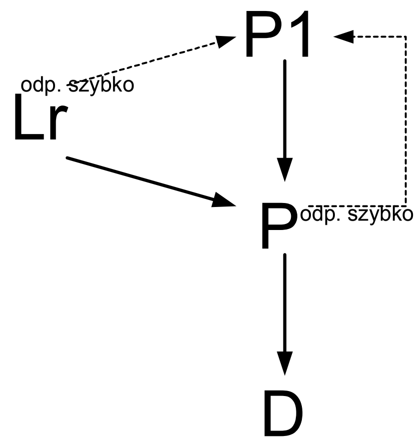
$$\mathbf{E}\varkappa_k^2 \le ce_k$$

Lemat 2 *Jeżeli $\varkappa_k = O(e_k)$ według średniej z kwadratem, to $\varkappa_k = O(\sqrt{e_k})$ według prawdopodobieństwa.*

Definicja 6 *Mówimy, że ciąg zmiennych losowych X_k jest zbieżny **według rozkładu** do zmiennej losowej X, gdy*

$$\lim_{k\to\infty} F_k(x) = F(x)$$

7. Relacje (związki) pomiędzy różnymi typami zbieżności



Dowód faktu $P1 \Longrightarrow P$

oczywisty, skoro $P(\lim_{k\to\infty} \varkappa_k = \varkappa^*) = 1$, to $\sum_{k=1}^{\infty} P(|\varkappa_k - \varkappa^*| < \varepsilon) < \infty$ i aby szereg ten był zbieżny, musi zachodzić $P(|\varkappa_k - \varkappa^*| < \varepsilon) \to 0$ dla $k \to \infty$

Dowód faktu $Lr \Longrightarrow P$

z definicji

$$\mathbf{E}\,|\varkappa_k - \varkappa^*|^r = \int_{\Omega} |\varkappa_k - \varkappa^*|^r\, d\omega \geqslant \int_{\{|\varkappa_k - \varkappa^*| > \varepsilon\}} |\varkappa_k - \varkappa^*|^r\, d\omega \geqslant \varepsilon^r \int_{\{|\varkappa_k - \varkappa^*| > \varepsilon\}} d\omega = \varepsilon^r P(|\varkappa_k - \varkappa^*| > \varepsilon)$$

a zatem

$$P(|\varkappa_k - \varkappa^*| > \varepsilon) \leqslant \frac{1}{\varepsilon^r}\mathbf{E}\,|\varkappa_k - \varkappa^*|^r$$

w szczególność dla $r = 2$ i $\varkappa^* = \mathbf{E}\varkappa$

$$P(|\varkappa - \mathbf{E}\varkappa| > \varepsilon) \leqslant \frac{1}{\varepsilon^2} var\,\varkappa$$

Dowód faktu $\sum_{k=1}^{\infty} P(|\varkappa_k - \varkappa^*| < \varepsilon) < \infty$ i $\varkappa_k \xrightarrow{p} \varkappa^* \Longrightarrow \varkappa_k \xrightarrow{p1} \varkappa^*$

$$P(\sup_{k \geqslant k_0} |\varkappa_k - \varkappa^*| \;>\; \varepsilon) = P(|\varkappa_k - \varkappa^*| > \varepsilon \text{ dla pewnego (konkretnego) } k \geqslant k_0) = P\left(\bigcup_{k=k_0}^{\infty} (|\varkappa_k - \varkappa^*| > \varepsilon)\right) \leqslant$$

$$\leqslant \sum_{k=k_0}^{\infty} P(|\varkappa_k - \varkappa^*| > \varepsilon) \to 0, \text{ bo szereg jest zbieżny, zaś } k_0 \to \infty$$

Przykład

jeśli $P(|\varkappa_k - \varkappa^*| < \varepsilon) = O(\frac{1}{k})$ wtedy zachodzi $\varkappa_k \xrightarrow{p} \varkappa^*$, ale nie zachodzi $\varkappa_k \xrightarrow{p1} \varkappa^*$

Problem

jeżeli $\varkappa_k \xrightarrow{p} \varkappa^*$, gdy $k \to \infty$, to czy wtedy zachodzi $g(\varkappa_k) \xrightarrow{p} g(\varkappa^*)$, gdy $k \to \infty$???

Tak – pod warunkiem, że $g()$ jest funkcją ciągłą w punkcie $\varkappa^*$

8. Mocne Prawo Wielkich Liczb Kołmogorowa (wersja podstawowa)

Założenia

(a) $X_1, X_2, ..., X_N$ – jest ciągiem zmiennych losowych typu i.i.d. – niezależnych i o tym samym rozkładzie (*ang* ***i****ndependent and* ***i****dentically* ***d****istributed sequence of random variables*)

(b) istnieje $\mathbf{E}X_i = m < \infty$

Teza

$$\frac{1}{N}\sum_{i=1}^{N} X_i \xrightarrow{p1} m, \text{ gdy } N \to \infty$$

inne wersje MPWL – patrz [Feller], [Krzyśko], [Ninness]

9. Mocne Prawo Wielkich Liczb Kołmogorowa (wersja bez wymogu i.i.d.)

Założenia

(a) $X_1, X_2, ..., X_N$ – jest ciągiem niezależnych zmiennych losowych, w ogólności o różnych rozkładach

(b) istnieją $\mathbf{E}X_i = m_i < \infty$

(c) istnieją $var X_i = \sigma_i^2 < \infty$

(d) $\sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} < \infty$

Teza

$$\frac{1}{N}\sum_{i=1}^{N} X_i - \frac{1}{N}\sum_{i=1}^{N} m_i \xrightarrow{p1} 0, \text{ gdy } N \to \infty$$

10. Centralne Twierdzenie Graniczne Lindenberga–Levy'ego

Założenia

(a) $X_1, X_2, ..., X_N$ – ciąg typu i.i.d. (mają ten nam, ale dowolny rozkład! – niekoniecznie normalny)

(b) istnieje $\mathbf{E}X_i = m < \infty$

(c) istnieje $var X_i = \sigma^2 < \infty$

Teza

$$\frac{\sum_{i=1}^{N} X_i - Nm}{\sigma\sqrt{N}} \xrightarrow{D} \mathcal{N}(0,1), \text{ gdy } N \to \infty$$

Wnioski – fundamentalne dla zagadnienia estymacji

$$\frac{\frac{1}{N}\sum_{i=1}^{N} X_i - m}{\frac{\sigma}{\sqrt{N}}} \xrightarrow{D} \mathcal{N}(0,1) \qquad \frac{1}{N}\sum_{i=1}^{N} X_i \xrightarrow{D} \mathcal{N}(m, \frac{\sigma^2}{N})$$

Oszacowanie dokładności przybliżenia – nierówność Barry-Essena

oznaczmy $\varkappa_N = \frac{\sum_{i=1}^{N} X_i - Nm}{\sigma\sqrt{N}}$

$$\sup_x |F_{\varkappa_N}(x) - \Phi(x)| \leqslant \frac{33}{4}\frac{\mathbf{E}\,|X_i - m|^3}{\sigma^3\sqrt{N}} = O\left(\frac{1}{\sqrt{N}}\right)$$

11. Analiza korelacyjna procesów

kowariancja – miara zależności liniowej

$$cov(X,Y) = \mathbf{E}\{(X - \mathbf{E}X)(Y - \mathbf{E}Y)\} \qquad\qquad |cov(X,Y)| \leqslant \sqrt{varXvarY}$$

korelacja (znormalizowana kowariancja)

$$\xi(X,Y) = \frac{cov(X,Y)}{\sqrt{varXvarY}} \qquad\qquad |\xi(X,Y)| \leqslant 1$$

pojęcie procesu losowego (stochastycznego)

$X(\omega,t)$ – dla ustalonego momentu czasu $t = t_0$ otrzymujemy zmienną losową $X_{t_0}(\omega)$

funkcja autokowariancji procesu losowego (stacjonarnego) – miara zależności liniowej pomiędzy X_{t_0} o przesuniętą o τ zmienną $X_{t_0+\tau}$

$$A_X(\tau) = cov(X_{t_0}, X_{t_0+\tau}), \qquad\qquad A_X(0) = \sigma_X^2$$

funkcja autokorelacji procesu losowego

$$r_X(\tau) = \frac{cov(X_{t_0}, X_{t_0+\tau})}{\sqrt{varX_{t_0}varX_{t_0+\tau}}} = \frac{A_X(\tau)}{\sigma_X^2}, \qquad\qquad r_X(0) = 1$$

funkcja kowariancji wzajemnej dwóch procesów – $X(\omega,t)$ i $Y(\omega,t)$

$$W_{X,Y}(\tau) = cov(X_{t_0}, Y_{t_0+\tau})$$

funkcja korelacji wzajemnej dwóch procesów – $X(\omega,t)$ i $Y(\omega,t)$

$$r_{X,Y}(\tau) = \frac{W_{X,Y}(\tau)}{\sigma_X\sigma_Y}$$

12. Przejście białego szumu przez układ dynamiczny

$$y_k = \sum_{i=0}^{\infty} \gamma_i u_{k-i}$$

Założenia

(a) $\{u_k\}$ – proces typu i.i.d.

(b) układ jest asymptotycznie stabilny tzn. $\sum_{i=0}^{\infty} |\gamma_i| < \infty$

(c) dla uproszczenia prezentacji niech $\mathbf{E}u_k = 0$ i $var u_k = 1$

Autokowariancja procesu u_k

$$\begin{aligned} A_u(\tau) &= \mathbf{E}u_k u_{k+\tau} = \begin{cases} = var u_k = 1, \text{ dla } \tau = 0 \\ = 0, \text{ dla } \tau \neq 0 \text{ (na podstawie niezależności } u_k \text{ i } u_{k+\tau} \text{ i założenia (c))} \end{cases} \\ r_u(\tau) &= A_u(\tau) \text{ (patrz założenie (c))} \end{aligned}$$

Własości procesu y_k

$$\begin{aligned} \mathbf{E}y_k &= \mathbf{E}\sum_{i=0}^{\infty} \gamma_i u_{k-i} = \sum_{i=0}^{\infty} \mathbf{E}\gamma_i u_{k-i} = \mathbf{E}u_k \sum_{i=0}^{\infty} \gamma_i = 0 \\ var y_k &= var\left(\sum_{i=0}^{\infty} \gamma_i u_{k-i}\right) = \sum_{i=0}^{\infty} var\left(\gamma_i u_{k-i}\right) = var u_k \sum_{i=0}^{\infty} \gamma_i^2 \\ A_y(\tau) &= \mathbf{E}y_k y_{k+\tau} = \mathbf{E}\left(\sum_{i=0}^{\infty} \gamma_i u_{k-i} \sum_{j=0}^{\infty} \gamma_j u_{k+\tau-j}\right) = \\ &= \mathbf{E}\left\{(\gamma_0 u_k + \gamma_1 u_{k-1} + \gamma_2 u_{k-2} + ...)\left(\gamma_0 u_{k+\tau} + \gamma_1 u_{k+\tau-1} + \gamma_2 u_{k+\tau-2} + ... + \gamma_\tau u_k + \gamma_{\tau+1} u_{k-1} + ...\right)\right\} = var u_k \sum_{i=0}^{\infty} \gamma_i \gamma_{i+\tau} \end{aligned}$$

13. Popularne nierówności

Nierówność Czebyszewa

$$P(|\varkappa - \mathbf{E}\varkappa| > \varepsilon) \leqslant \frac{1}{\varepsilon^2} var \varkappa$$

Nierówność Barry-Essena
oznaczmy $\varkappa_N = \frac{\sum_{i=1}^{N} X_i - Nm}{\sigma\sqrt{N}}$

$$\sup_x |F_{\varkappa_N}(x) - \Phi(x)| \leqslant \frac{33}{4} \frac{\mathbf{E}\,|X_i - m|^3}{\sigma^3\sqrt{N}} = O\left(\frac{1}{\sqrt{N}}\right)$$

Nierówność Jensena
$g()$ – funkcja wypukła

$$\mathbf{E}g(X) \geqslant g(\mathbf{E}X)$$

Nierówność Höldera

$$\begin{aligned} \|X\|_p &= (\mathbf{E}X^p)^{1/p} \text{ tzw. p-norma zmiennej losowej} \\ \mathbf{E}\,|XY| &\leqslant \|X\|_p \|Y\|_{p'}\text{, gdzie } \frac{1}{p} + \frac{1}{p'} = 1 \end{aligned}$$

Nierówność Schwartza $(p = 2,\ p' = 2)$

$$|\mathbf{E}XY| \leqslant \mathbf{E}\,|XY| \leqslant \sqrt{\mathbf{E}X^2\mathbf{E}Y^2}$$

Nierówność Rao-Cramera

$$\mathbf{E}(\theta_N - \theta^*)^{\mathbf{2}} \geqslant \frac{1}{N\int_{-\infty}^{\infty}\left(\frac{\partial f(x,\theta^*)}{\partial\theta^*}\right)^2 f(x,\theta^*)dx}$$