metody statystyczne wykład 2

kacper.topolnicki@uj.edu.pl

$\boxed{\text{plan}}$

- przypomnienie
- ciągłe rozkłady losowe
- zmiana zmiennych losowych 1D
- wielowymiarowe zmienne losowe
- zmiana zmiennych losowych 2D

przypomnienie

prawdopodobieństwo

- (Ω, F, P)
  - P – funkcja prawdopodobieństwa
  - F – zbiór zdarzeń losowych
  - Ω – zbiór zdarzeń elementarnych
- aksjomaty Kolomogorowa

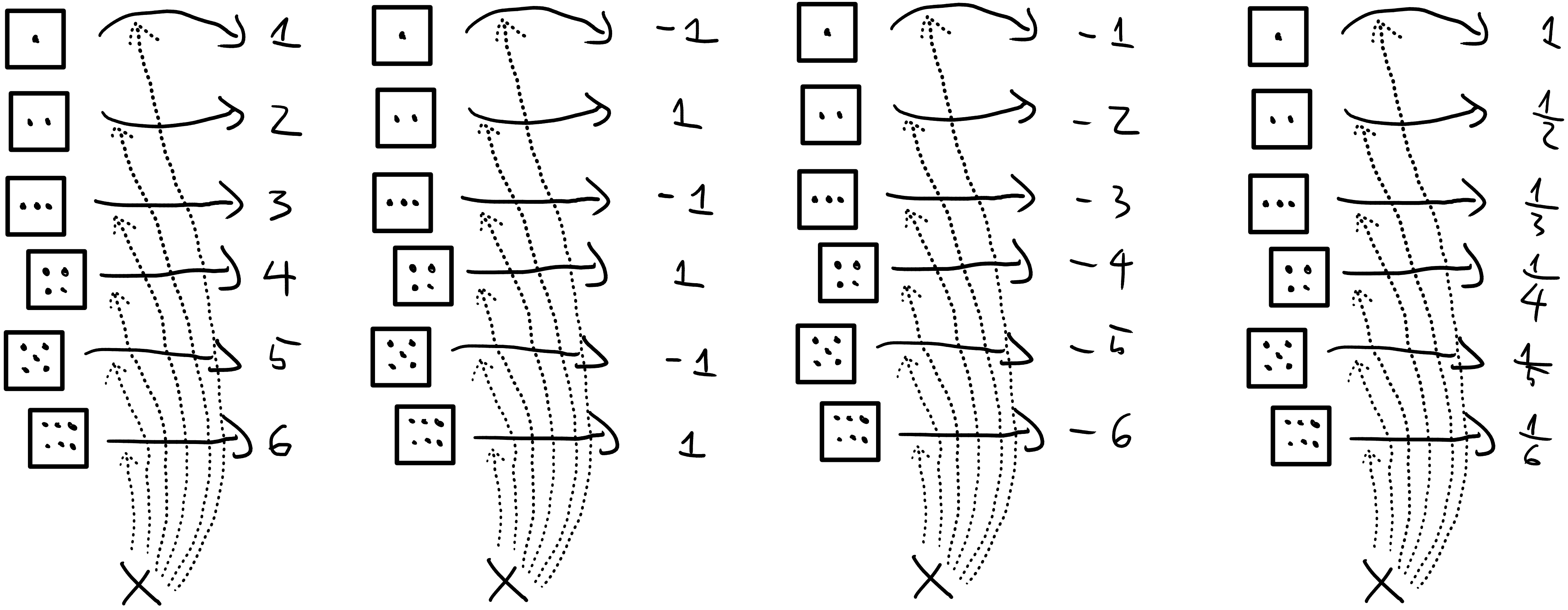
zmienna losowa X:

- funkcja
- $X : \Omega \to E$

  Ω – zbiór zdarzeń elementarnych

  E – zbiór „mierzalny”
- zazwyczaj $E = \mathbb{R}$

Kilka przykładów:



- typowo
- parzystość
- odwrotność
- mamy dużo swobody!

ciągłe rozkłady prawdopodobieństwa

- Ω nie musi być skończonym zbiorem np.:

$$\Omega = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}$$

- może być „nieprzeliczalnym zbiorem" np.:

$$\Omega = \mathbb{R} \qquad \Omega = \underbrace{\mathbb{R} \times \mathbb{R}} \qquad \Omega = \mathbb{R} \times \mathbb{R} \times \ldots$$

zbiór liczb rzeczywistych

zbiór par (wektorów) liczb

- funkcja prawdopodobieństwa oparta na całkowaniu Funkcji Gęstości Prawdopodobieństwa

- 1D

zmienna losowa X

wartość X : x

$$\underbrace{P(a<x<b)}_{\substack{\text{prawdopodobieństwo} \\ \text{wylosowania } x \\ \text{w przedziale } (a,b)}} = \int_a^b f_X(x)\,dx$$

FGP

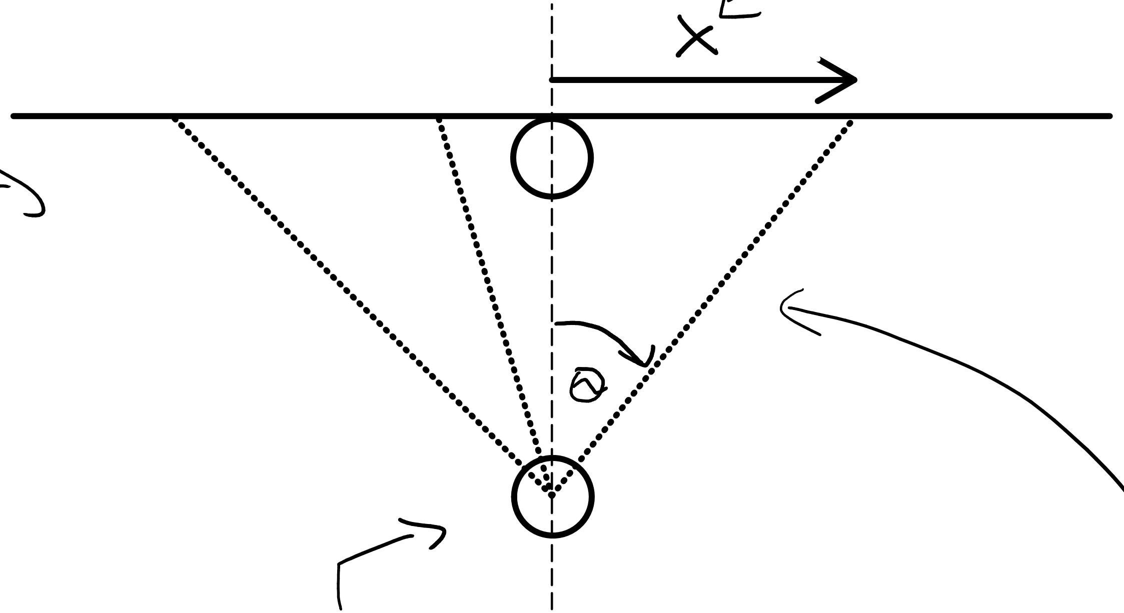
Zmiana zmiennych losowych 1D

rozkład Cauchiego

pozycja trafienia

mur

nieszczęśnik



szaleniec z bronią strzela w losowych kierunkach

$$\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \quad f(\theta) = \frac{1}{\pi}$$

rozkład jednorodny FGP

- znamy FGP kąta strzału: $f(\alpha)$
- ile wynosi FGP pozycji trafienia w mur: $g(x)$?
- zachowanie prawdopodobieństwa:

$$\underbrace{f(\alpha)\, d\alpha}_{\text{prawdopodobieństwo oddania strzału w „nieskończenie mały" zakres kąta}} = \underbrace{g(x)\, dx}_{\text{prawdopodobieństwo, że strzał trafi w „nieskończenie mały" fragment muru}}$$

$d\alpha$: „nieskończenie mały" zakres kąta strzału

dx: „nieskończenie mały" odcinek muru

$$\underbrace{f(\theta)\,d\theta = g(x)\left|\frac{dx}{d\theta}\right|}_{dx} d\theta$$

interesują nas szerokości dx a nie znak

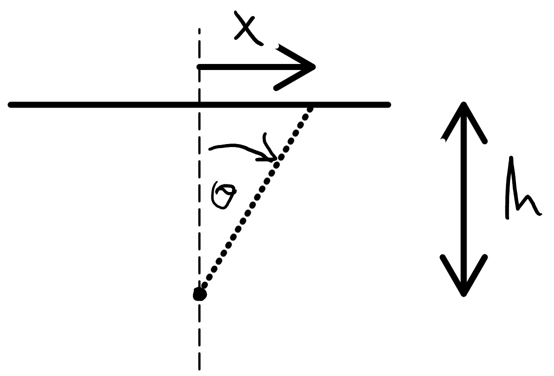
pochodna $x(\theta)$

- możemy porównać
- równanie powinno być spełnione dla wszystkich wartości θ, x

$$f(\theta(x)) = g(x)\left|\frac{dx}{d\theta}\right|$$

można zapisać jako funkcja x

- możemy policzyć $g(x)$!



$$\tan(\theta) = \frac{x}{h}$$

$$x = x(\theta) = h \cdot \tan\theta$$

$$\theta = \theta(x) = \arctan\left(\frac{x}{h}\right)$$

$$\frac{dx}{d\theta} = h \left(\frac{1}{\cos(\theta)}\right)^2 > 0 \quad \text{możemy opuścić } |\ |$$

$$\frac{1}{\pi} = f(\theta)\, d\theta = g(x)\, h \frac{1}{\cos^2(\theta)}\, d\theta$$

$$g(x) = \frac{1}{\pi} \frac{\cos^2(\theta)}{h} = \frac{1}{\pi} \frac{\cos^2\left(\arctan\left(\frac{x}{h}\right)\right)}{h} = \ldots$$

$$\sin^2(\theta) + \cos^2(\theta) = 1$$

$$\tan^2(\theta) + 1 = \frac{1}{\cos^2(\theta)}$$

$$\cos^2(\theta) = \frac{1}{1 + \tan^2(\theta)}$$

$$\dots = \underbrace{\frac{h}{h^2\pi + \pi x^2}} = g(x)$$

rozkład Cauchiego !

$$\boxed{\begin{array}{l}\text{proszę sprawdzić :} \\ \displaystyle\int_{-\infty}^{+\infty} g(x)\, dx \overset{?}{=} 1\end{array}}$$

pojawi
się na
ćwiczeniach

- ile wynosi $\mathbb{E}(x)$?

- ile wynosi $\mathrm{var}(x)$?

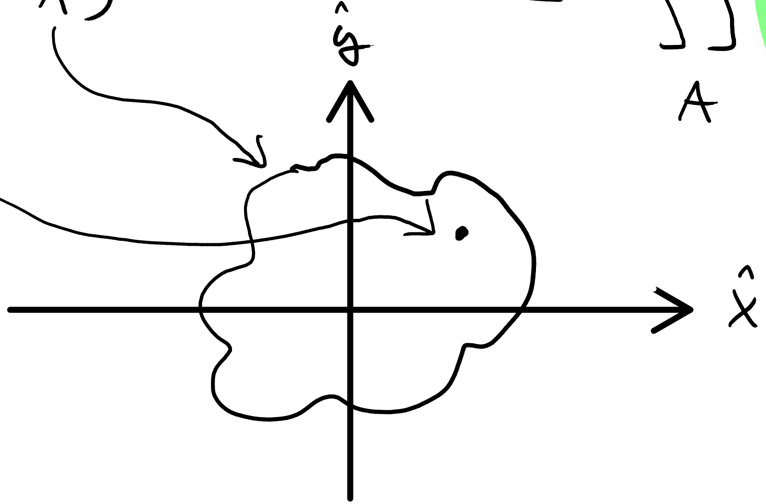
wielowymiarowe zmienne losowe

- (x, y) $x, y \in \mathbb{R}$ $\big((x,y,z), (x,y,z,u), \ldots\big)$

nowa zmienna losowa

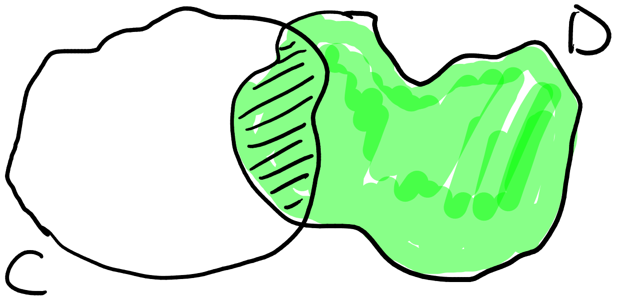
- FGP

$$P(r \in A) = \iint_A f(x,y)\,dx\,dy$$



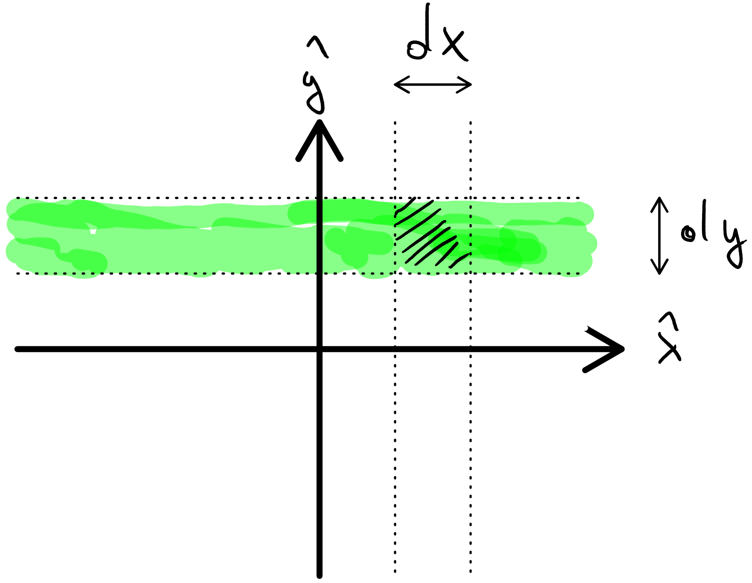
• prawdopodobieństwo warunkowe

$$P(C|D) \equiv \frac{P(C \cap D)}{P(D)}$$



$$\overbrace{f(x|y)}^{\text{gęstość praw. warunk.}}\,dx = \frac{f(x,y)\,dx\,dy}{\underbrace{f(y)}_{\text{gęstość prawdopodob. brzegowego}}\,dy}$$

trafienie w
pod warunkiem



$$f(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx$$

- cowariancja

$$\mathrm{cov}(X, Y) = E\big((X - E(X))(Y - E(Y))\big) =$$

zmienne losowe

$$= \iint (x - E(X))(y - E(Y)) \overbrace{f(x, y)}^{\text{FGP}} \, dx \, dy$$

wartości zmiennych losowych

$$\mathrm{cov}(X, X) = \mathrm{var}(X)$$

$$\mathrm{cov}(Y, Y) = \mathrm{var}(Y)$$

$$\mathrm{cov}(X, Y) = \mathrm{cov}(X, X)$$

- macierz kowariancji

$$K_{X,Y} = \begin{pmatrix} var(X) & cov(X,Y) \\ cov(Y,X) & var(Y) \end{pmatrix}$$

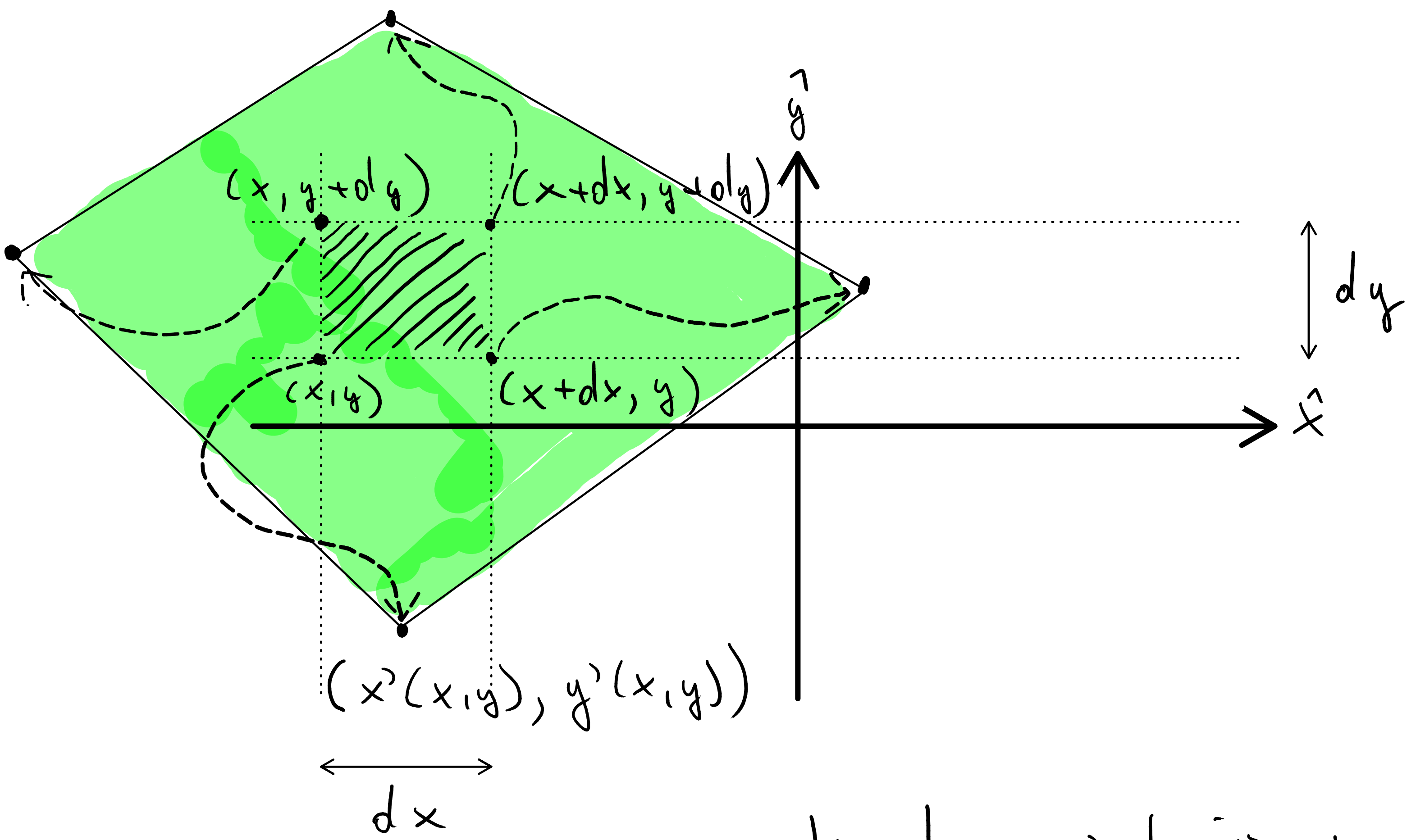
- współczynnik korelacji

$$\rho_{X,Y} = corr(X,Y) = \frac{cov(X,Y)}{var(X)\, var(Y)}$$

zmiana zmiennych 2D

$$\underbrace{f(x,y)\,dx\,dy}$$

prawdopodobieństwo, że punkt wpadnie ////



- nowe zmienne

$$x' = x'(x,y)$$
$$y' = y'(x,y)$$

dx, dy – „nieskończenie małe"

- zachowanie prawdopodobieństwa:

$$f(x,y)\,dx\,dy = g(x',y') \cdot \langle \text{powierzchnia} \rangle$$

- szukany FGP nowych zmiennych:

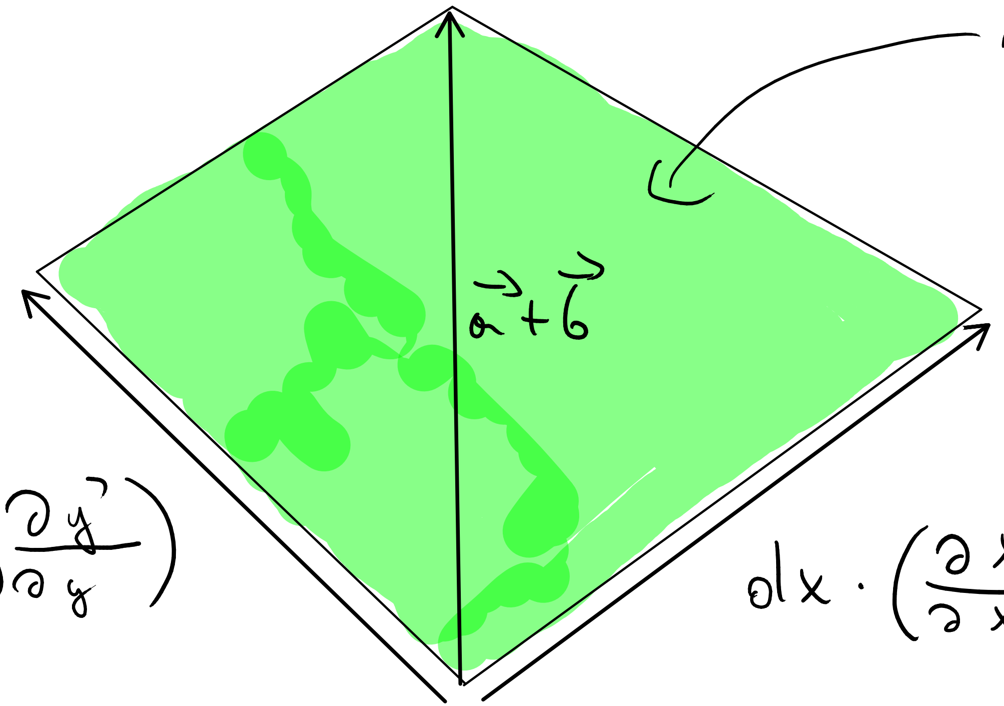
$$g(x', y')$$

- ile wynosi $\langle$powierzchnia $\rangle$:

$$\underbrace{\begin{aligned}
&x'(x+dx, y) = x'(x,y) + dx\,\frac{\partial x'}{\partial x} + \dots; \quad y'(x+dx, y) = y'(x,y) + dx\,\frac{\partial y'}{\partial x} + \dots; \\
&x'(x, y+dy) = x'(x,y) + dy\,\frac{\partial x'}{\partial y} + \dots; \quad y'(x, y+dy) = y'(x,y) + dy\,\frac{\partial y'}{\partial y} + \dots; \\
&x'(x+dx, y+dy) = x'(x,y) + dx\,\frac{\partial x'}{\partial x} + dy\,\frac{\partial x'}{\partial y} + \dots; \\
&y'(x+dx, y+dy) = y'(x,y) + dx\,\frac{\partial y'}{\partial x} + dy\,\frac{\partial y'}{\partial y} + \dots;
\end{aligned}}_{\text{rozwinięcie w szereg w okolicy } (x,y)}$$

- ponieważ dx oraz dy są małe to możemy zaniedbać "..."

równoległobok



$$\vec{a} = dy \cdot \left(\frac{\partial x'}{\partial y}, \frac{\partial y'}{\partial y}\right)$$

$$dx \cdot \left(\frac{\partial x'}{\partial x}, \frac{\partial y'}{\partial x}\right) = \vec{b}$$

• powierzchnia

interesuje nas powierzchnia a nie to jak jest ona skierowana

$$dx\,dy \left| \det \underbrace{\begin{pmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} \end{pmatrix}}_{\text{macierz Jacobiego}} \right| = dx \cdot dy \, |\det(J)|$$

– można inaczej, np licząc iloczyn wektorowy

– ten sposób łatwo jednak uogólnić na 4D, 5D, ...

• wracając do zachowania prawdopodobieństwa

$$f(x,y)\,dx\,dy = g(x',y') \cdot$$ <powierzchnia >

$$f(x,y)\,dx\,dy = g(x',y') \cdot |\det(J)|$$

możemy wyliczyć FGP nowych zmiennych!