

$X_1, X_2, X_3, \dots, X_p$ - predyktoły, zmienne mierzone,
 metryczne
 Bill Gates
 input variables, features, cechy
 zmiennej
 rzeczywistej
 prawdopodobieństwo
 prawdopodobieństwo
 prawdopodobieństwo
 prawdopodobieństwo
 \vdots
 RZECZYWISTOŚĆ

$$\vec{X} = (X_1, X_2, X_3, \dots, X_p) \in \mathbb{R}^p$$

\vec{Y} - output, response, dependent variable

$$Y = f(\vec{X}) + \epsilon$$

? niepoznawalna
 systematyczna
 zależność
 losowość, szum
 $\mathcal{N}(0, \sigma^2)$

MODEL

$$\hat{Y} = \hat{f}(x)$$

przydzielona

wartości
zmiennych zależnych

\ estymata (przybliżenia) f

E - błąd (arg. error)

$$\begin{aligned} E(Y - \hat{Y})^2 &= E[f(x) + \varepsilon - \hat{f}(x)]^2 \\ &= |f(x) - \hat{f}(x)|^2 + \text{Var}(\varepsilon) \end{aligned}$$

Składownica
redukowalna

Składownica
nieusuwalna

Metody parametryczne

Poznajemy, i.e.:

$$\hat{f}(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

b a

$$y = ax + b$$

Na podstawie danych rozscian

wyznaczymy, i.e. $X = (X_1)$

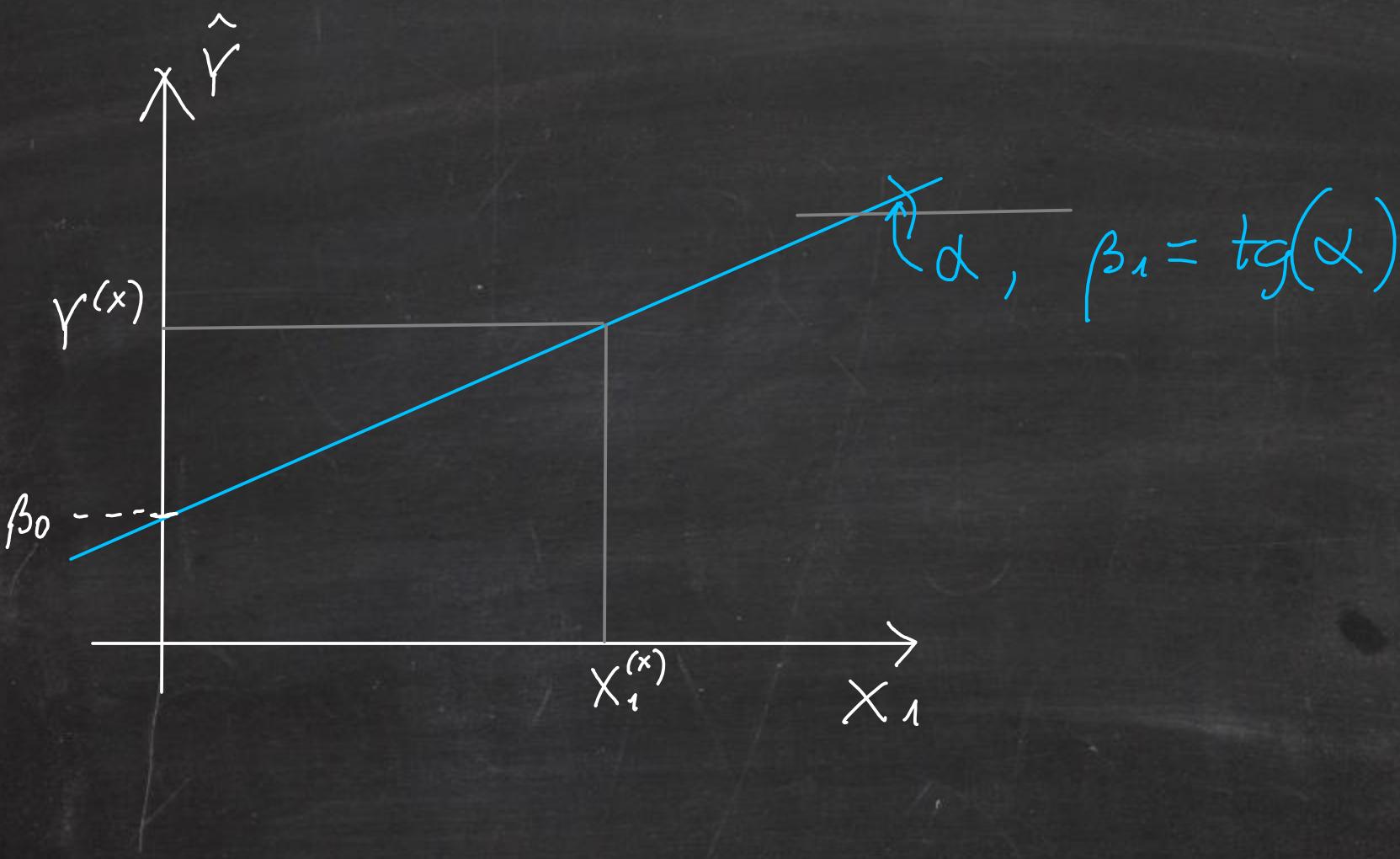
gdzie X_1 - powierzchnia mieszkania

INTERCEPT SLOPE

$$\hat{Y} = \hat{f}(x) = \beta_0 + \beta_1 X_1$$

? ? metrae

TODD: wyznaczyć β_0 i β_1



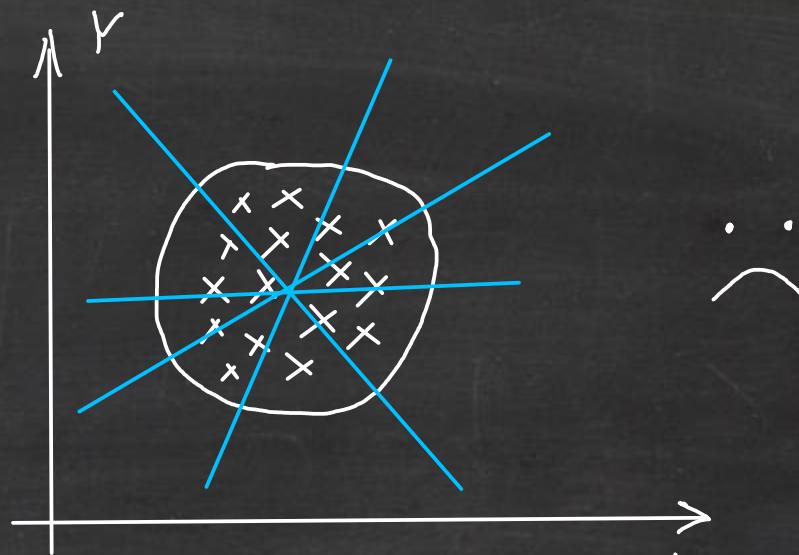


$$\hat{\beta}_1 = t_y(x)$$

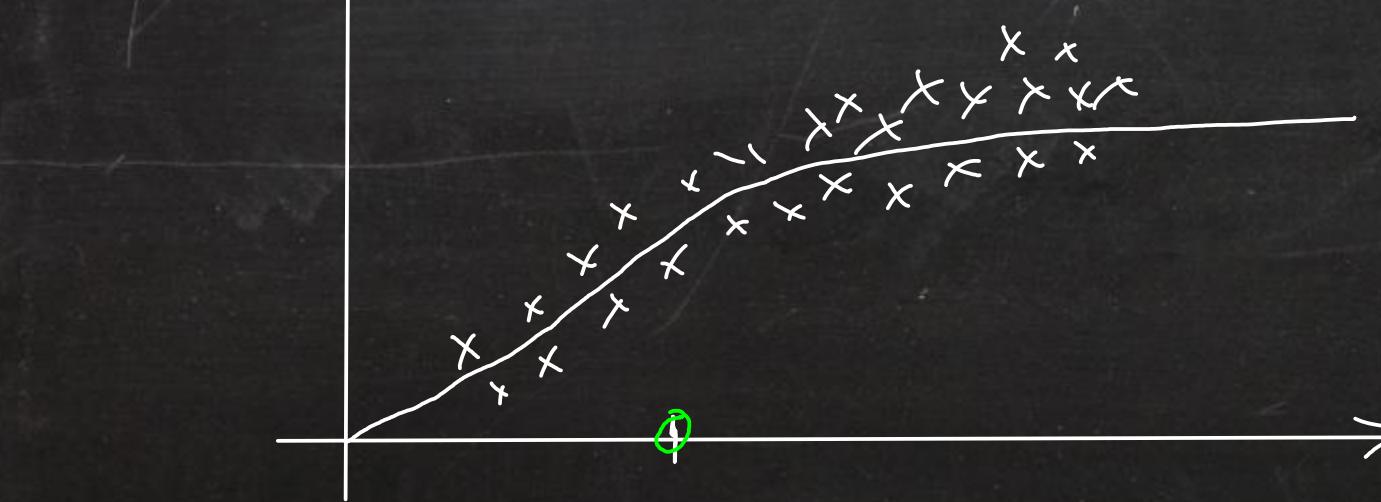
Metoda najmniejszych kwadratów



Dobrobyt
outlier



brak korelacji pomiędzy X i Y



Lata edukacji

