Regression basics

dark horse of statistics

Chuvakin Sergey

«School of Advanced Studies»

December 21, 2020

Outline

- Why do we need it?
- Applications
- Formal representation
- Intuition behind
- Estimators of regression
- Optimzations for regression
- Assumtions (Gauss-Markov)

Why do we need it?

- *Forcast* future
- *Explain* present and the past

Applications

- Social Sciences (in broad sense)
- Economics
- Natural Sciences (originally came from Phisics)
- Business
- Medicine (Pharmacology)
- Urban Studies
- Criminology
- etc

Formal representation

Things in «real» abstract world:

$$\hat{y} = \beta X \tag{1}$$

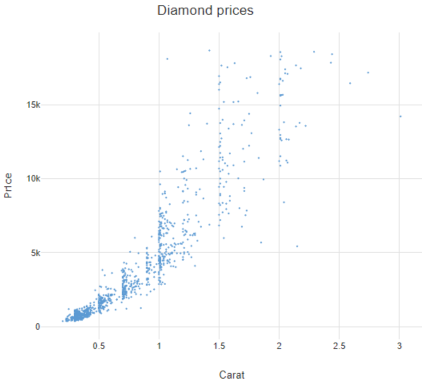
Things we are doing in statistics:

$$\hat{y} = \beta X + \hat{\epsilon} \tag{2}$$

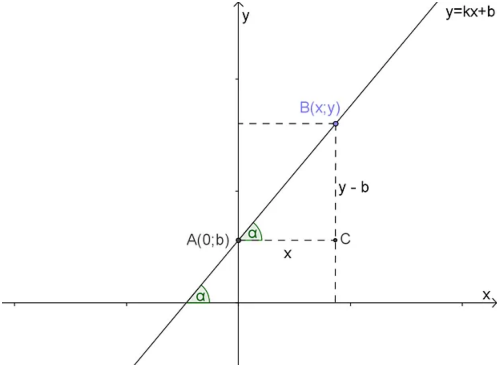
More common representation:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_n x_n + \hat{\epsilon} \tag{3}$$

Intuition behind

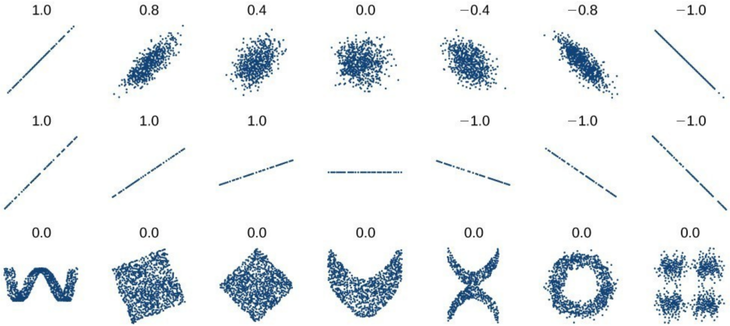


Intuition behind



Intuition behind

What if we have smth like this?



Intuition behind

Once we decided that our data can be fitted by line model, let's select β coefs. Let's follow here.

Estimators

Estimation method (aka LOSS functions) - equation that help to find coefs. The most common techniques:

- OLS: $\beta = \Sigma(\hat{y} - y)^2 \rightarrow min$
- MLE: $\theta = P(X|\theta) \rightarrow max$
- Other purpose specified (Ridge, Quantile, PLS, Bayes Models)

Estimators

When and what

What should I use for my model?

- OLS (and ols related): only continious variables
- MLE (and mle related): all other varaibles

How it works inside

- Analytical solution $\beta = (X^T X)^{-1} X^T y$
- (Stochastic) Gradient Descent (related variations)
- MCMC (for Bayes equations)

Gauss-Markov Assumptions

- Linearity of data
- Sample should be *randomly* selected for population
- X matrix should not be correlated within
- X marrix should not be correlated with error
- Variance of error should be constant