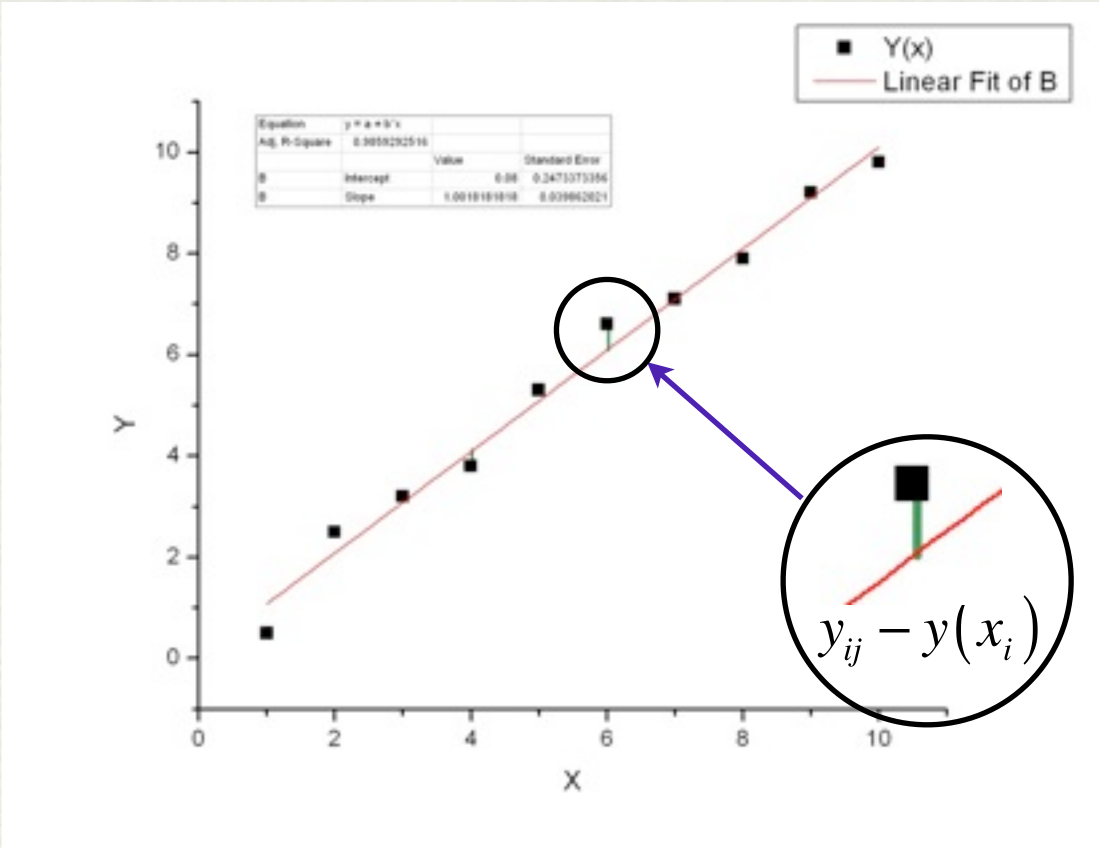
Metoda najmanjih kvadrata

Metoda najmanjih kvadrata (MNK)

MNK smo već uveli u proučavanju linearne korelacije; gdje smo tražili da suma kvadrata odstupanja eksperimentalnih točaka od pravca koji ih na "najbolji" način opisuje bude minimalna:

$$\boxed{S^2 = \sum_{i,j} f_{ij}\left(y_{ij} - ax_i - b\right)^2 = \min}$$

Točka: (x_i, y_{ij})
Pravac: $y = ax + b$
Teorijski y_{ij}: $y_{ij} = ax_i + b$



Ponekad nam u pokusu nije od interesa grupirati podatke u razrede i sl., pa možemo svaku točku (x_i, y_j) gledati zasebno, tj $f_{ij} = 1$, za svaki i, j.

Ukoliko se neki par (x_i, y_j) ponovi, jednostavno ga ubrojimo kao novu točku.

Metoda najmanjih kvadrata (MNK)

Taj je pristup u znanosti vrlo čest, naročito ako radimo s relativno malim brojem parova.

Dakle, stvar se pojednostavljuje:

$$\boxed{S^2 = \sum_{i=1}^{n}\left[y_i - y(x_i)\right]^2 = \sum_i \left[y_i - ax_i - b\right]^2 = \min}$$

Kao i prije, rješavamo:

$$\frac{\partial S^2}{\partial a} = 0; \quad \frac{\partial S^2}{\partial b} = 0$$

Metoda najmanjih kvadrata (MNK)

Rješenja su:

$$a = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2},$$

$$b = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

Metoda najmanjih kvadrata (MNK)

Jasno, možemo pisati (kao i ranije):

$$a = \frac{\sigma_{xy}}{\sigma_x^2} \qquad b = \overline{y} - \frac{\sigma_{xy}}{\sigma_x^2}\,\overline{x}$$

Suma kvadrata odstupanja: $S^2 = \sigma_y^2\left(1 - r^2\right)$

Često a i b (naročito a) imaju fizikalno značenje.

Gore izraženi a i b su *najvjerojatnije vrijednosti* nagiba i odsječka pravca.

Uz pretpostavku gausijanskog odstupanja (y_i - ax_i - b) to su i srednje vrijednosti.

Metoda najmanjih kvadrata (MNK)

Primjer: analiza jednolikog gibanja po pravcu

$$x(t) = \underbrace{v_0}_{a} \cdot t + \underbrace{x_0}_{b}$$

Dakle, ako je $\bar{a} = \sigma_{xy} / \sigma_x^2$, $\bar{b} = \bar{y} - \overline{ax}$ zanima nas i nepouzdanost koeficijenata, kako bismo u potpunosti izračunali naše rezultate.

Metoda najmanjih kvadrata (MNK)

Bez izvoda:

- srednja pogreška (pogreška jednog mjerenja) dana je izrazom:

$$m^2 = \frac{\sum_i \left(y_i - y(x_i)\right)^2}{n-2} = \frac{\sum_i \left[y_i - ax_i - b\right]^2}{n-2}$$

Intuitivni "dokaz":

$m^2 \propto S^2 \longrightarrow$ što veća suma kvadrata odstupanja to veća pogreška

$m^2 \propto 1/(n-2) \longrightarrow$ - $m \to \infty$ ako je $n = 2$, jer za pravac kroz dvije točke ne možemo znati pogrešku

- s druge strane, $m^2 \propto 1/(\text{broj događaja})$

Metoda najmanjih kvadrata (MNK)

Doslovno raspisivanje m^2 vodi do komplicirane formule, pa možemo stvar pojednostaviti:

$$\sum_i \left(y_i - (ax_i + b)\right)^2 = \sum_i \left(y_i^2 + (ax_i + b)^2 - 2(ax_i + b)y_i\right) =$$

$$= \sum_i \left\{ y_i^2 + (ax_i + b)\left[\underbrace{(ax_i + b)}_{\approx y_i} - 2y_i \right] \right\} \approx \sum_i \left\{ y_i^2 - y_i(ax_i + b) \right\}$$

Prema tome je:

$$m^2 = \frac{1}{n-2} \sum_i \left\{ y_i^2 - ax_i y_i - by_i \right\} =$$

$$= \frac{1}{n-2} \sum_i \left\{ y_i^2 - ax_i y_i - (\bar{y} - a\bar{x}) y_i \right\}$$

Metoda najmanjih kvadrata (MNK)

$$m^2 = \frac{1}{n-2}\left\{\sum_i y_i^2 - \bar{y}\sum_i y_i - a\left(\sum_i x_i y_i - \bar{x}\sum_i y_i\right)\right\} =$$

$$= \frac{1}{n-2}\left\{\sum_i y_i^2 - \frac{1}{n}\left(\sum_i y_i\right)^2 - a\left[\sum_i x_i y_i - \frac{1}{n}\sum_i x_i \sum_i y_i\right]\right\} =$$

$$= \frac{1}{n(n-2)}\left\{n\sum_i y_i^2 - \left(\sum_i y_i\right)^2 - a\left[n\sum_i x_i y_i - \sum_i x_i \sum_i y_i\right]\right\}$$

Iz izraza za *a*:

$$n\sum_i x_i y_i - \sum_i x_i \sum_i y_i = a\left[n\sum_i x_i^2 - \left(\sum_i x_i\right)^2\right]$$

Metoda najmanjih kvadrata (MNK)

$$m^2 = \frac{1}{n(n-2)}\left\{n\sum_i y_i^2 - \left[\sum_i y_i\right]^2 - a^2\left[n\sum_i x_i^2 - \left(\sum_i x_i\right)^2\right]\right\} =$$
$$= \frac{n}{n-2}\left\{\sigma_y^2 - a^2\sigma_x^2\right\}$$

No, nas najviše zanima nepouzdanost. Ovaj put ne možemo koristiti izraz $M = m/\sqrt{n}$ već je stvar malo kompliciranija.

Vrijedi:

$$M^2\left(C_1x_1 + C_2x_2 + \ldots C_nx_n\right) = C_1^2M^2\left(x_1\right) + C_2^2M^2\left(x_2\right) + \ldots + C_n^2M^2\left(x_n\right)$$

Metoda najmanjih kvadrata (MNK)

Iskoristimo to, napisavši *a* malo drugačije:

$$a = \sum_i \left[\frac{n\left(x_i - \overline{x}\right)}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} \right] \cdot y_i$$

$$\boxed{M_a^2 = \sum_i \left[\frac{n\left(x_i - \overline{x}\right)}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} \right]^2 M^2\left(y_i\right)}$$

Pretpostavka: sve $M^2(y_i)$ su međusobno jednake te vrijednosti m^2 (pogreška jednog mjerenja)

Metoda najmanjih kvadrata (MNK)

$$M_a^2 = \frac{m^2 n^2}{\left[n\sum_i x_i^2 - \left(\sum_i x_i\right)^2\right]^2} \sum_i \left(x_i - \bar{x}\right)^2$$

$$\sum_i \left(x_i - \bar{x}\right)^2 = \sum_i x_i^2 - n\bar{x}^2 = \sum_i x_i^2 - n\left(\frac{\sum_i x_i}{n}\right)^2 = \frac{1}{n}\left[n\sum_i x_i^2 - \left(\sum_i x_i\right)^2\right]$$

$$\boxed{M_a^2 = \frac{n \cdot m^2}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2}}$$

Metoda najmanjih kvadrata (MNK)

Stoga je:

$$M_a^2 = \frac{1}{n-2}\left\{\frac{n\sum_i y_i^2 - \left(\sum_i y_i\right)^2}{n\sum_i x_i^2 - \left(\sum_i x_i\right)^2} - a^2\right\} =$$

$$= \frac{1}{n-2}\left\{\frac{\sigma_y^2}{\sigma_x^2} - a^2\right\}$$

Nepouzdanost za *b* (bez izvoda):

$$M_b^2 = M_a^2 \overline{x}^2$$

Metoda najmanjih kvadrata (MNK)

Općeniti slučaj polinoma:

Pristupom metode najmanjih kvadrata moguće je naći koeficijente polinoma proizvoljnog stupnja, iz jednadžbi:

$$\frac{\partial S^2}{\partial \alpha_i} = 0, \text{ gdje je } y(x) = \sum_n \alpha_n \cdot x^n$$

Oprez! Općenita formula za polinom k-tog stupnja mora se koristiti uz sve $\alpha_i \neq 0$!

Primjer: $y = ax$

$$b = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

Stavimo $b = 0$ u $y = ax$ i to implicira

$$(*) \quad \sum_i x_i^2 \sum_i y_i - \sum_i x_i \sum_i x_i y_i = 0$$

zašto bi to uistinu bilo tako?

Metoda najmanjih kvadrata (MNK)

Puno je lakše i 'sigurnije' učiniti ovo:

$$\frac{\partial}{\partial a}\left\{\sum_i \left(y_i - ax_i\right)\right\} = -2\sum_i \left(y_i - ax_i\right) = 0$$

$$a = \frac{\sum_i x_i y_i}{\sum_i x_i^2}$$

Taj rezultat može se dobiti i uvrštavanjem (*) u izraz za *a*, no učinimo li to, radimo *konceptualnu pogrešku*

$$\boxed{a = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}}$$

Metoda najmanjih kvadrata (MNK)

Primjer polinoma: $y = \alpha_0 + \alpha_1 x + \alpha_2 x^2$

$$\alpha_0 n + \alpha_1 \sum_i x_i + \alpha_2 \sum_i x_i^2 = \sum_i y_i$$

$$\alpha_0 \sum_i x_i + \alpha_1 \sum_i x_i^2 + \alpha_2 \sum_i x_i^3 = \sum_i x_i y_i$$

$$\alpha_0 \sum_i x_i^2 + \alpha_1 \sum_i x_i^3 + \alpha_2 \sum_i x_i^4 = \sum_i x_i^2 y_i$$

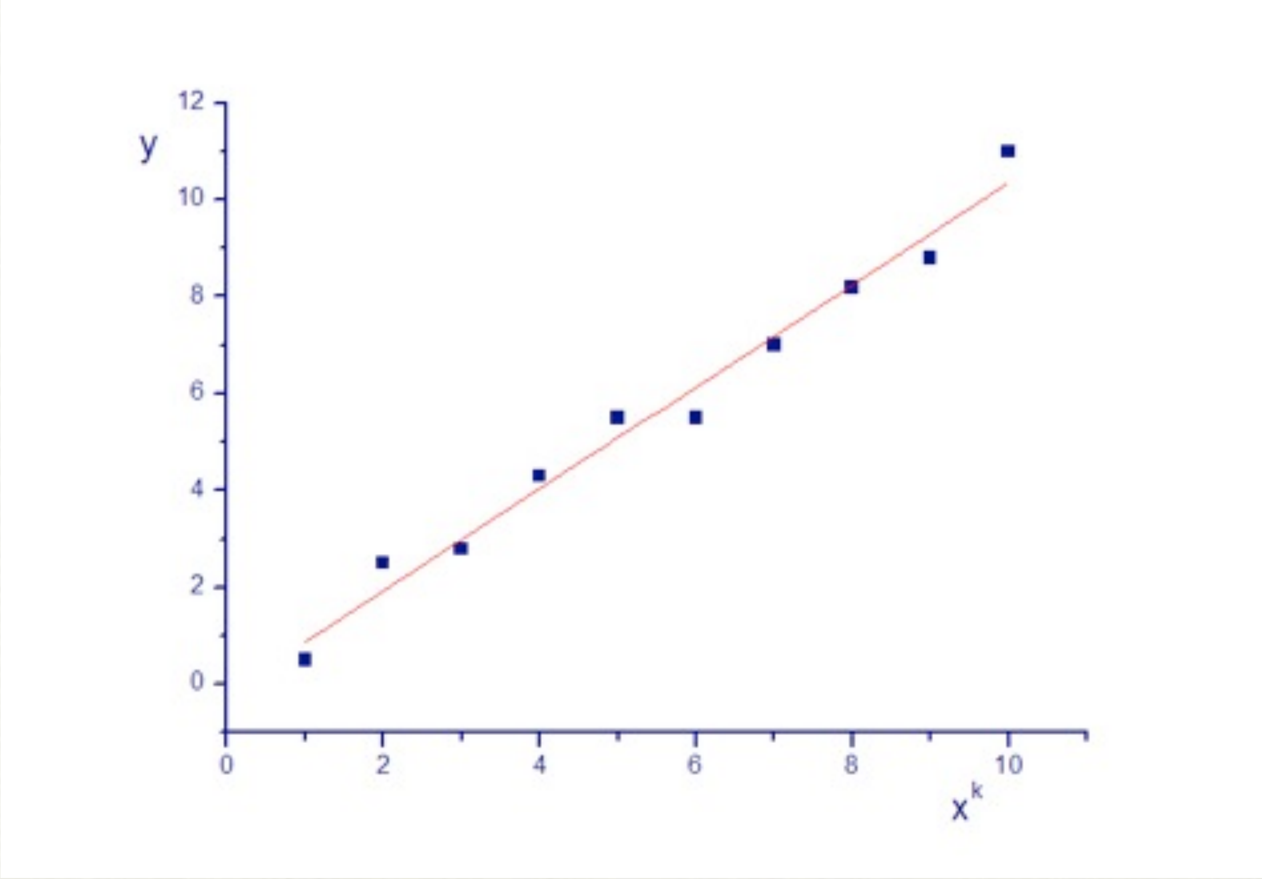
Općenito, za polinom s k koeficijenata moramo riješiti sustav od k jednadžbi s k nepoznanica

Metoda najmanjih kvadrata (MNK)

Pregled slučajeva koji se daju riješiti MNK

(1) 'polinom' tipa $y(x) = ax^k + b$, gdje je k *znani broj*

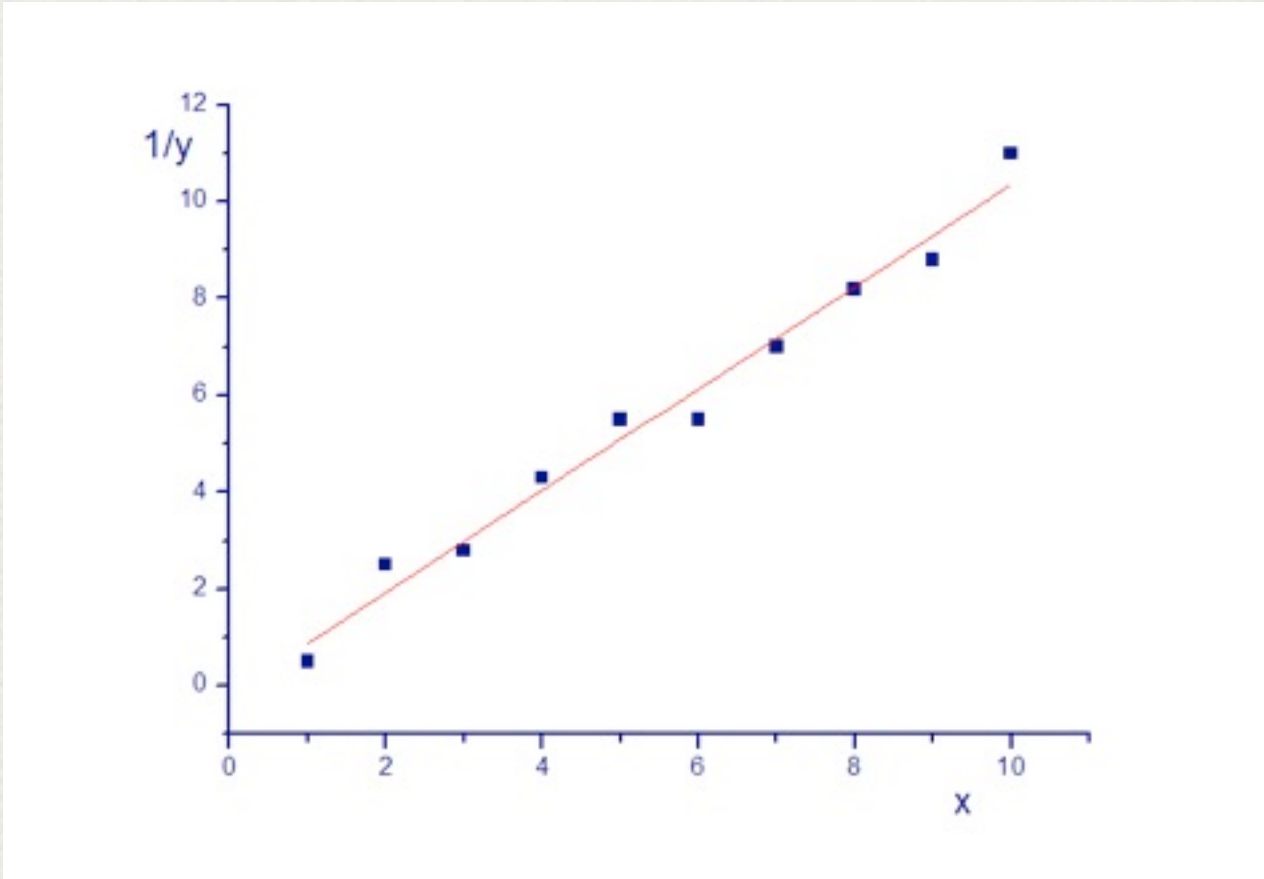
Supstitucija: $\xi = x^k \Rightarrow y(\xi) = a\,\xi + b$
Grafički prikaz:



Metoda najmanjih kvadrata (MNK)

(2) 'Inverzni' polinomi, npr. $y = 1/(ax + b)$

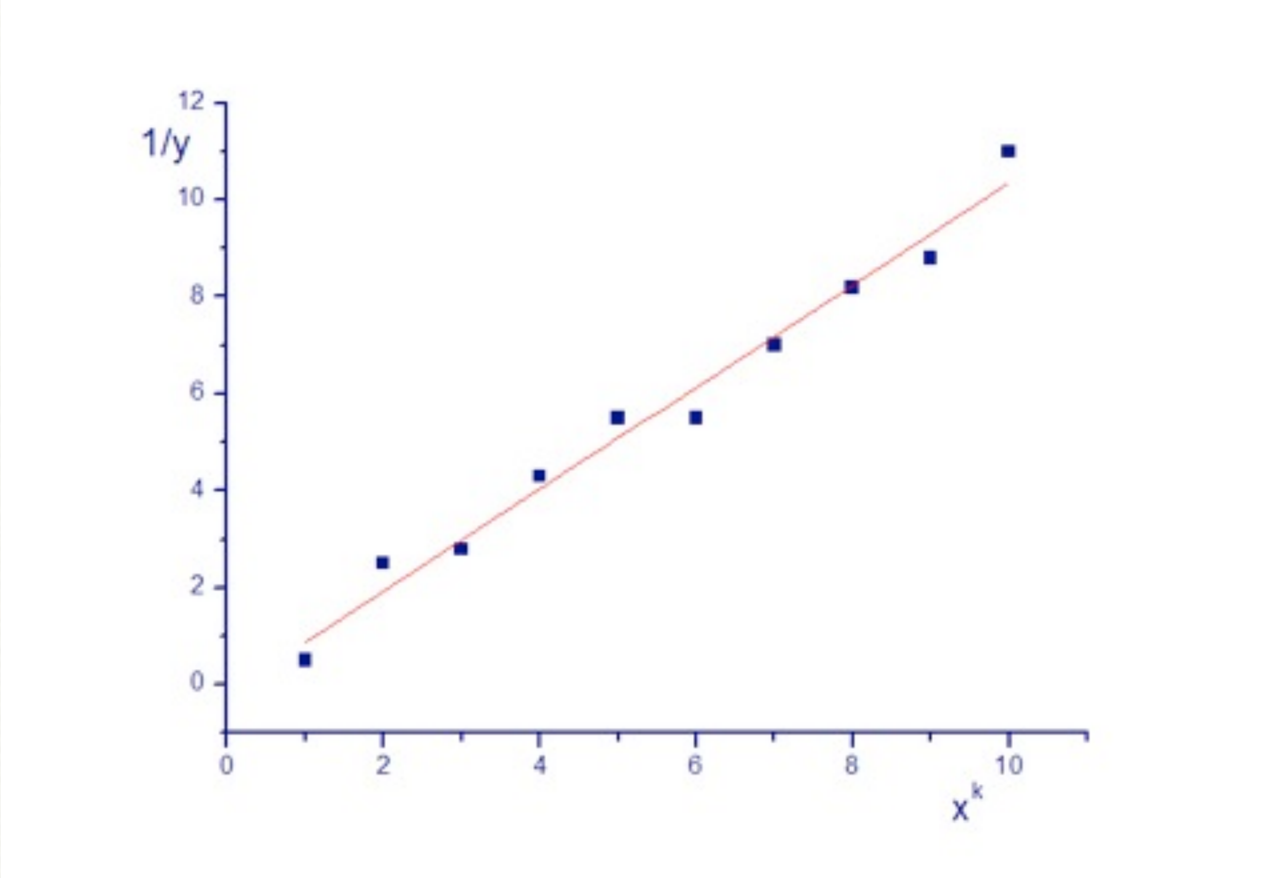
Supstitucija: $\eta = 1/y \Rightarrow \eta(x) = a\,x + b$
Grafički prikaz:



Metoda najmanjih kvadrata (MNK)

Povezivanje (1) i (2): npr. $y = 1/\left(ax^4 + b\right)$ rješava se supstitucijom $\eta = 1/y$ i $\xi = x^4$, a grafički prikaz je:

$\eta = a\xi + b$



Metoda najmanjih kvadrata (MNK)

(3) potencijska funkcija $y = ax^b$, b je nepoznato

$$y = ax^b \quad / \begin{cases} \log_{10} \\ \ln \end{cases}$$

$$\log_{10} y = \log a + b \log x$$

$$\ln y = \ln a + b \ln x$$

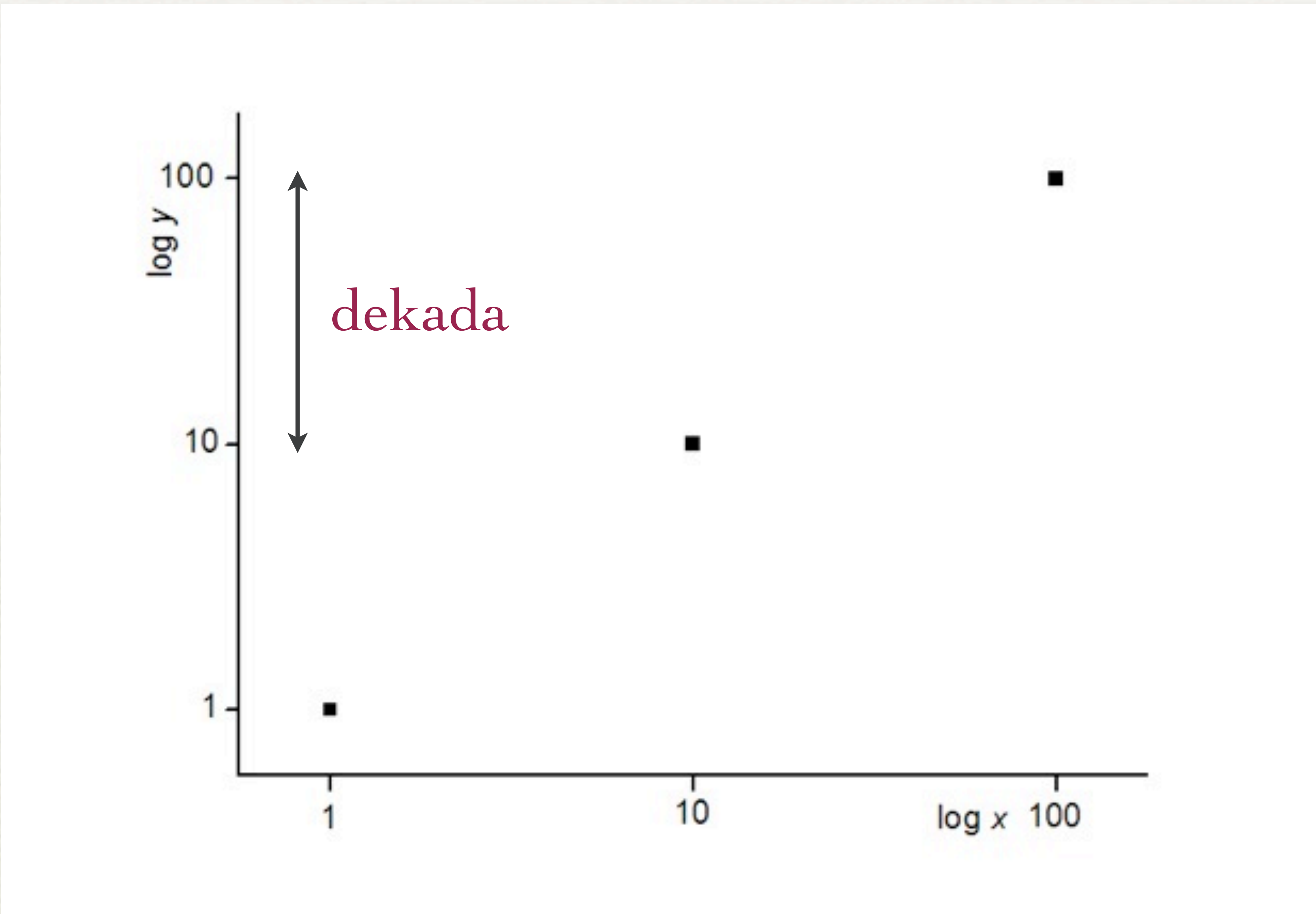
$$Y = \begin{cases} \log_{10} y \\ \ln y \end{cases} ; \quad X = \begin{cases} \log_{10} x \\ \ln x \end{cases} ; \quad A = \begin{cases} \log_{10} a \\ \ln a \end{cases}$$

$\boxed{Y = A + bX}$ $\quad b = \bar{b} \pm M_b \longrightarrow$

- ne ovisi o zboru $\log_{10}$ ili ln
- M_A nije trivijalno naći, ali b najčešće ima fizikalno značenje

Metoda najmanjih kvadrata (MNK)

- grafički prikaz: LOG-LOG plot



ln-ln plot je u praksi rijeđi

Vrlo važan način prikazivanja rezultata. Omogućava prikaz vrijednosti koje se protežu kroz nekoliko redova veličina.

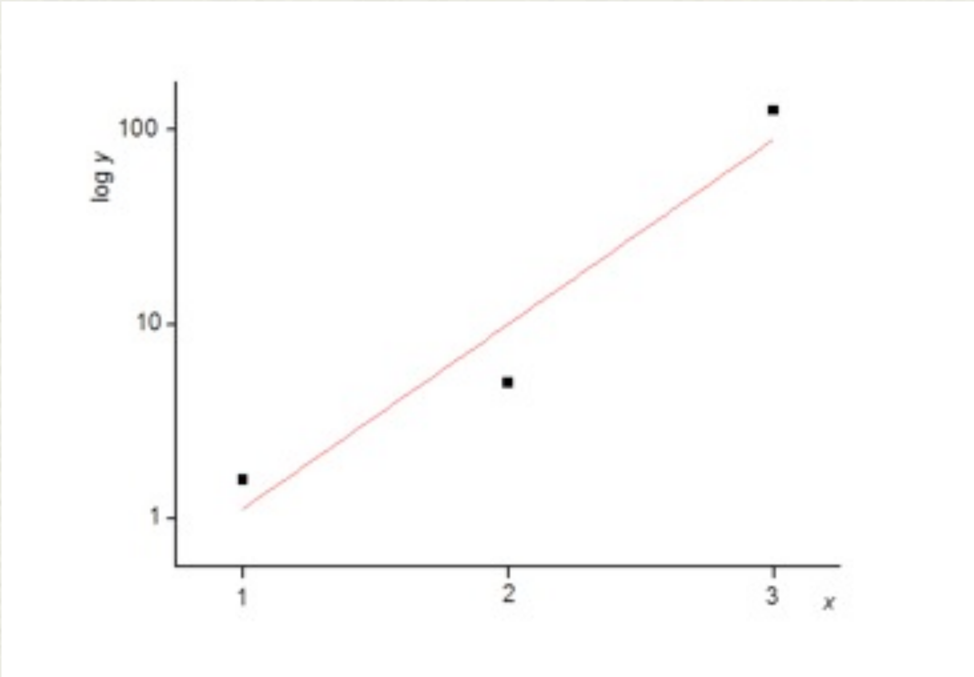
Metoda najmanjih kvadrata (MNK)

(3) eksponencijalna funkcija $y = a \cdot e^{bx}$

$$\left.\begin{aligned} \ln y &= \ln a + bx \\ \log_{10} y &= \log_{10} a + \frac{b}{\ln 10} x \end{aligned}\right\} \quad Y = A + BX$$

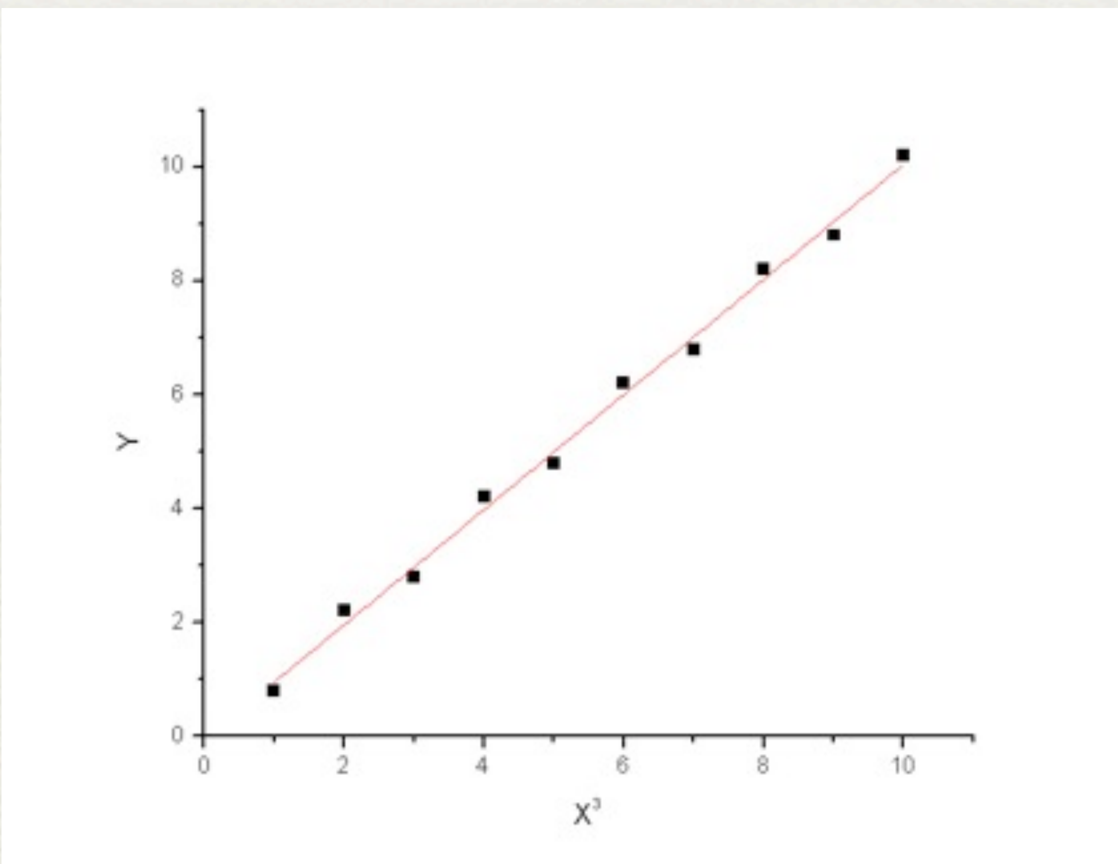
Y A B

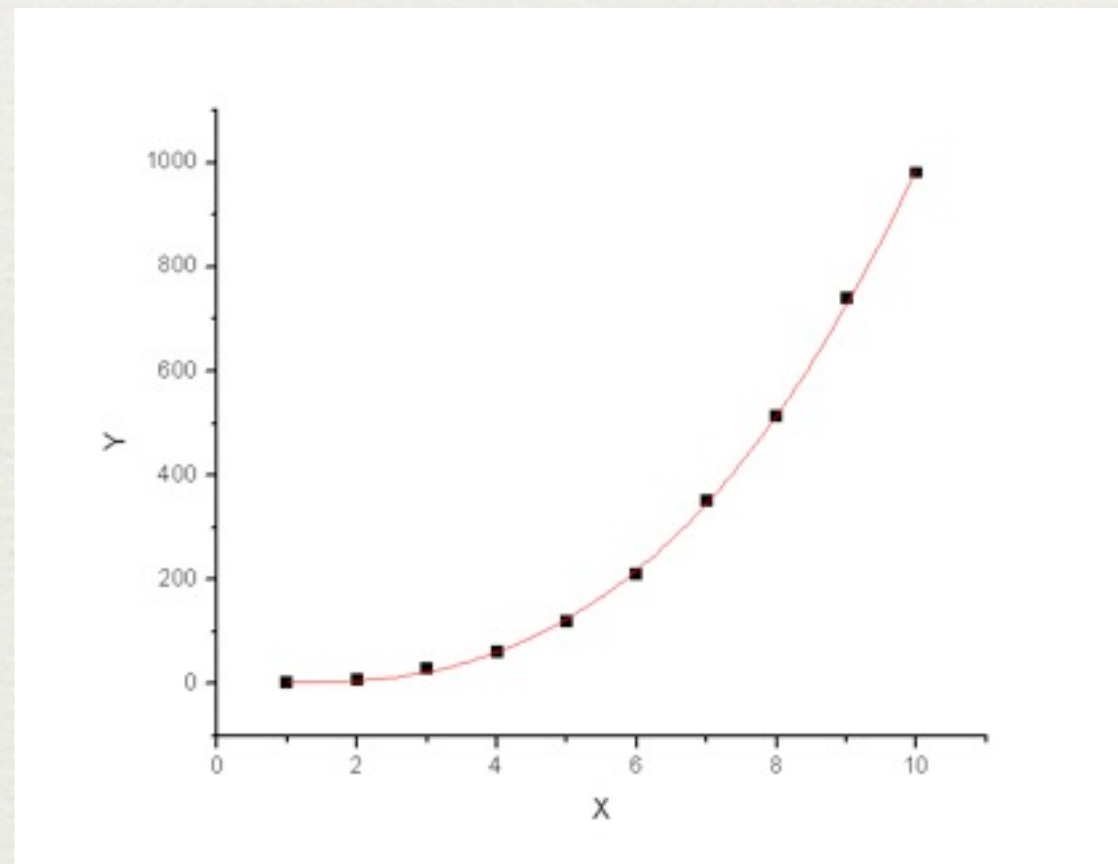
-grafički prikaz: LIN-LOG plot



Metoda najmanjih kvadrata (MNK)

- različitim kombinacijama opisanih slučajeva može se riješiti veliki broj funkcija
- općenito se pri prikazu rezultata treba što više oslanjati na pravac
- npr. $y = ax^3$ je puno bolje analizirati crtanjem y vs. x^3 nego li y vs. x





jasno se vidi leže li točke dobro uz pravac ili ne

procjena slaganja je nepouzdanija