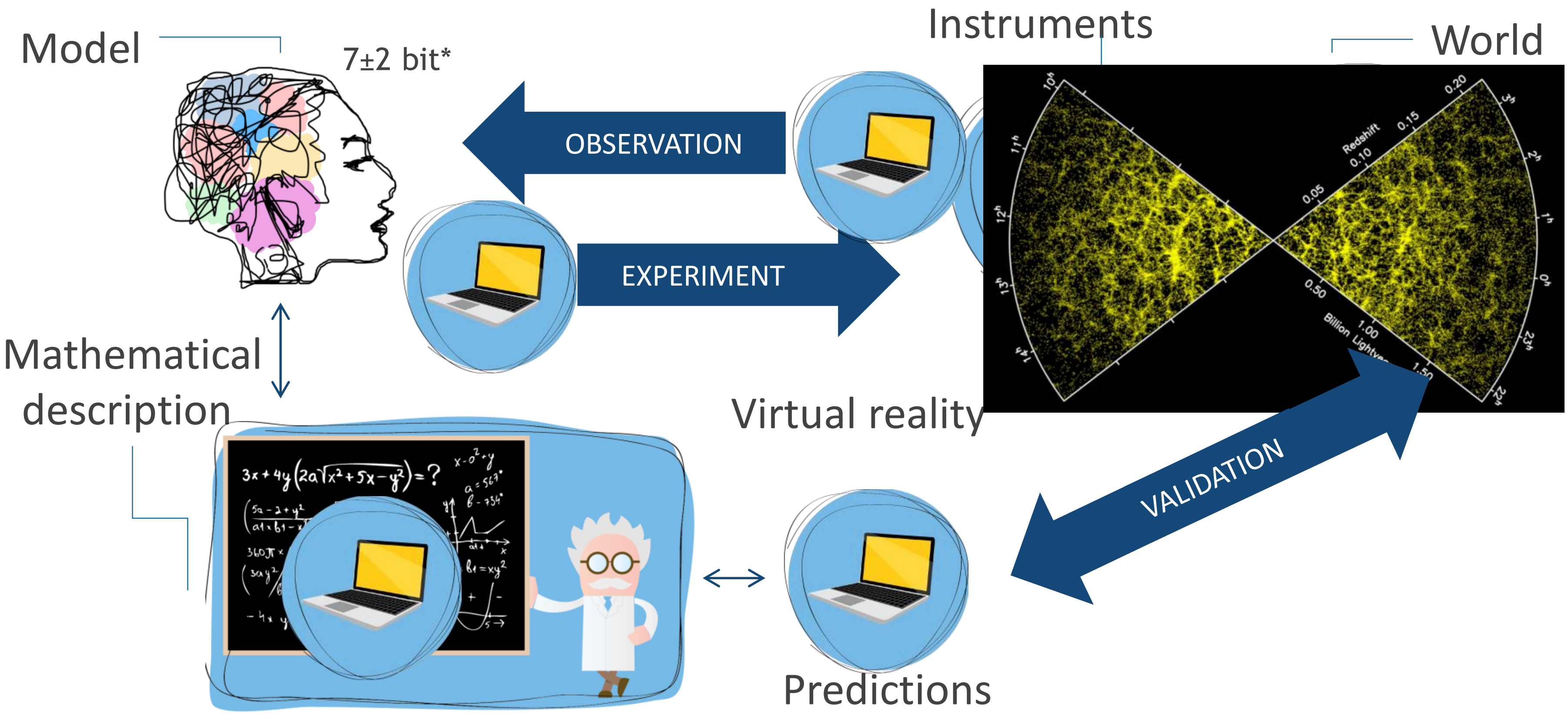
Least Squares and Maximum Likelihood

Istvan Csabai

Modern data science



Initial values

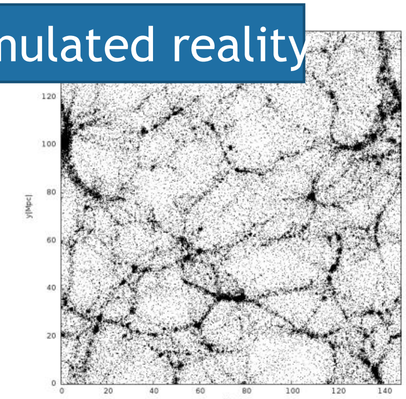
“laws”, equations

Simulated reality

Λ=0.7
Ω_m=0.3

$$F = G\frac{m_1 m_2}{r^2}$$

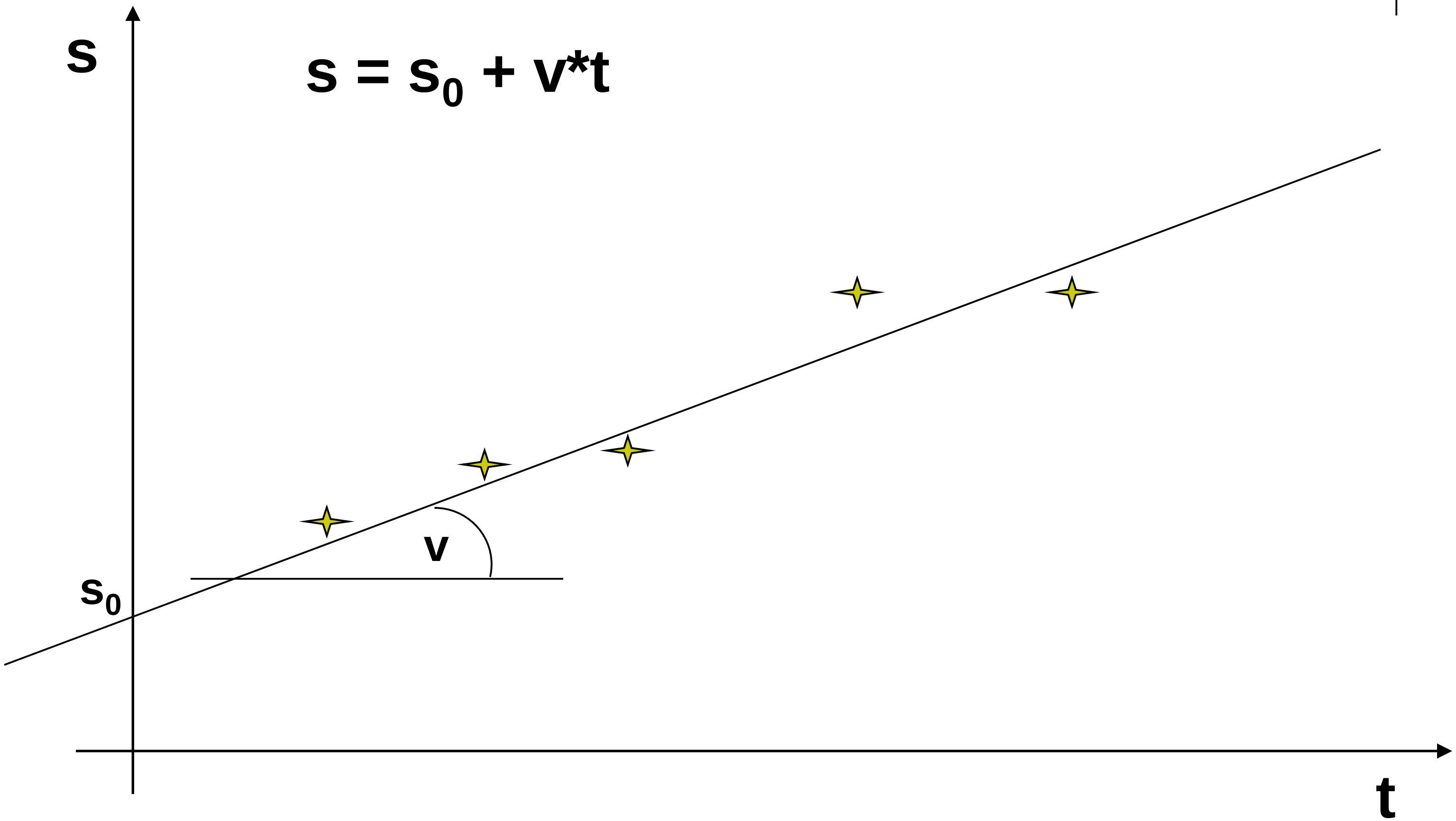
$$R_{\mu\nu} - \frac{1}{2}R\,g_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}$$



Introduction

- Find a MODEL for measurement/simulation data, that:
  - Fits some theory
  - Reproduce measurement data
  - Can be parametrized
- Parameters may have „physical” meaning
  - Displacement/time diagram – linear fit – slope: speed

Example: speed



Errors

- Measurements contain **errors** (in each variables!)
  - noise, inaccuracy
  - Systematic errors
  - outliers
- Hence: the calculated parameters will have errors, too

Least square fit

Given: $(x_i, y_i);\ i=1..N$ measurement points. Fit some function :

$$y(x) = y(x; a_1, a_2, \ldots, a_M)$$

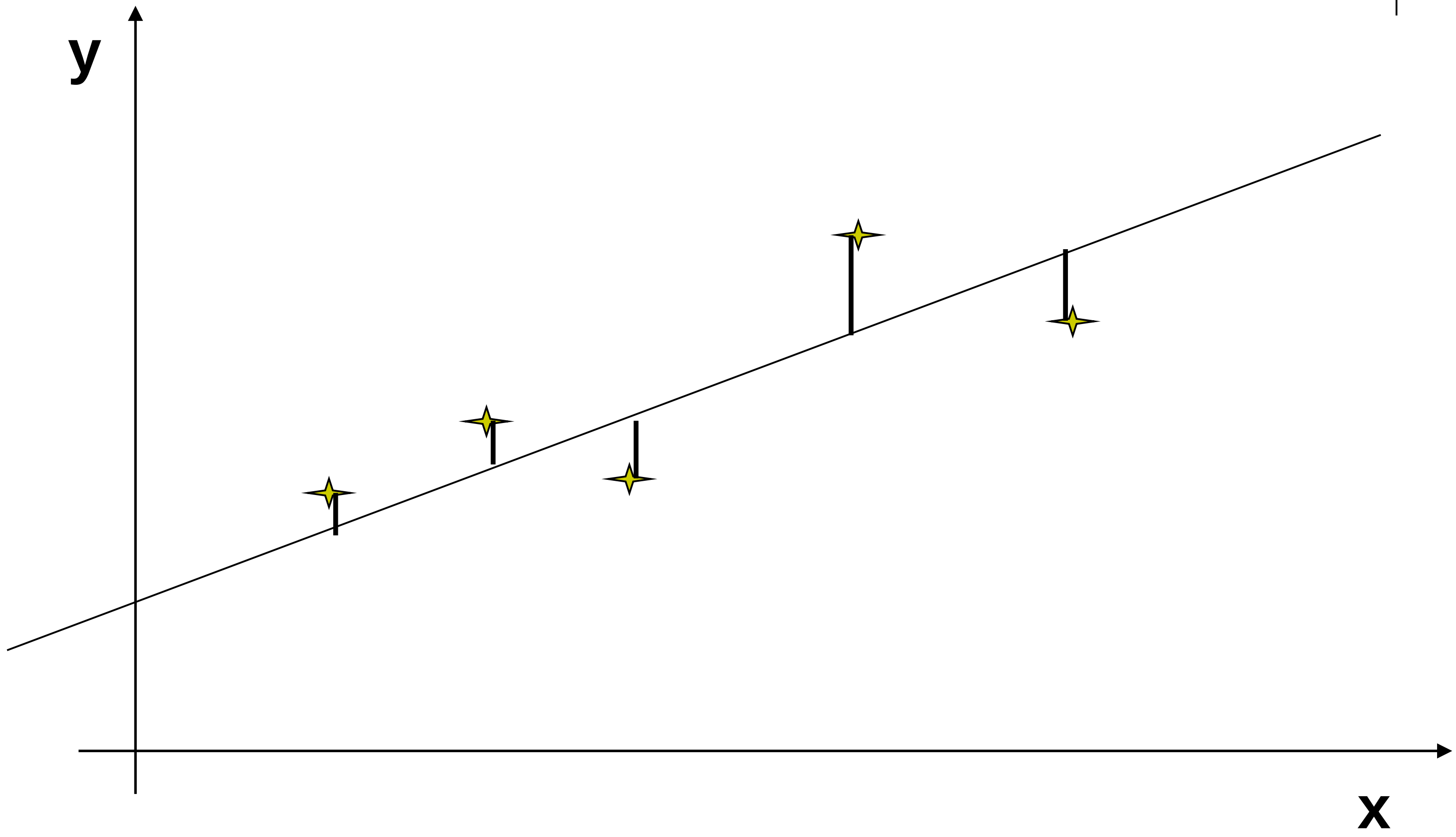
a_j parameters, $j=1..M$

Least square fit:

$$\min_{a_1 \ldots a_M} \left(\sum_{i=1}^{N} \left[y_i - y(x_i; a_1 \ldots a_M) \right]^2 \right)$$

Take derivatives by $a_1, a_2, \ldots, a_M$ -> set of linear equations (!! Also for some nonlin func.)

Least squares



Why not least absolute values, fourth powers, etc?

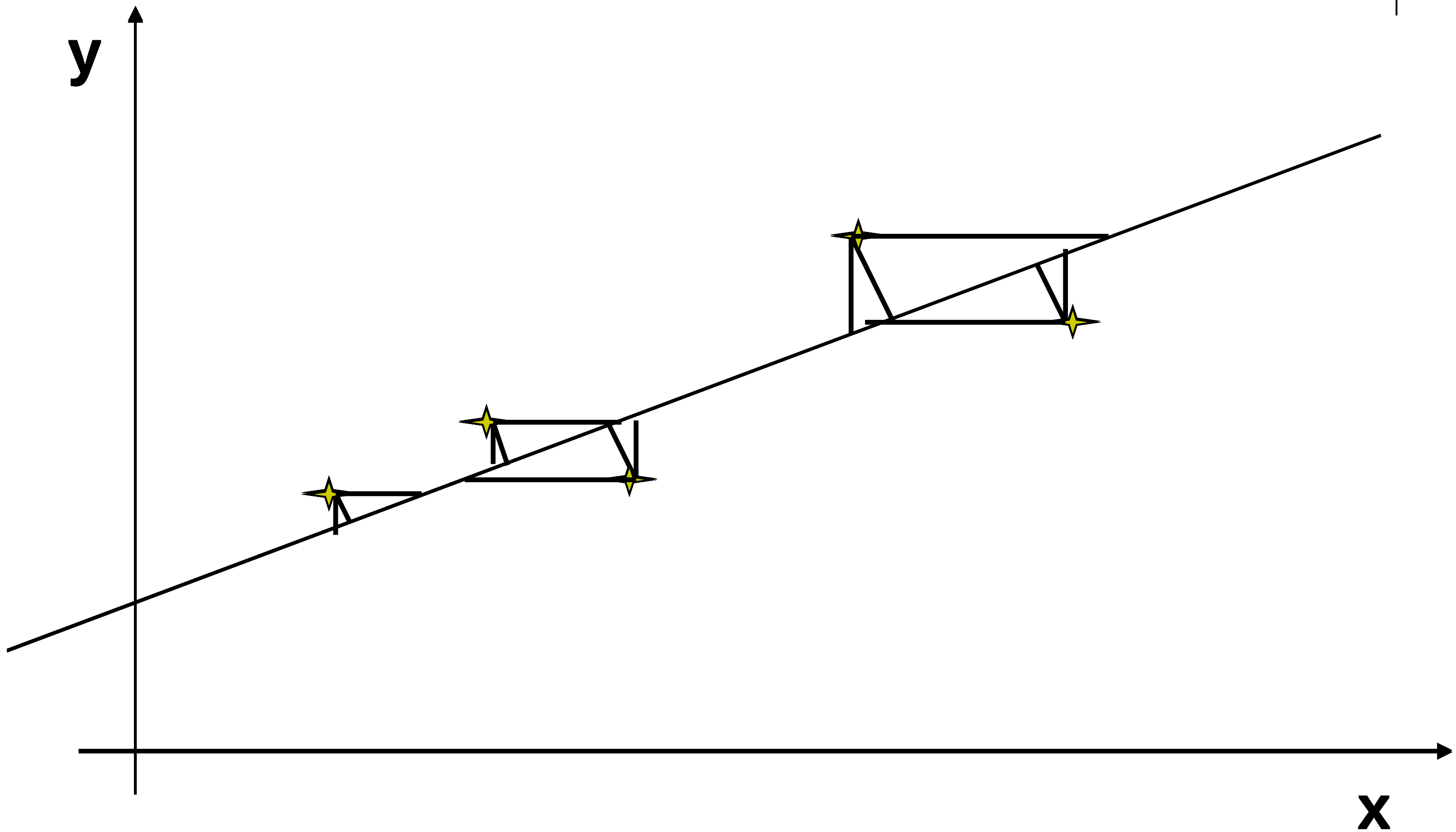
$$\min_{a_1 \ldots a_M} \left(\sum_{i=1}^{N} \left[y_i - y(x_i; a_1 \ldots a_M) \right]^2 \right)$$

$$\min_{a_1 \ldots a_M} \left(\sum_{i=1}^{N} \left| y_i - y(x_i; a_1 \ldots a_M) \right| \right)$$

$$\min_{a_1 \ldots a_M} \left(\sum_{i=1}^{N} \left[y_i - y(x_i; a_1 \ldots a_M) \right]^4 \right)$$

$$\min_{a_1 \ldots a_M} \left(\sum_{i=1}^{N} \sqrt[8]{\left| y_i - y(x_i; a_1 \ldots a_M) \right|} \right)$$

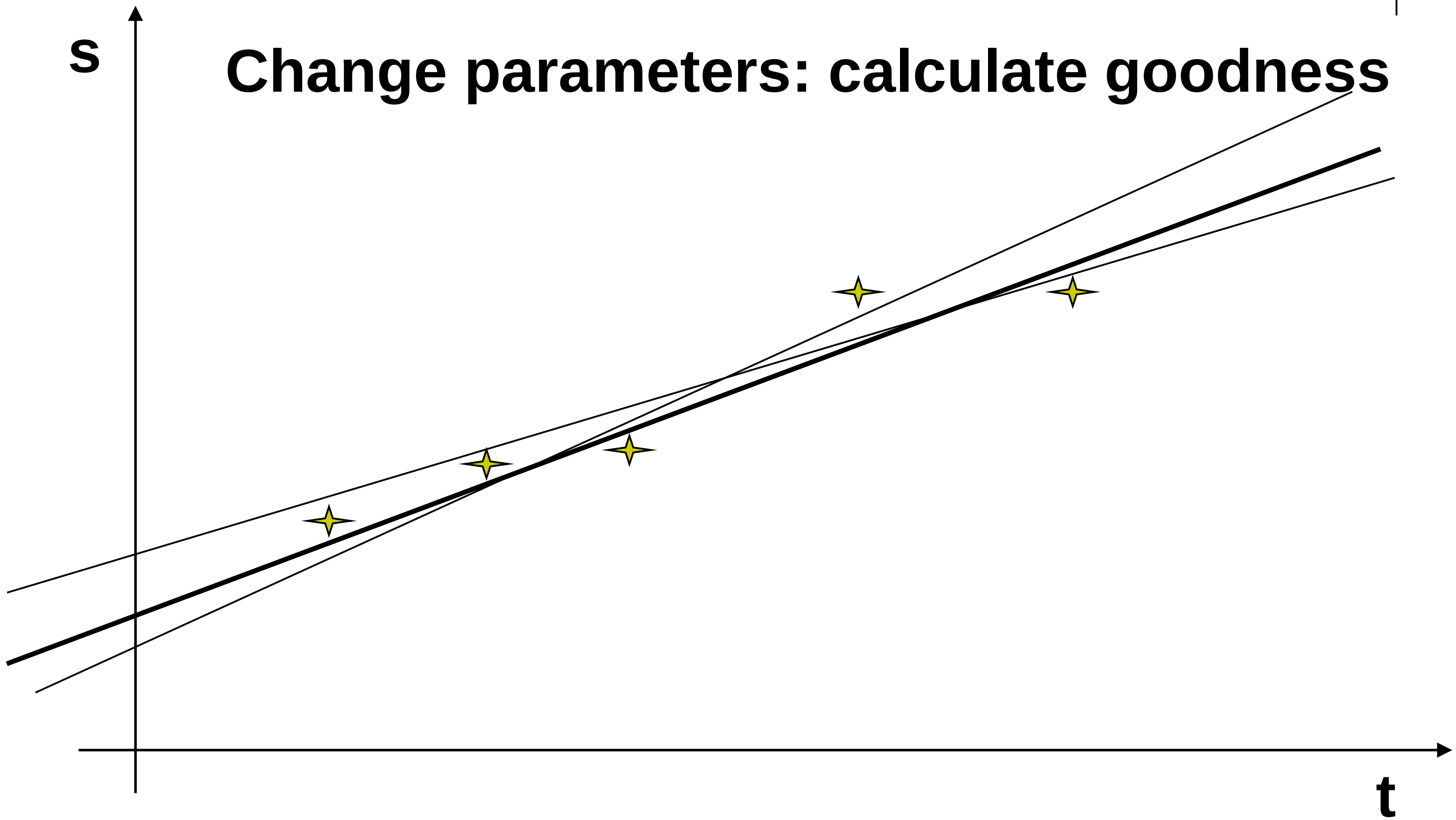
Least distances?



Why least squares?

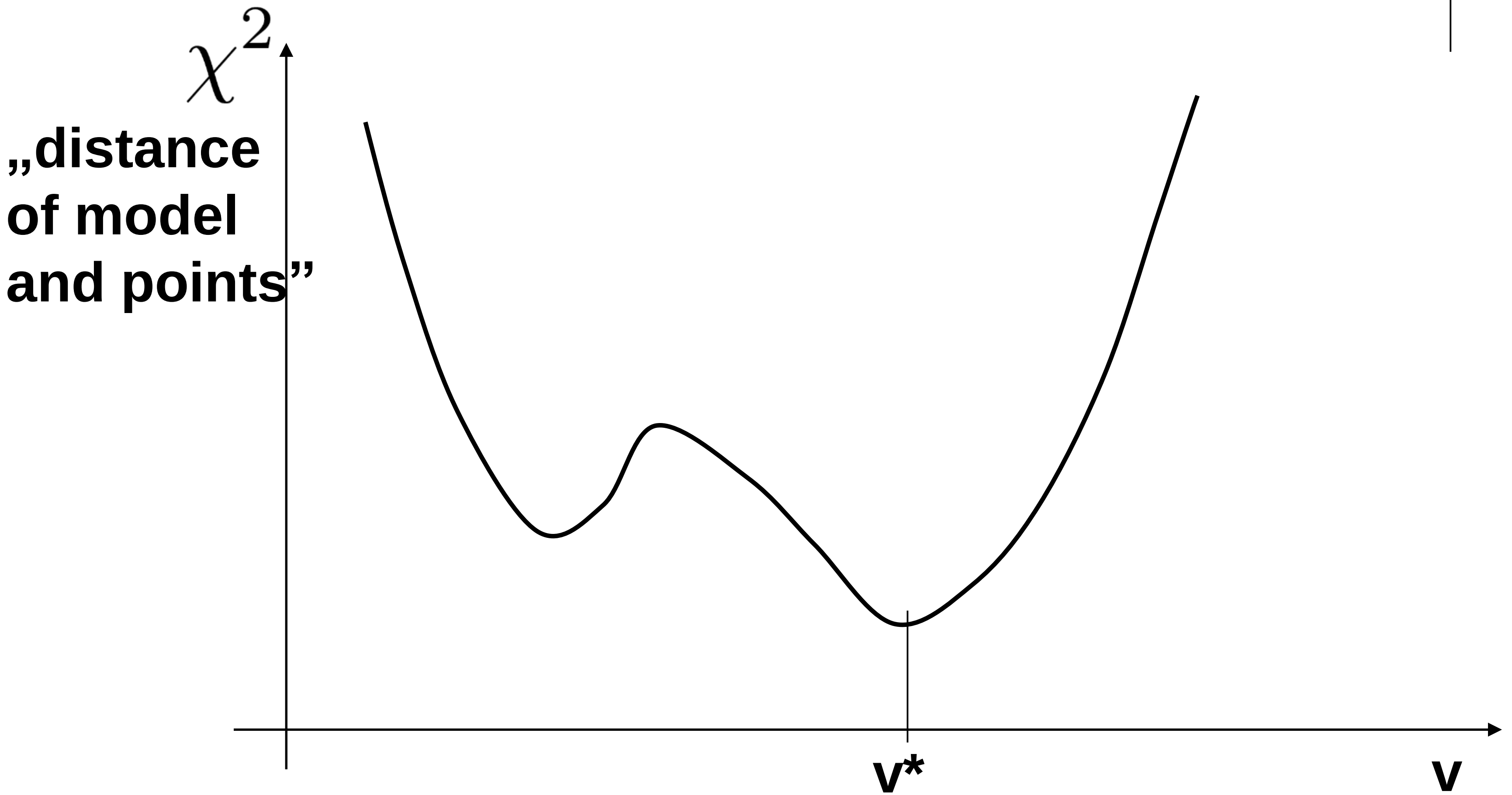
- Maximum likelihood estimation
- Least quare gives best parameters, if :
  - Errors only in y_i
  - Errors are normal (Gauss) distributed
  - Errors are independent

Cost function



Find optimal parameters,
with largest goodness / smallest „cost/energy”

Cost function



Finding minimum: in general leads to a nonlinear optimization problem

In special case: linear set of equations

Maximum likelihood

$$P \propto \prod_{i=1}^{N} \left\{ \exp\left[-\frac{1}{2} \left(\frac{y_i - y(x_i)}{\sigma} \right)^2 \right] \Delta y \right\}$$

Find parameters where P is maximal.

The log() func. is monotonuous, so find maximum of log(P), or minimum of –log(P)

-log(P)= $\left[\sum_{i=1}^{N} \frac{[y_i - y(x_i)]^2}{2\sigma^2} \right] - N \log \Delta y$

This is Least Squares!

Complex example: cosmology