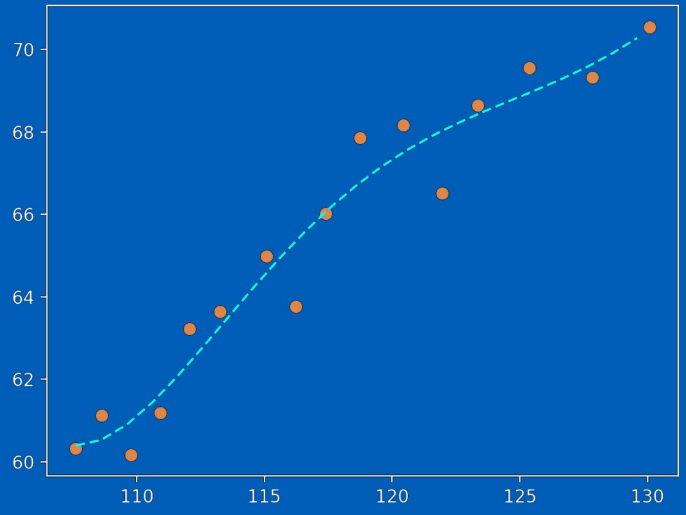
UNIVERSITÀ DI PARMA

Model Fitting



Summary

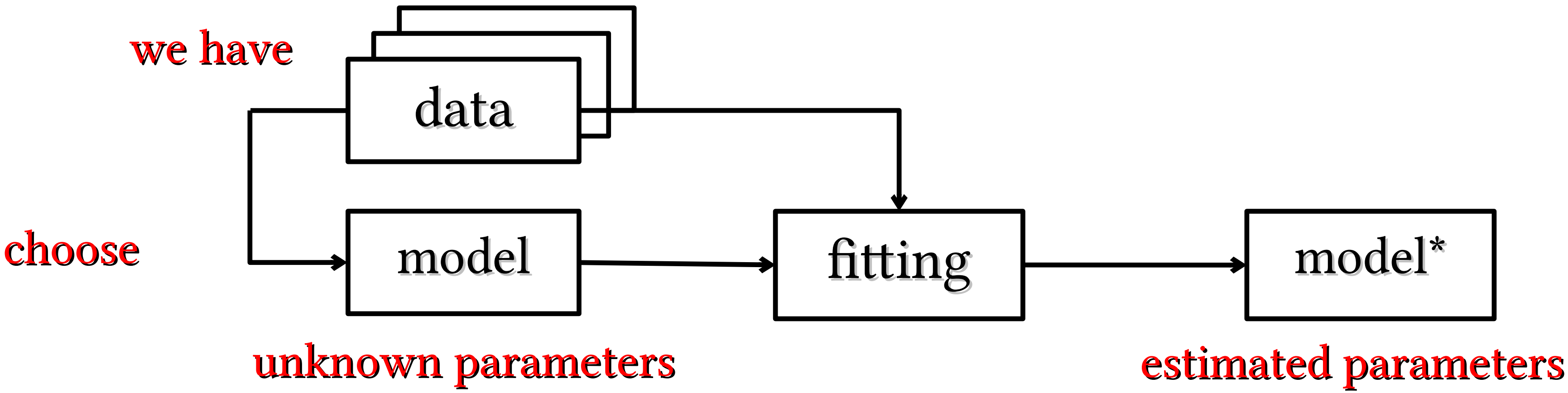
- What is Model Fitting?
- Is it difficult?
- Least Squares (LS)
  - Line fitting example
  - LS issues (outliers)
- Robust Statistics
- RANSAC

Model Fitting

- We have data
  - A lot of them
- Is there a model that generalize data distribution?
  - Choose a model
  - Compute model parameters

Model Fitting

UNIVERSITÀ DI PARMA

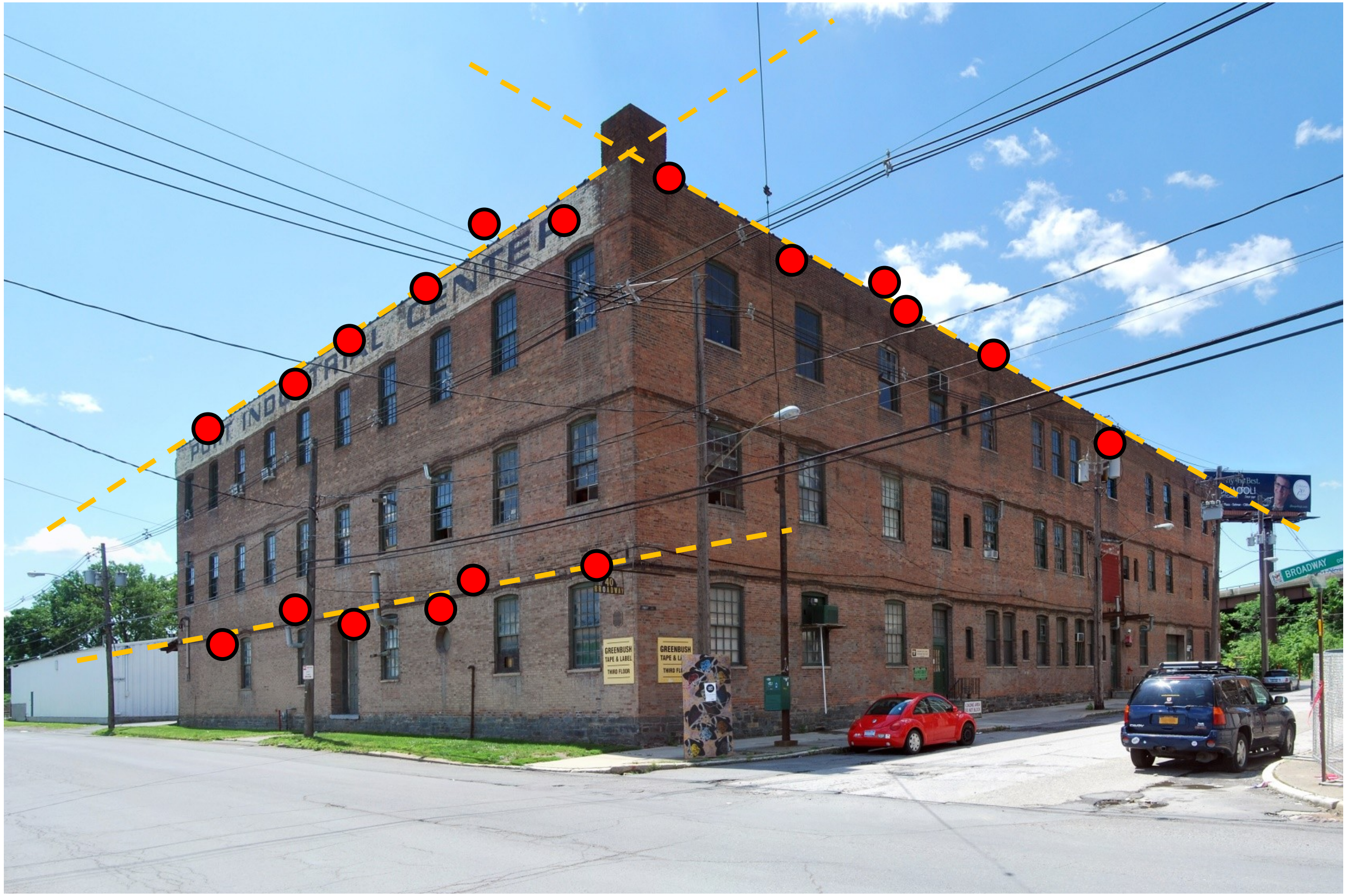


Models

UNIVERSITÀ DI PARMA

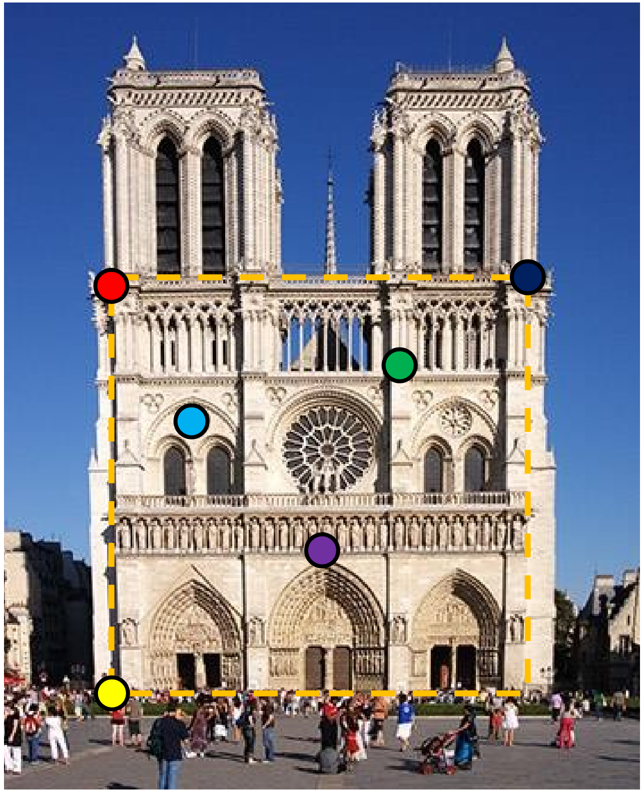
- Lines, curves, planes…
- Homographies, Calibration Matrices…
- …

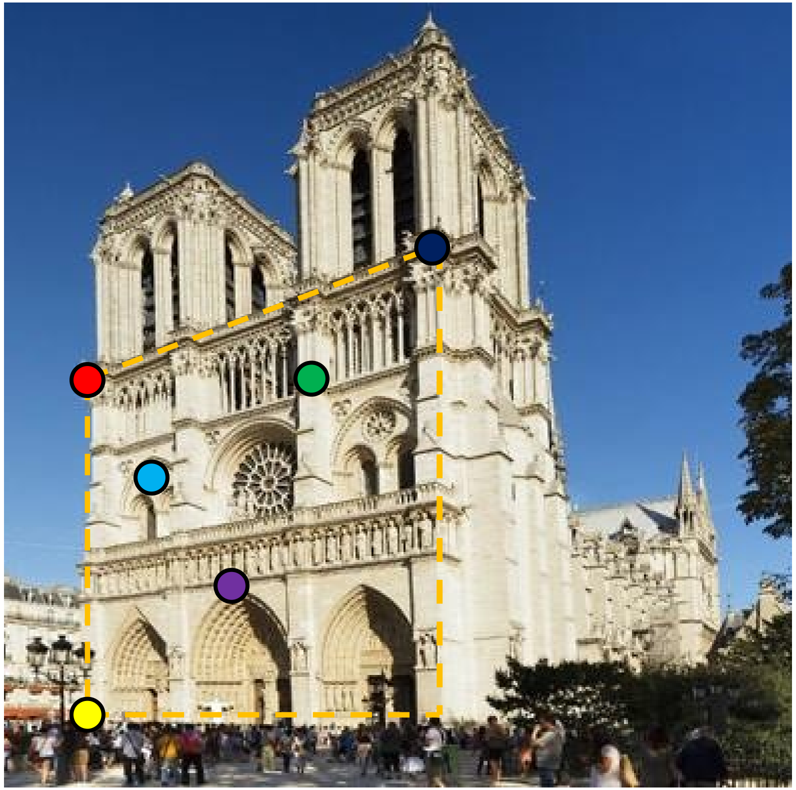
Example: lines fitting



Example: homograpy

UNIVERSITÀ DI PARMA





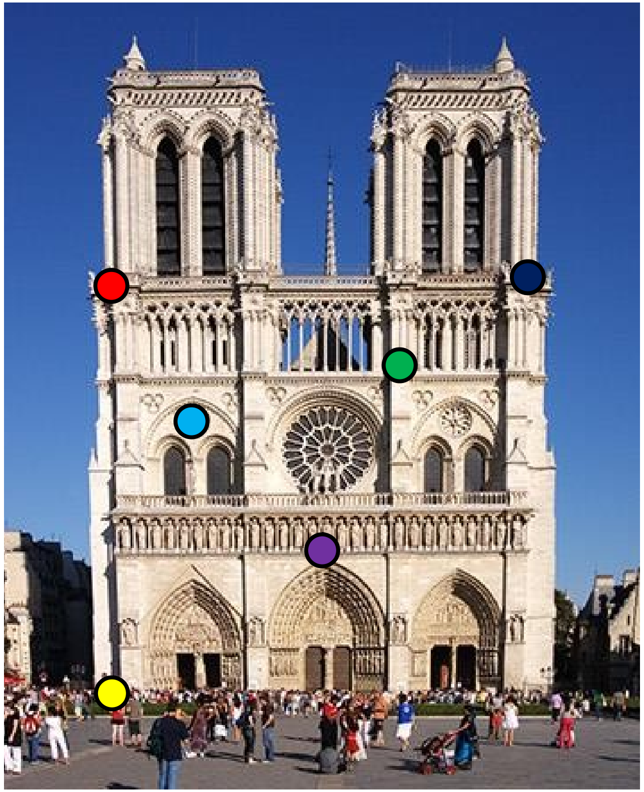
Problems

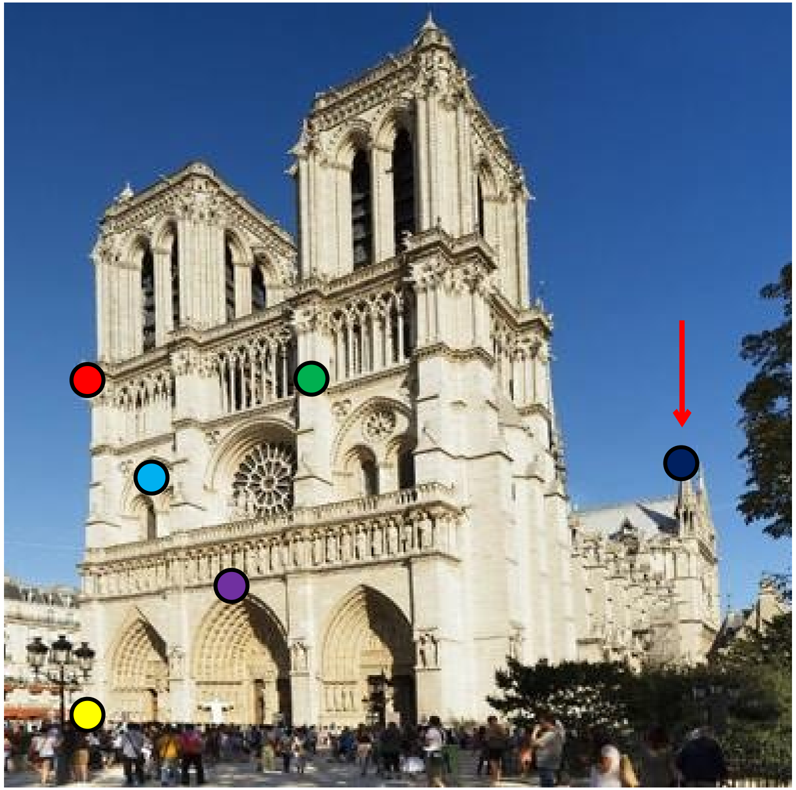
- Noise
- Outliers
- Missing data

Noisy data

UNIVERSITÀ
DI PARMA

Outliers



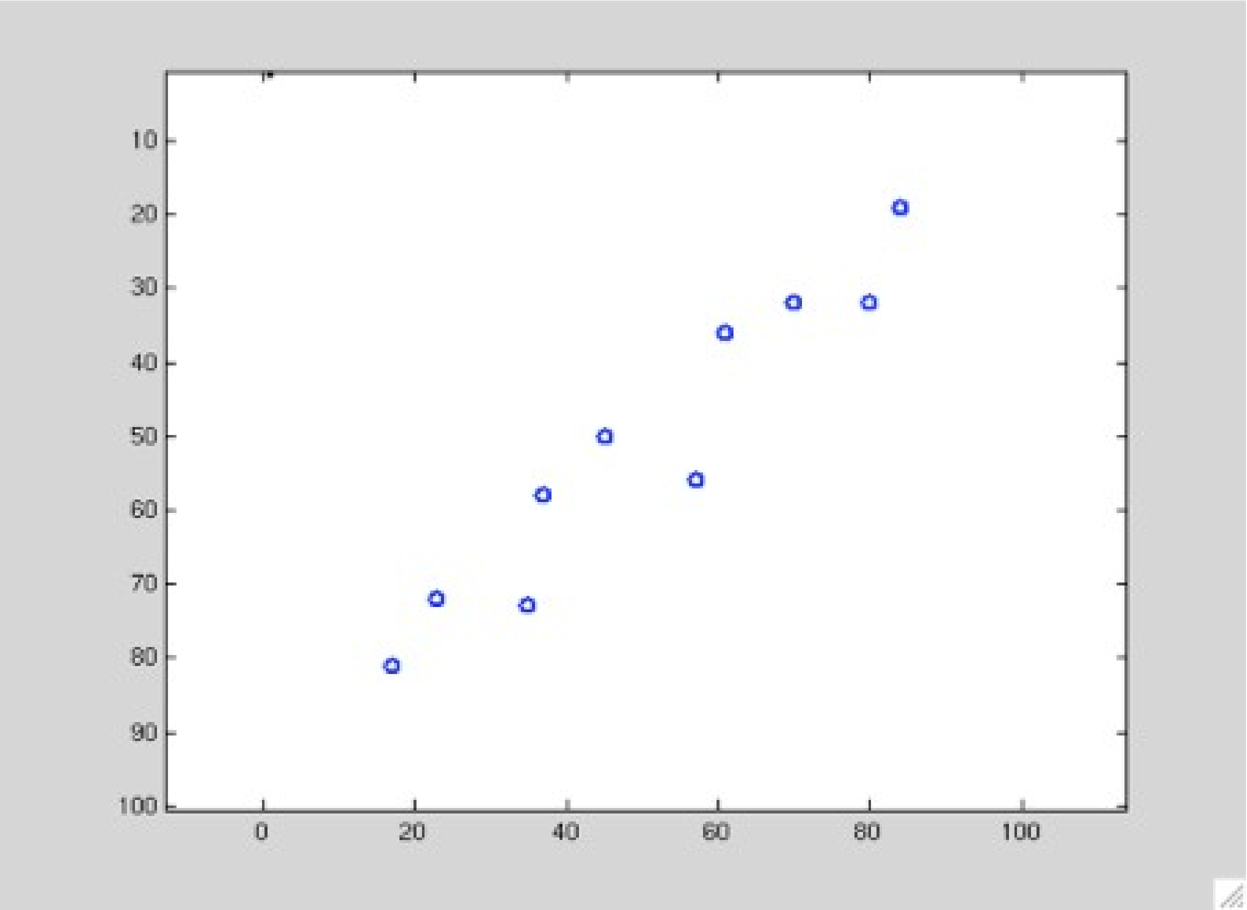


Fitting

UNIVERSITÀ DI PARMA

- We need to estimate parameters for a model that **optimally** generalize data distribution
- Potential approaches:
  - Least Squares (LS)
  - Robust Statistics
  - RANSAC

LS line fitting



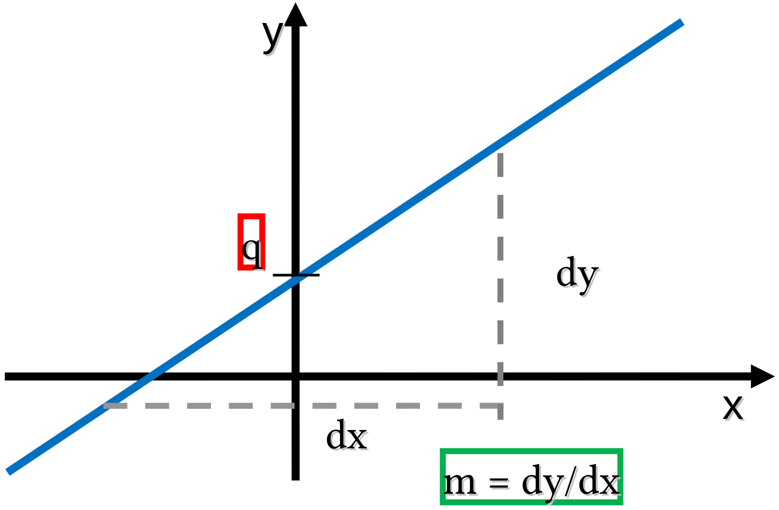
LS line fitting

UNIVERSITÀ DI PARMA

- Choose a model and its parameters
- Define a error function E($model_i$, data) to evaluate $model_i$
- Adapt model to data
  - find the parameters set that minimize E()

LS Line Fitting

- Model: $y = m\,x + q$
- Parameters:



LS line fitting

- Choose a model and its parameters
- Define a error function E($model_i$, data) to evaluate $model_i$
- Adapt model to data
  - find the parameters set that minimize E()

LS Line Fitting

- Error function is E($model_i$, data)
  - The value should indicate how much well $model_i$ fits given data
- For LS Error Function is the sum of **squared residuals**

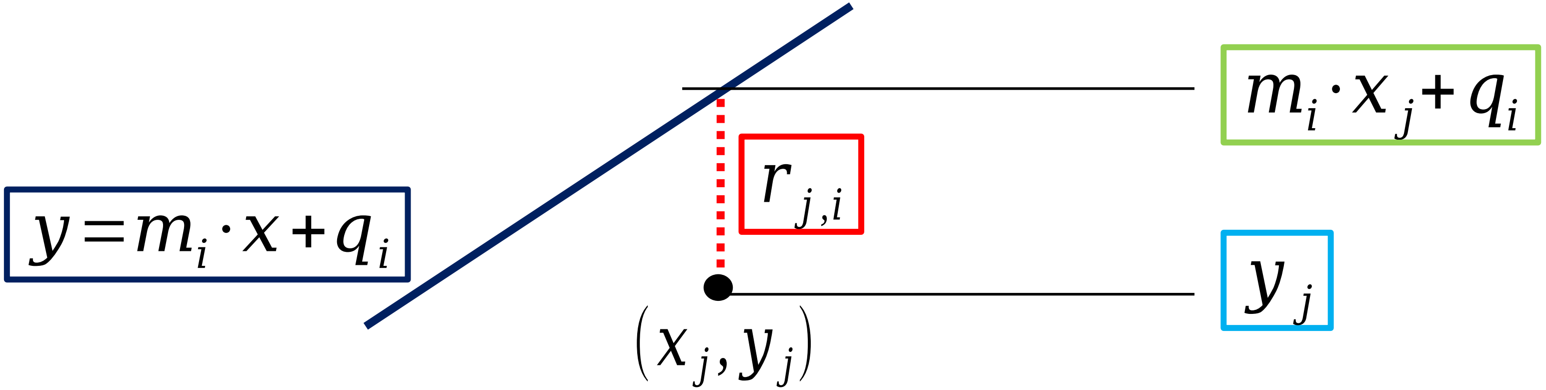
$$E(model_i, data) = \sum_{j=1}^{n} r_{i,j}^2$$

- $r_{i,j}$ is the "difference" between $model_i$ and $data_j$
- We need to define the residual *function*

LS Line Fitting

- For a line fitting we can define residuals as

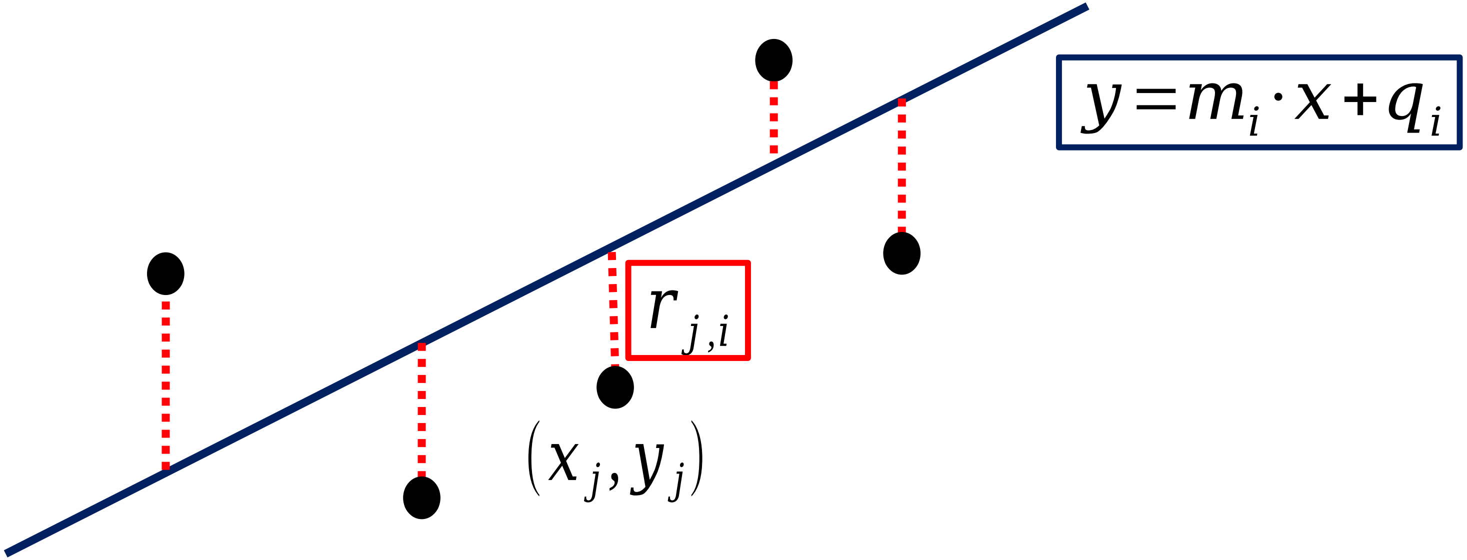
$$\underbrace{r_{i,j}}_{\text{Distance between data}_j\text{ and model}_i} = \underbrace{y_i}_{\text{Real value of data}_j} - \underbrace{(m_i \cdot x_j + q_i)}_{\text{Estimated value for data}_j\text{ given by model}_i}$$



LS Line Fitting

- Given Error Function definition:

$$E\left(model_i, data\right) = \sum_{j=1}^{n} r_{i,j}^2 = \sum_{j=1}^{n} \left[y_j - \left(m_i \cdot x_j + q_i\right)\right]^2$$

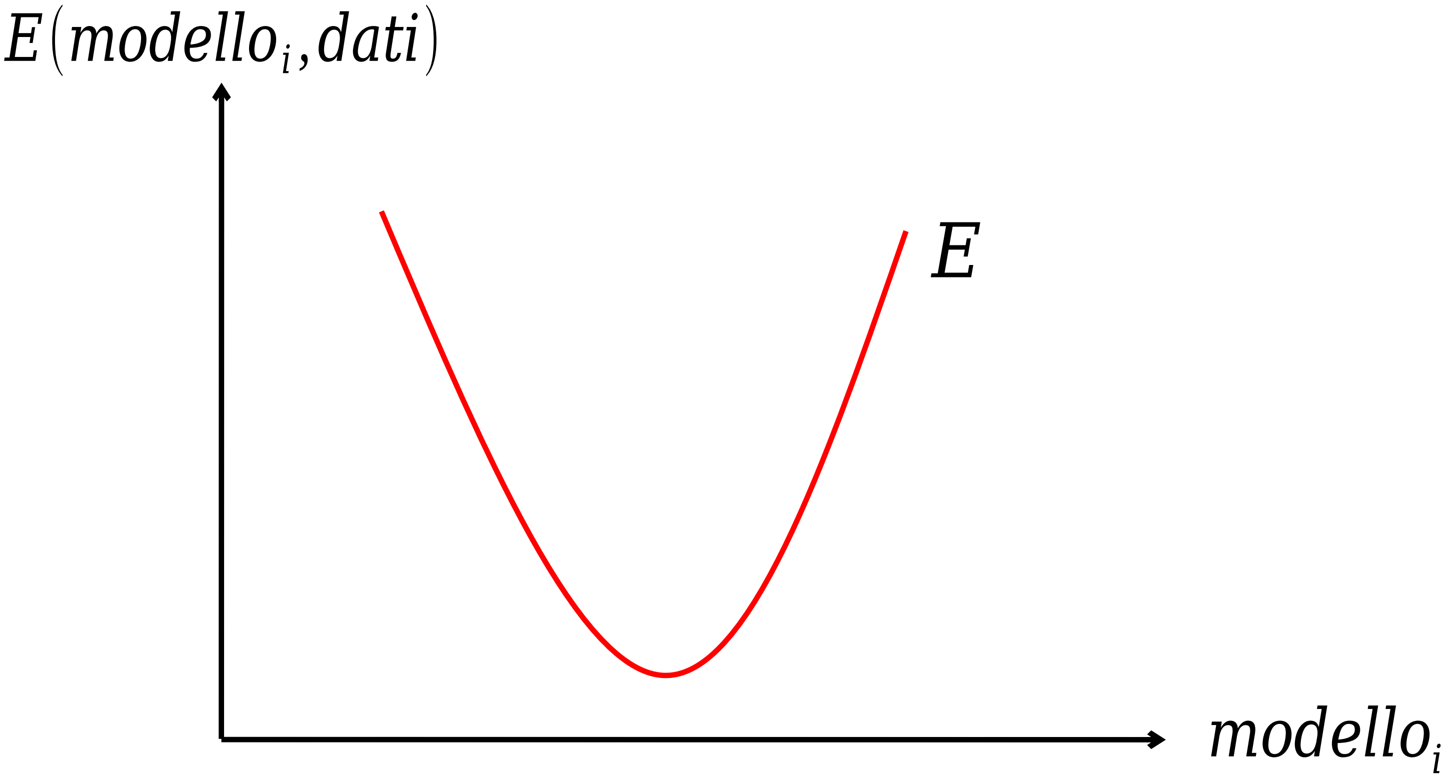


LS line fitting

- Choose a model and its parameters
- Define a error function E($model_i$, data) to evaluate $model_i$
- Adapt model to data
  - find the parameters set that minimize E()

LS Line Fitting

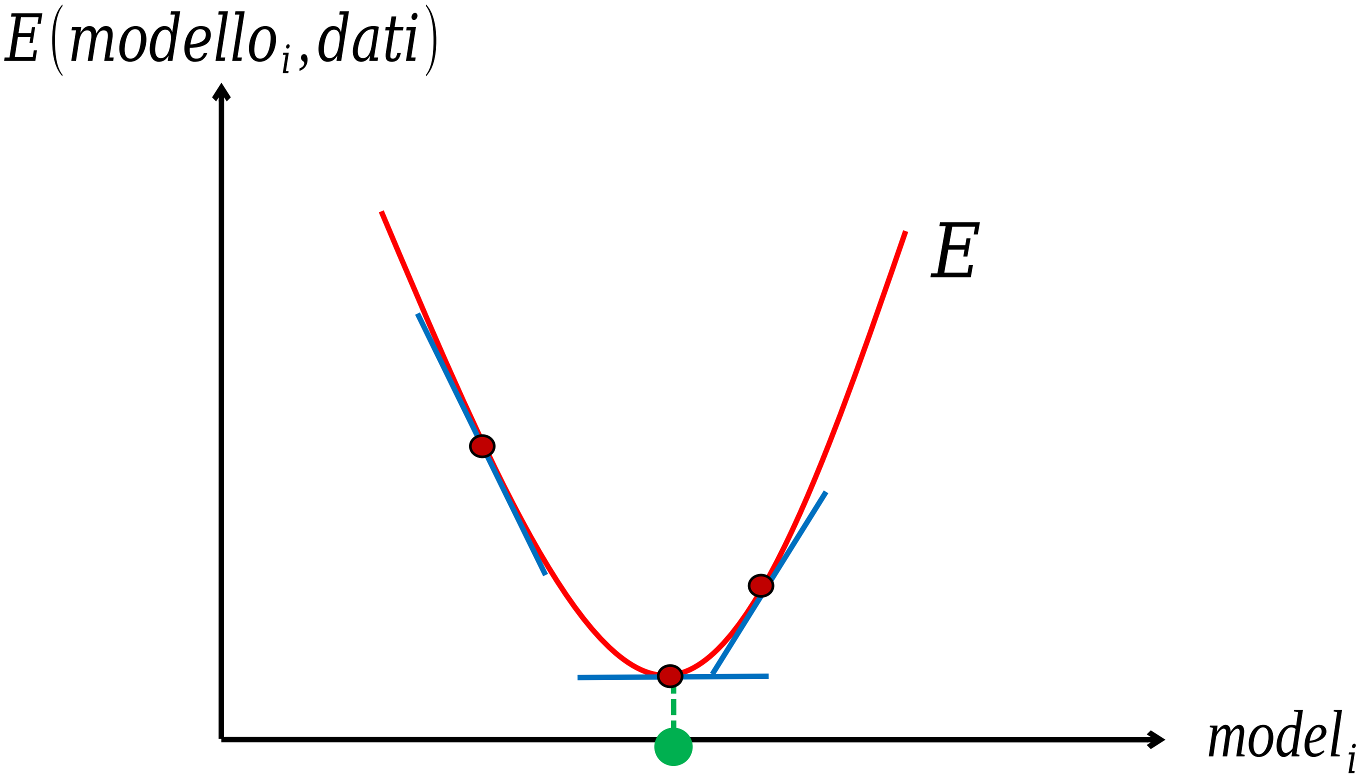
- Consider E() graph



LS Line Fitting

UNIVERSITÀ DI PARMA

- We need to find the parameter set that minimize E()

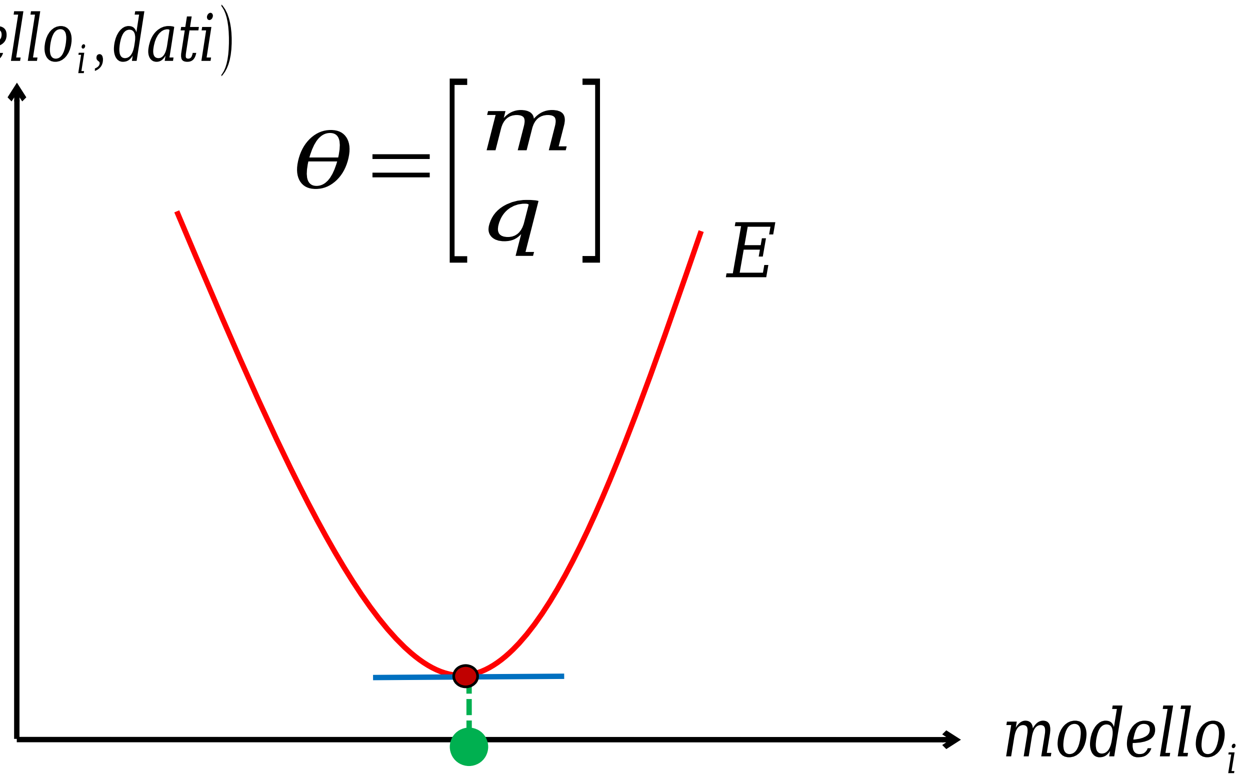


LS Line Fitting

- Compute gradient and try to obtain 0

$$\frac{\partial E}{\partial \theta} = \begin{bmatrix} \dfrac{\partial E}{\partial m} \\ \dfrac{\partial E}{\partial q} \end{bmatrix} = 0$$

$$\theta = \begin{bmatrix} m \\ q \end{bmatrix}$$



LS Line Fitting

UNIVERSITÀ DI PARMA

- Partial derivative of E with respect to m

$$\frac{\partial E}{\partial m} = \frac{\partial}{\partial m} \sum_{j=1}^{n} \left(y_j - (m x_j + q)\right)^2 = 0$$

$$.= \frac{\partial}{\partial m} \sum_{j=1}^{n} \left(y_j^2 - 2 y_j (x_j m + q) + (m x_j + q)^2\right)$$

$$.= \sum_{j=1}^{n} \frac{\partial}{\partial m} \left(y_j^2 - 2 x_j y_j m - 2 y_j q + x_j^2 m^2 + 2 x_j q m + q^2\right)$$

$$.= \sum_{j=1}^{n} \left(2 x_j^2 m - 2 x_j y_j + 2 x_j q\right)$$

LS Line Fitting

- Partial derivative of E with respect to m

$$\frac{\partial E}{\partial m}=\sum_{j=1}^{n}\left(2x_j^2 m-2x_j y_j+2x_j q\right)=0$$

$$\left(\sum_{j=1}^{n} x_j^2\right)m+\left(\sum_{j=1}^{n} x_j\right)q=\sum_{j=1}^{n} x_j y_j \qquad \text{(Eq. 1)}$$

LS Line Fitting

- Partial derivative of E with respect to q

$$\frac{\partial E}{\partial q} = \frac{\partial}{\partial q} \sum_{j=1}^{n} \left(y_j - (m x_j + q)\right)^2 = 0$$

$$.= \frac{\partial}{\partial q} \sum_{j=1}^{n} \left(y_j^2 - 2 y_j (x_j m + q) + (m x_j + q)^2\right)$$

$$.= \sum_{j=1}^{n} \frac{\partial}{\partial q} \left(y_j^2 - 2 x_j y_j m - 2 y_j q + x_j^2 m^2 + 2 x_j q m + q^2\right)$$

$$.= \sum_{j=1}^{n} \left(2 x_j^2 m + 2 q - 2 y_j\right)$$

LS Line Fitting

- Partial derivative of E with respect to q

$$\frac{\partial E}{\partial q} = \sum_{j=1}^{n} \left(2x_j m + 2q - 2y_j \right) = 0$$

$$\left(\sum_{j=1}^{n} x_j \right) m + \underbrace{\sum_{j=1}^{n} 1}_{n} q = \left(\sum_{j=1}^{n} x_j \right) m + nq = \sum_{j=1}^{n} y_j \qquad \text{(Eq. 2)}$$

LS Line Fitting

- We can write equations 1 & 2 in a matrix form

$$\begin{bmatrix} \sum x_j^2 & \sum x_j \\ \sum x_j & n \end{bmatrix} \begin{bmatrix} m \\ q \end{bmatrix} = \begin{bmatrix} \sum x_j y_j \\ \sum y_j \end{bmatrix}$$

LS Line Fitting

$$X = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \qquad \theta = \begin{bmatrix} m \\ q \end{bmatrix}$$

LS Line Fitting

UNIVERSITÀ DI PARMA

- We can see previous equation as:

$$\underbrace{\begin{bmatrix} \sum x_j^2 & \sum x_j \\ \sum x_j & n \end{bmatrix}}_{X^T X} \underbrace{\begin{bmatrix} m \\ q \end{bmatrix}}_{\theta} = \underbrace{\begin{bmatrix} \sum x_j y_j \\ \sum y_j \end{bmatrix}}_{X^T Y}$$

LS Line Fitting

UNIVERSITÀ DI PARMA

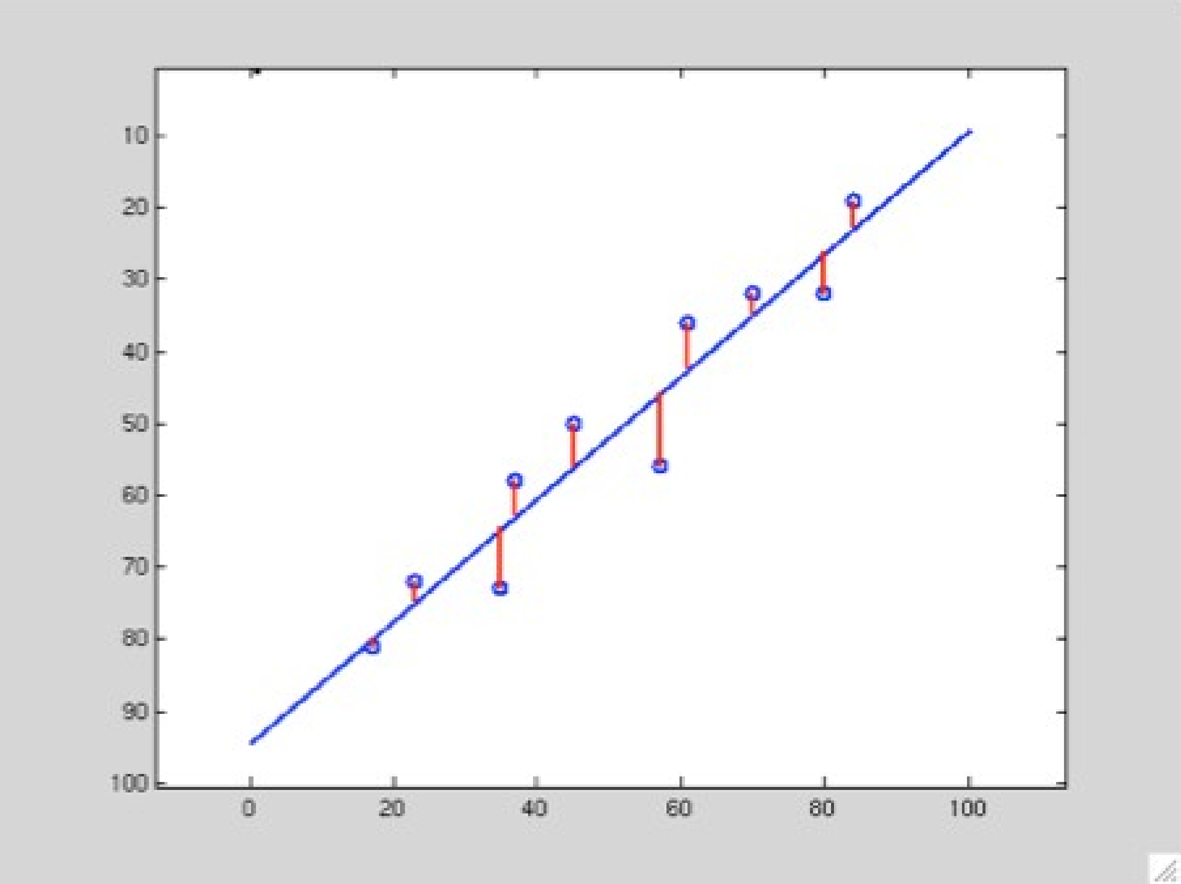
- We can then solve this as

$$X^T X \theta = X^T Y$$

$$\Downarrow$$

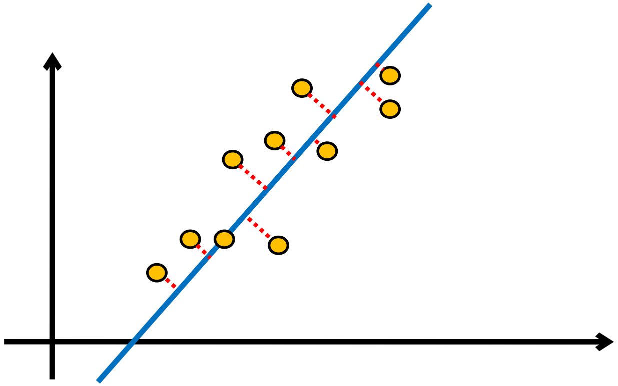
$$\theta = \left(X^T X\right)^{-1} X^T Y$$

LS Line Fitting



LS Line Fitting – algebraic distance

- (m,q) are not always the best parameter choiche
  - Vertical lines?
- A more general approach is $ax+by+c=0$
- We can also use algebraic distance as estimator for fitting



LS Line Fitting – algebraic distance

- Cost function E can now be the actual distance

$$E_i = \sum_{j=1}^{n} \left(a_i x_j + b_i y_j + c_i \right)^2$$

- Let's find the parameter set that minimize E_i

LS Line Fitting – algebraic distance

UNIVERSITÀ DI PARMA

- Partial derivatives can be written as

$$\frac{\partial E}{\partial a} = \sum_{j=1}^{n} \left(2x_j^2 a + 2x_j y_j b + 2x_j c\right)$$

$$\frac{\partial E}{\partial b} = \sum_{j=1}^{n} \left(2x_j y_j a + 2y_j^2 b + 2y_j c\right)$$

$$\frac{\partial E}{\partial c} = \sum_{j=1}^{n} \left(2x_j a + 2y_j b + 2c\right)$$

LS Line Fitting – algebraic distance

- Again, put them as homogeneous equation

$$\underbrace{\begin{bmatrix} \sum x_j^2 & \sum x_j y_j & \sum x_j \\ \sum x_j y_j & \sum y_j^2 & \sum y_j \\ \sum x_j & \sum y_j & n \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}_{0}$$

$$A \cdot x = 0$$

LS Line Fitting – algebraic distance

UNIVERSITÀ DI PARMA

- Then we should solve the homogeneous equation

$$A \cdot x = 0$$

- There is always a solution?
  - No!
- Moreover, ill posed when $A \in R_{3\times3}$ has a rank<3

LS Line Fitting – algebraic distance

UNIVERSITÀ DI PARMA

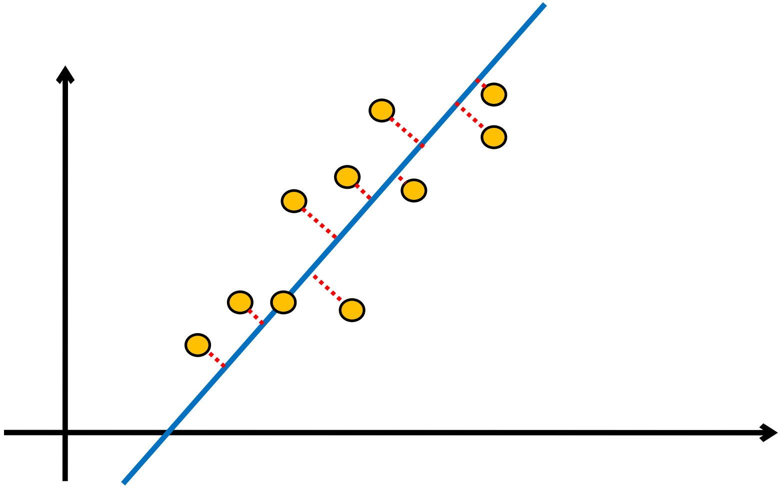
- We can use (again) SVD for a minimization

$$A = UDV^T$$

- x is the last column of V

LS Summary

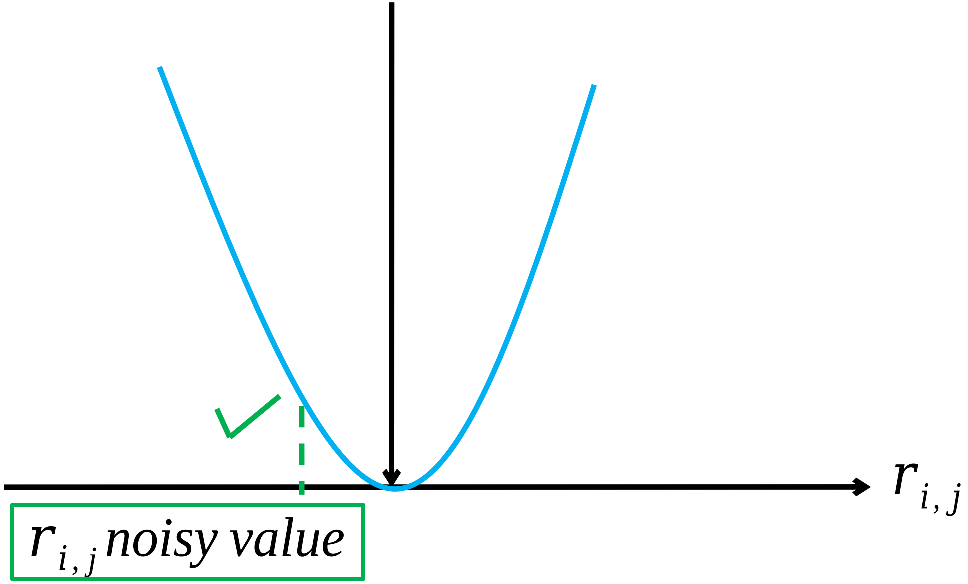
- Summary
  - Simple approach
    - n ≤ 2 ⟶ closed form
    - n > 2 ⟶ SVD
  - No thresholds
  - Results is good when until we have Gaussian noise…



LS issue: outliers

- Not always Gaussian noise
- LS is highly affected by outliers

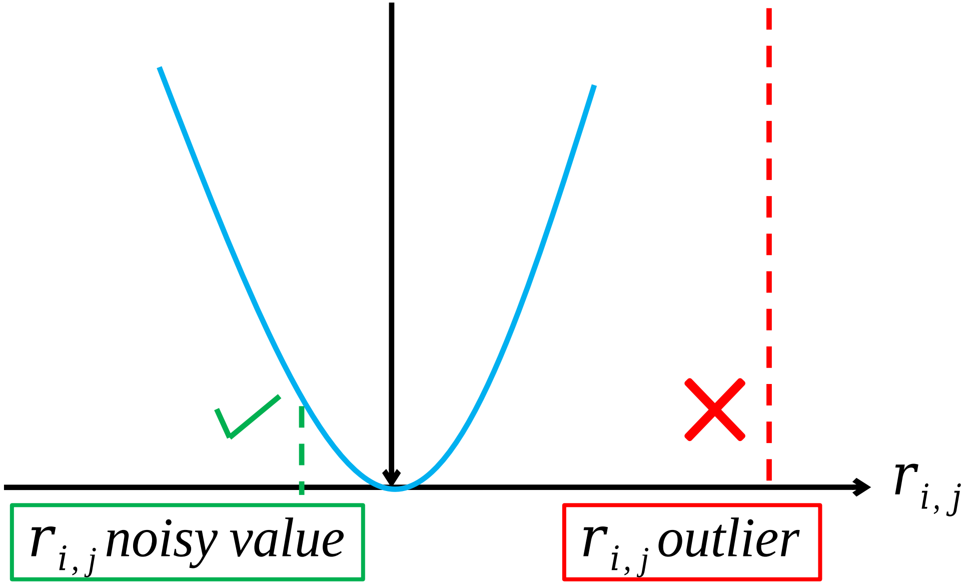
$$f\left(r_{i,j}\right) \propto r_{i,j}^2$$



LS issue: outliers

- Not always Gaussian noise
- LS is highly affected by distant points

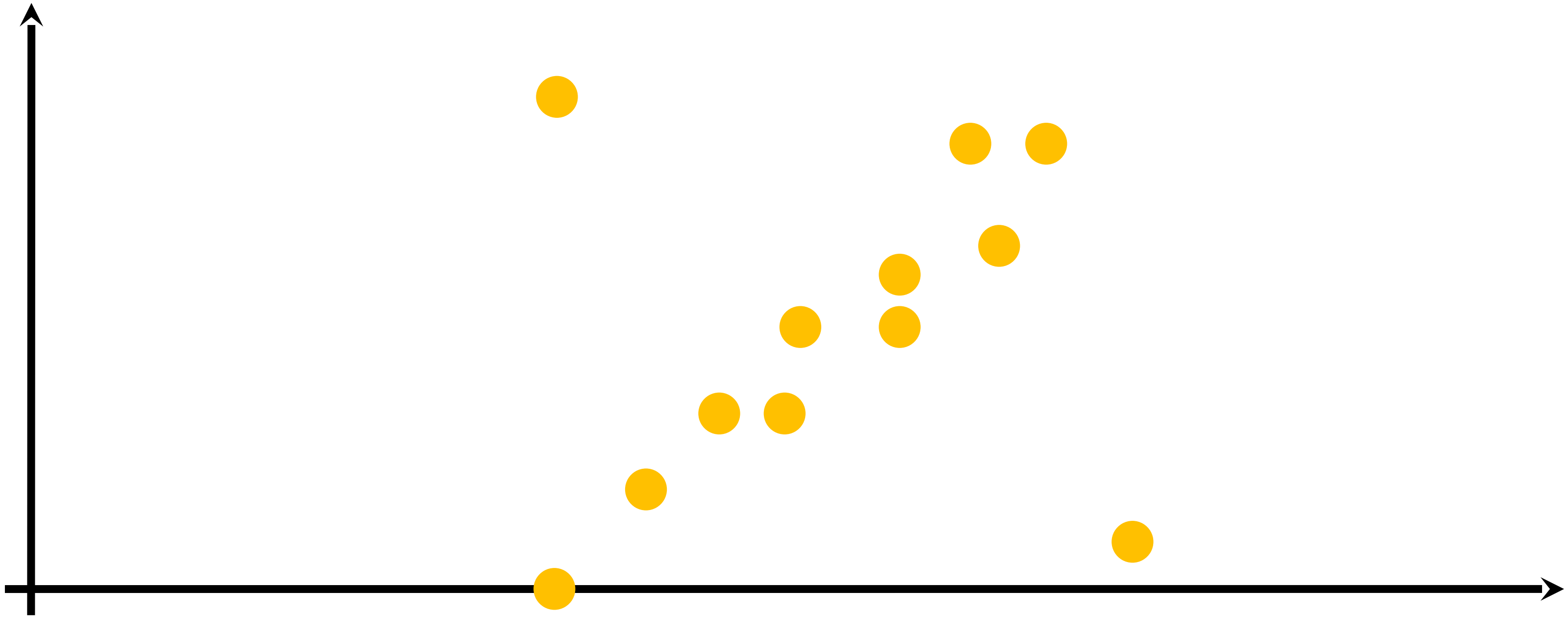
$$f\left(r_{i,j}\right) \propto r_{i,j}^2$$



LS issue: outliers

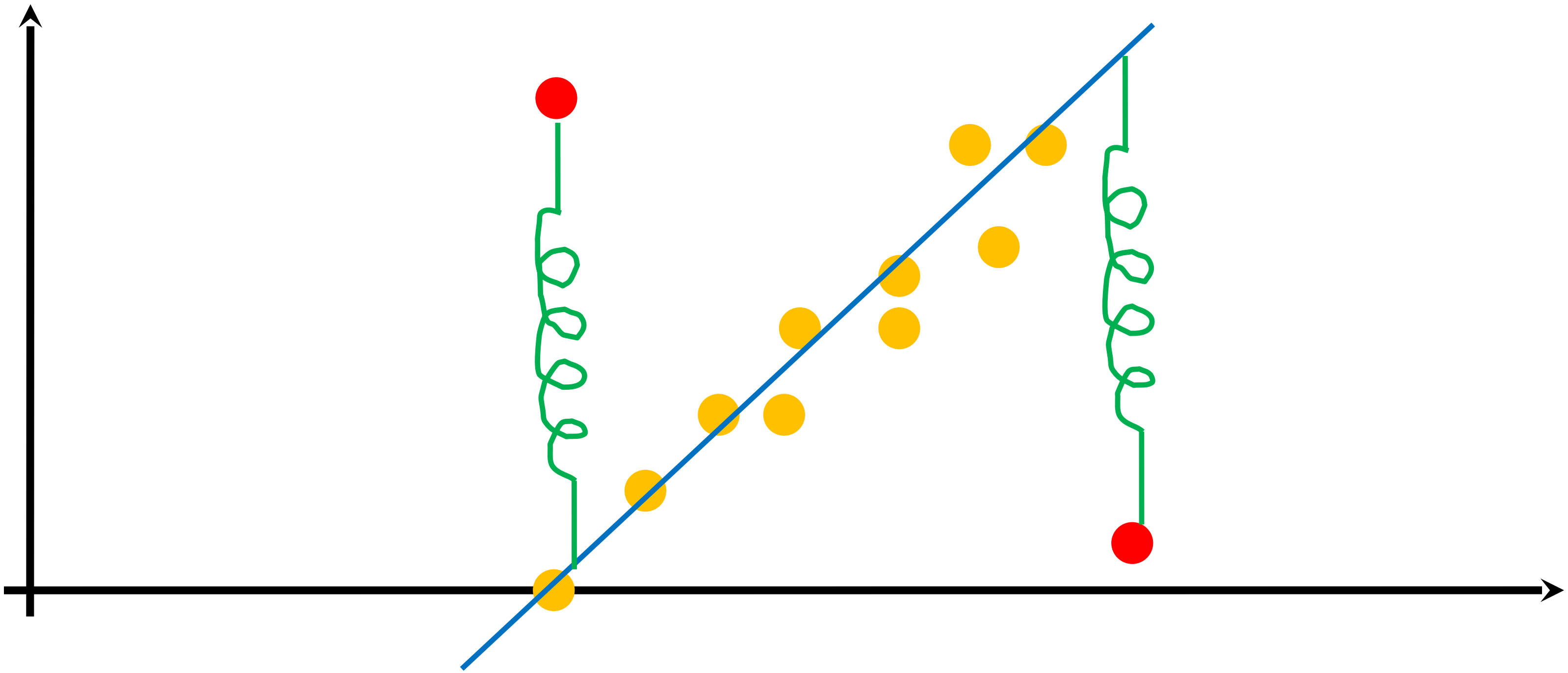
UNIVERSITÀ DI PARMA

- Example: what is the right fitting in this case?



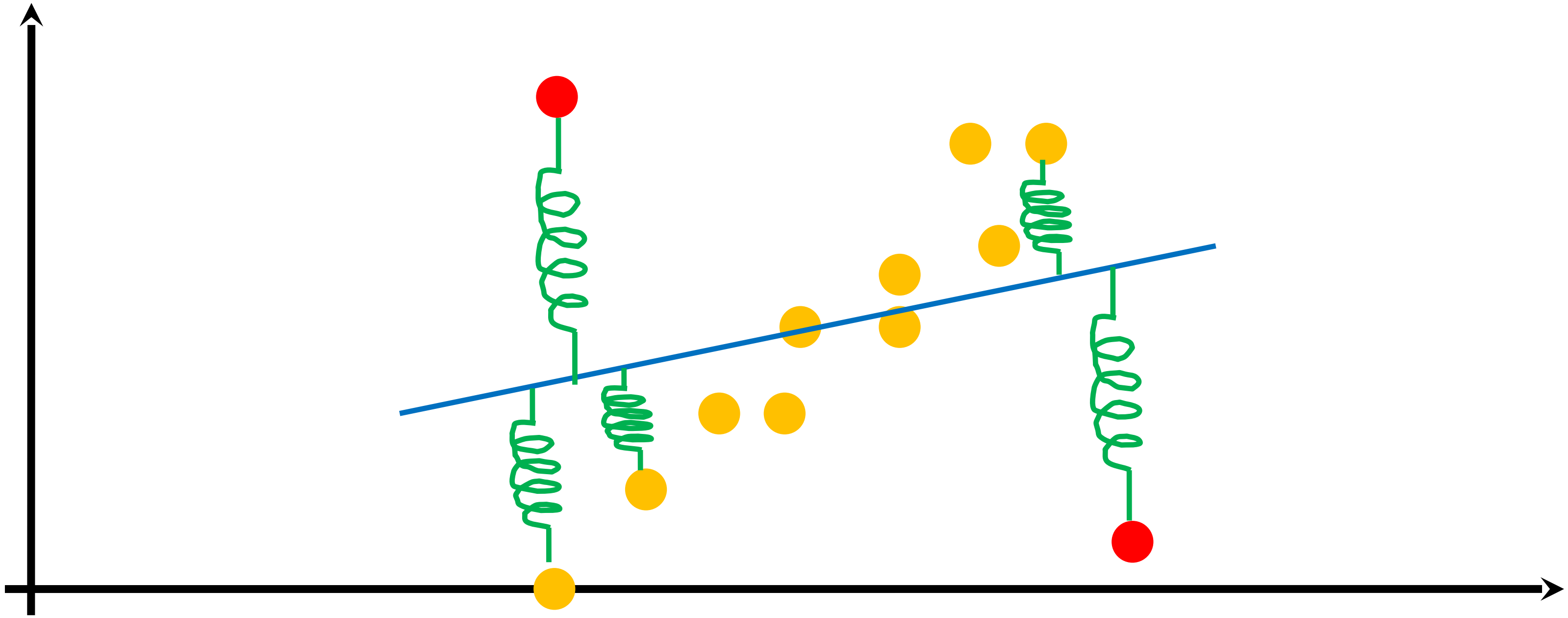
LS issue: outliers

- Example: what is the right fitting in this case?



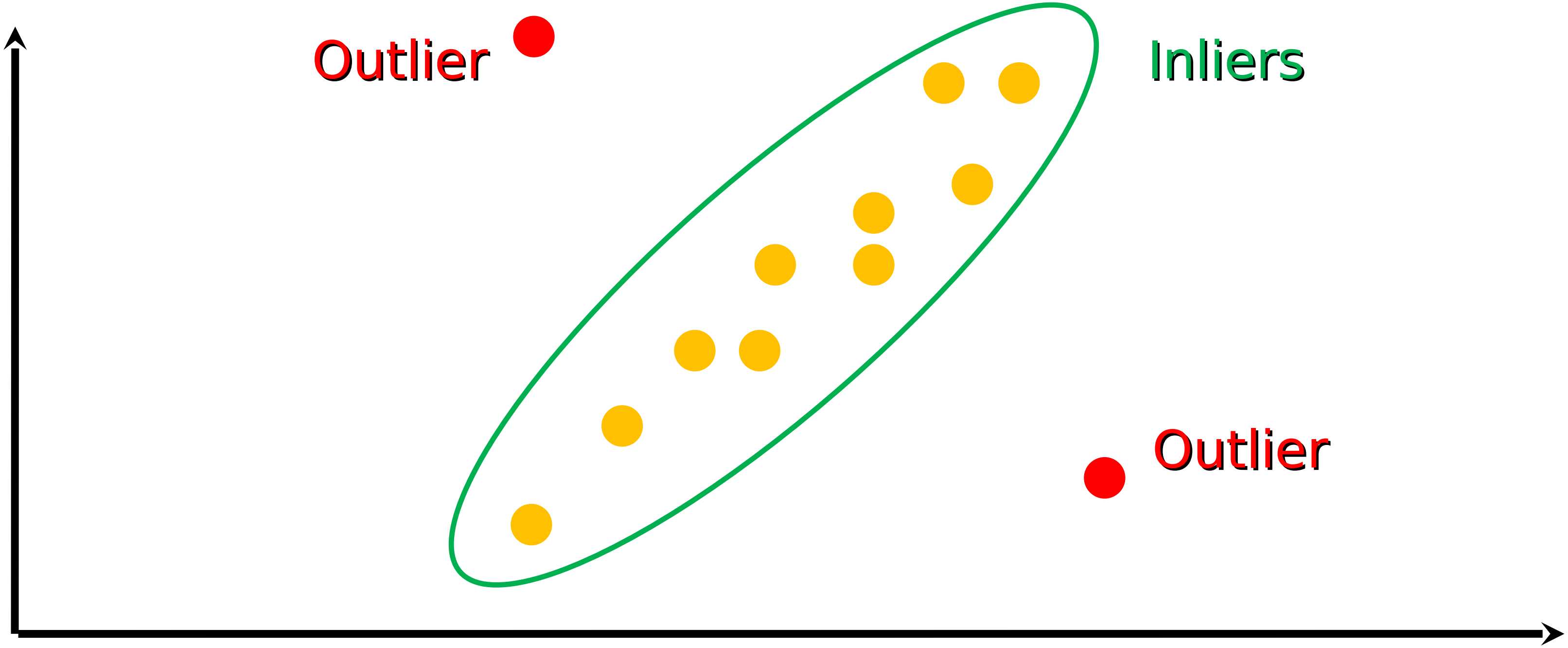
LS issue: outliers

- Example: what is the right fitting in this case?



Outliers

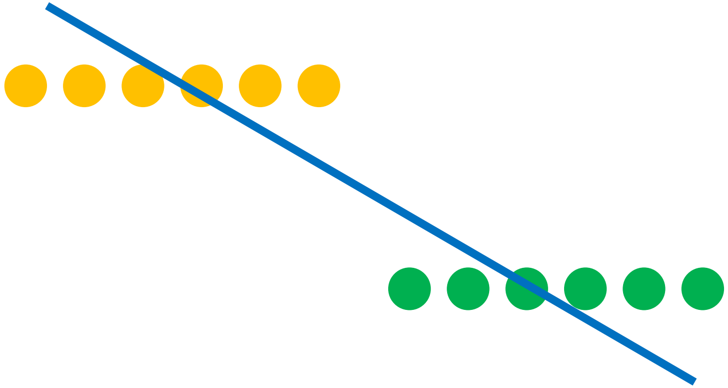
- Inliers: data which belong to the model
- Outliers: data not belonging to the model



Wrong model

UNIVERSITÀ DI PARMA

- Outliers are not the only issue



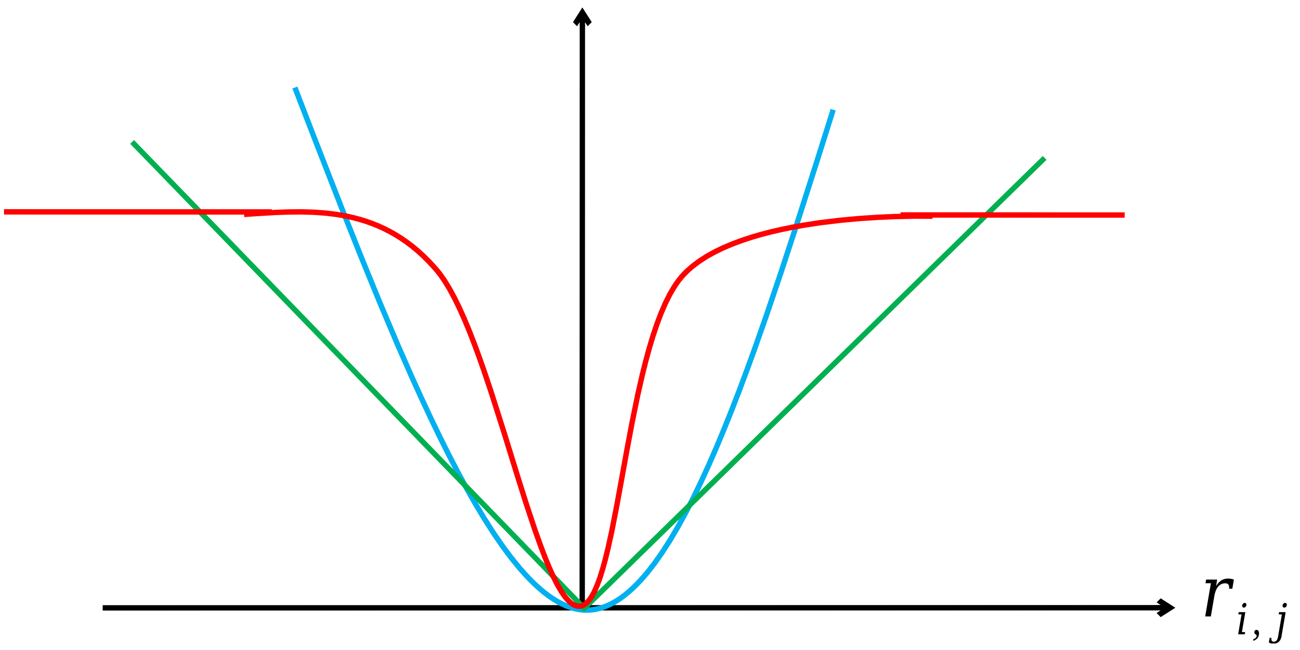
Robust Statistics

- Underlying idea: modify data weight depending on their distance from the model

$$f\left(r_{i,j}\right) \propto r_{i,j}^2$$

$$f\left(r_{i,j}\right) \propto \left|r_{i,j}\right|$$

$$f\left(r_{i,j}\right) \propto \frac{r_{i,j}^2}{r_{i,j}^2 + \sigma^2}$$



Robust Statistics

- Instead of squared residuals sum we use a function ρ()

$$min \sum_{j=1}^{n} r_i^2 \quad \rightarrow \quad min \sum_{j=1}^{n} \rho(r_i)$$

- The function is symmetrical and features a minimum in zero
  - Loss function

Robust Statistics RANSAC

UNIVERSITÀ DI PARMA

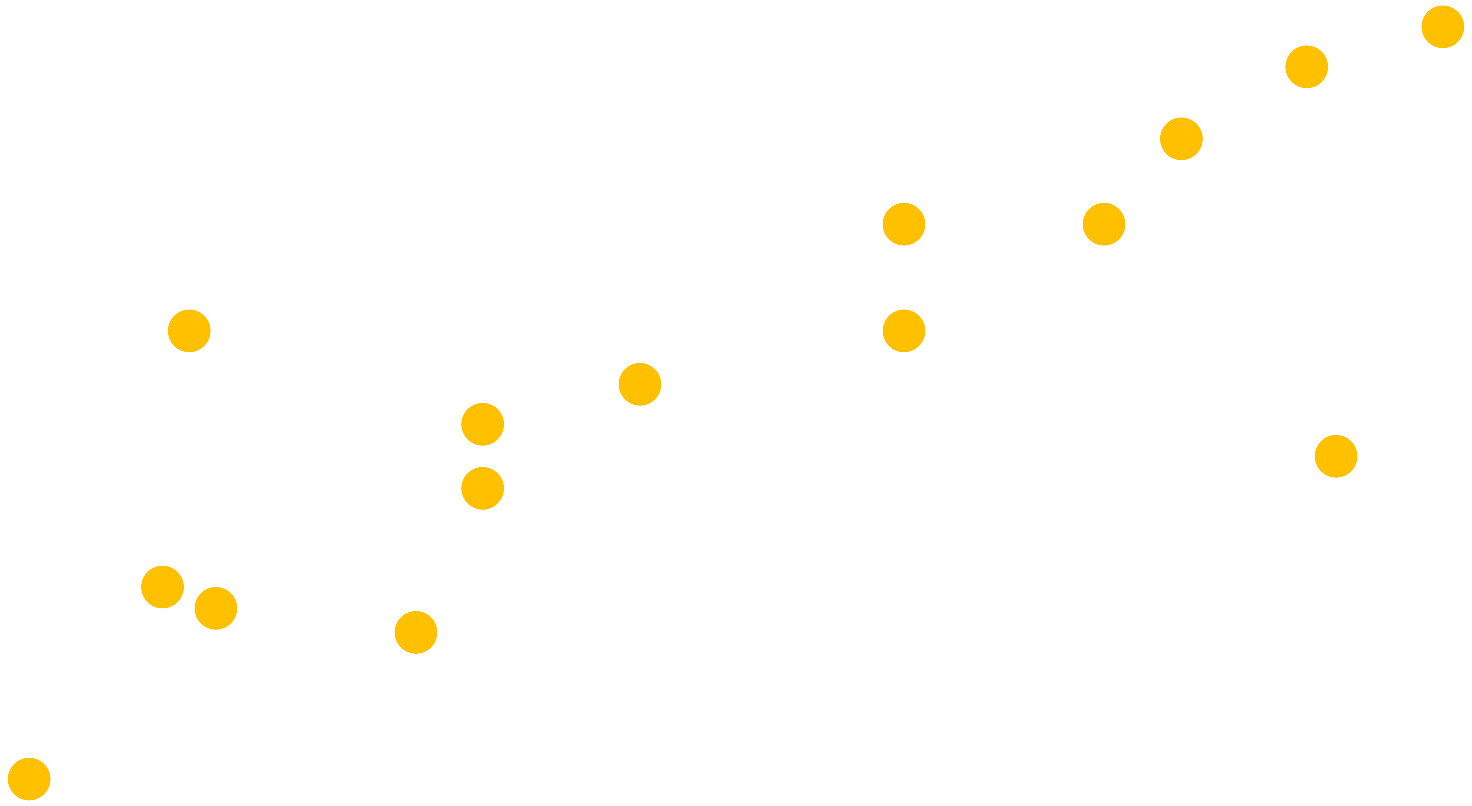
- Estimation can also divided in two steps
  - Classify data among Inliers and Outliers
  - Generate a model using inliers only
- RANSAC (RANdom SAmple Consensus) is a widely used approach
  - M. A. Fischler and R. C. Bolles (June 1981). *"Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography"*. Comm. of the ACM 24: 381--395

RANSAC

- Given a set of points
  - $(x_1, y_1), (x_2,y_2), \ldots, (x_n,y_n)$
  - Potentially including outliers…
- Estimate parameters
  - Example: $y=mx+q$ or $ax+by+c=0$

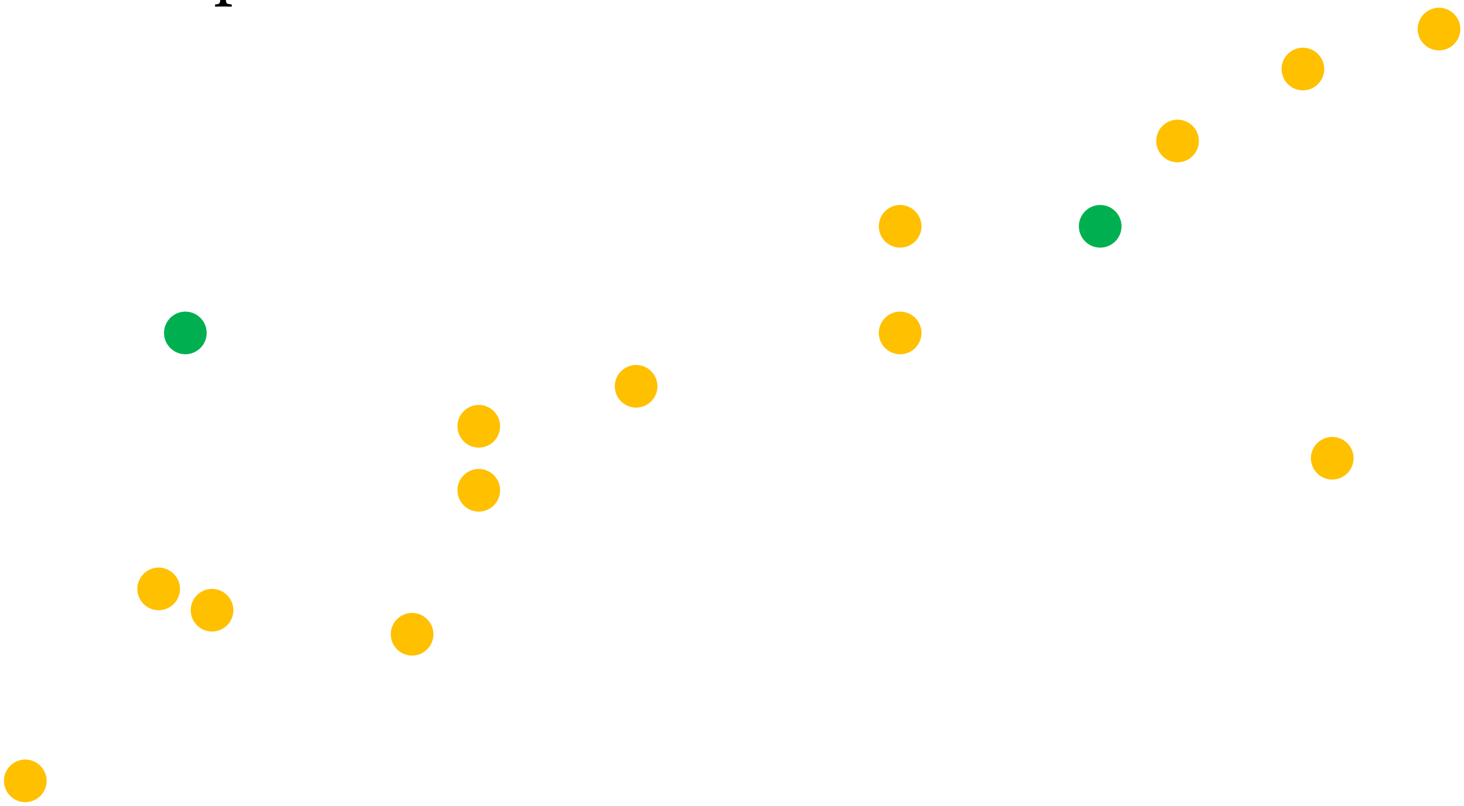
RANSAC line fitting

- We have some outliers…



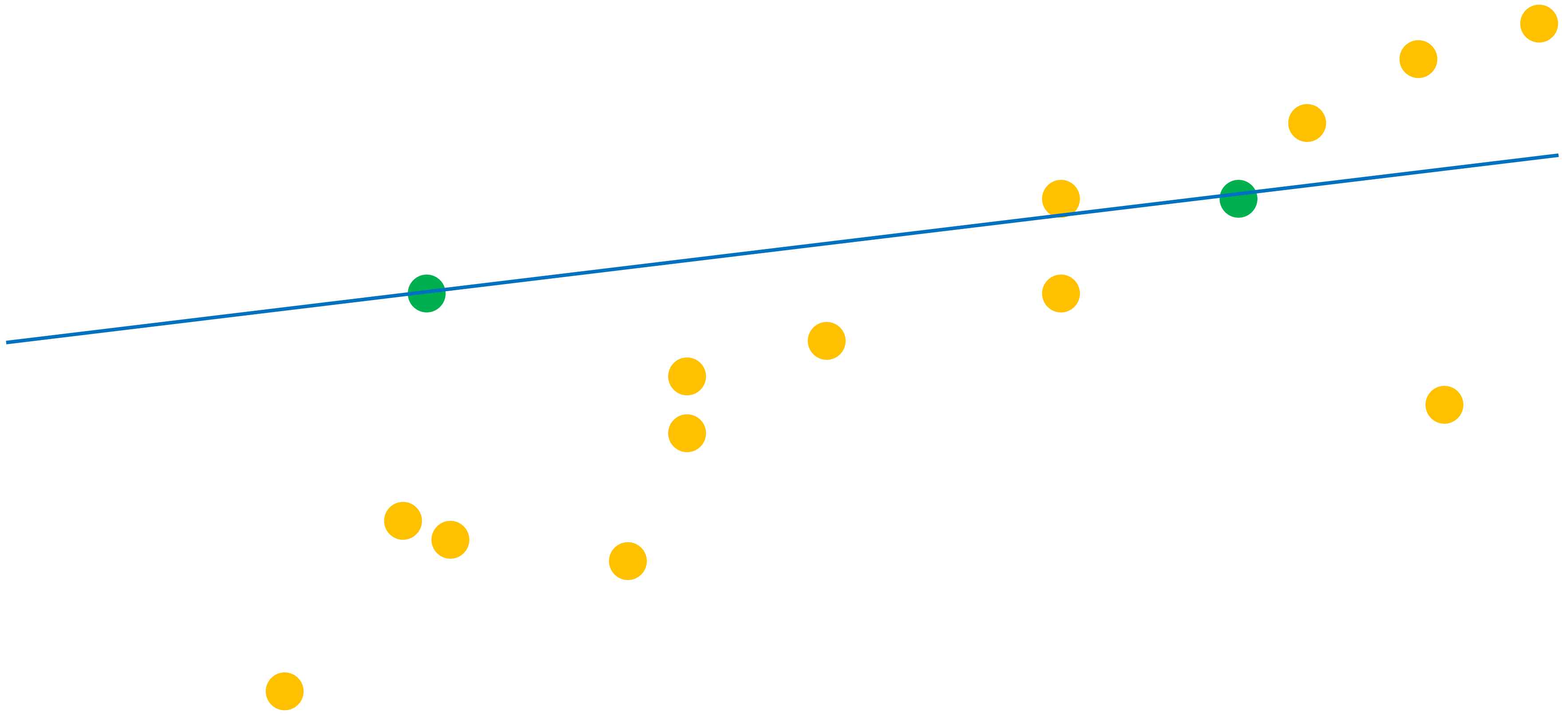
RANSAC line fitting

- Randomly choose 2 points



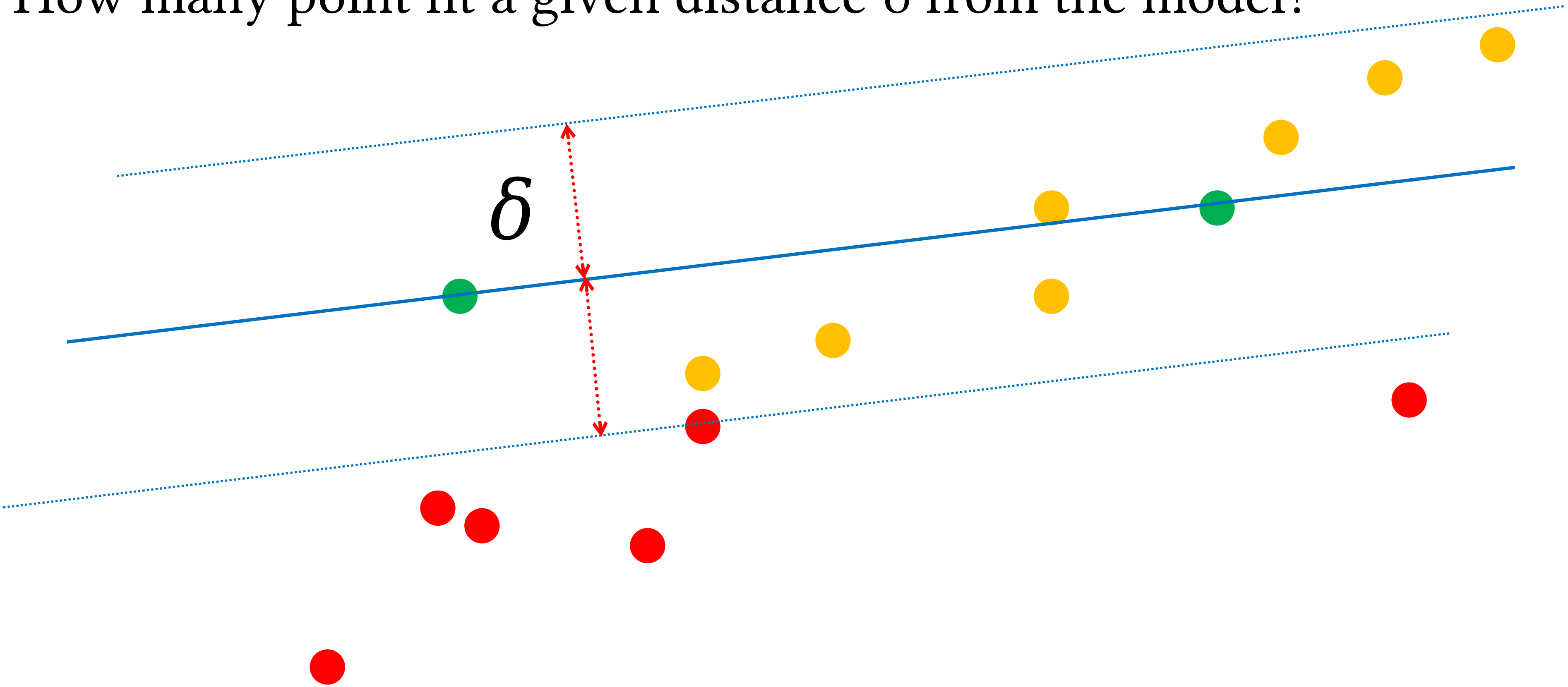
RANSAC line fitting

- Estimate model



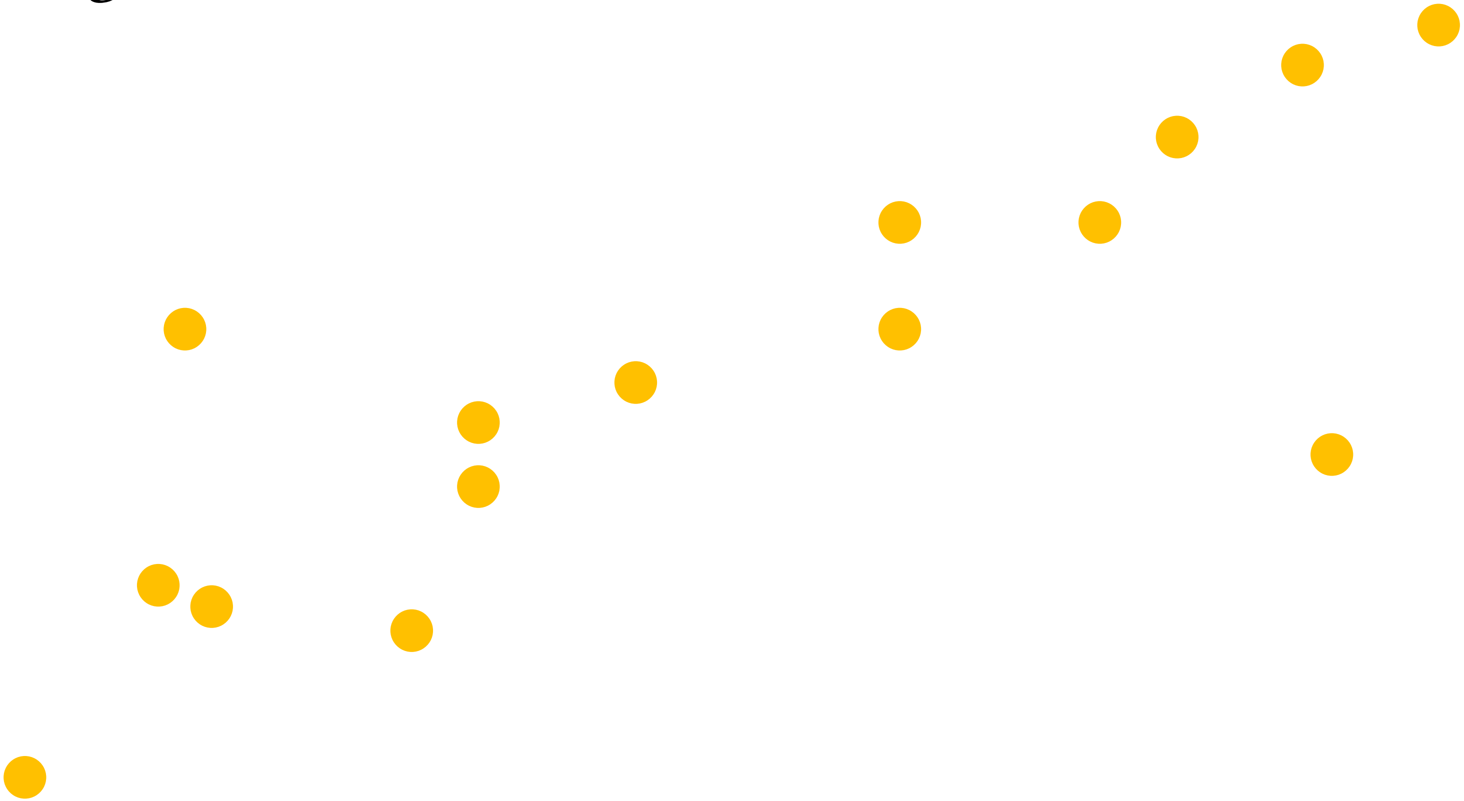
RANSAC line fitting

- How many point fit a given distance δ from the model?



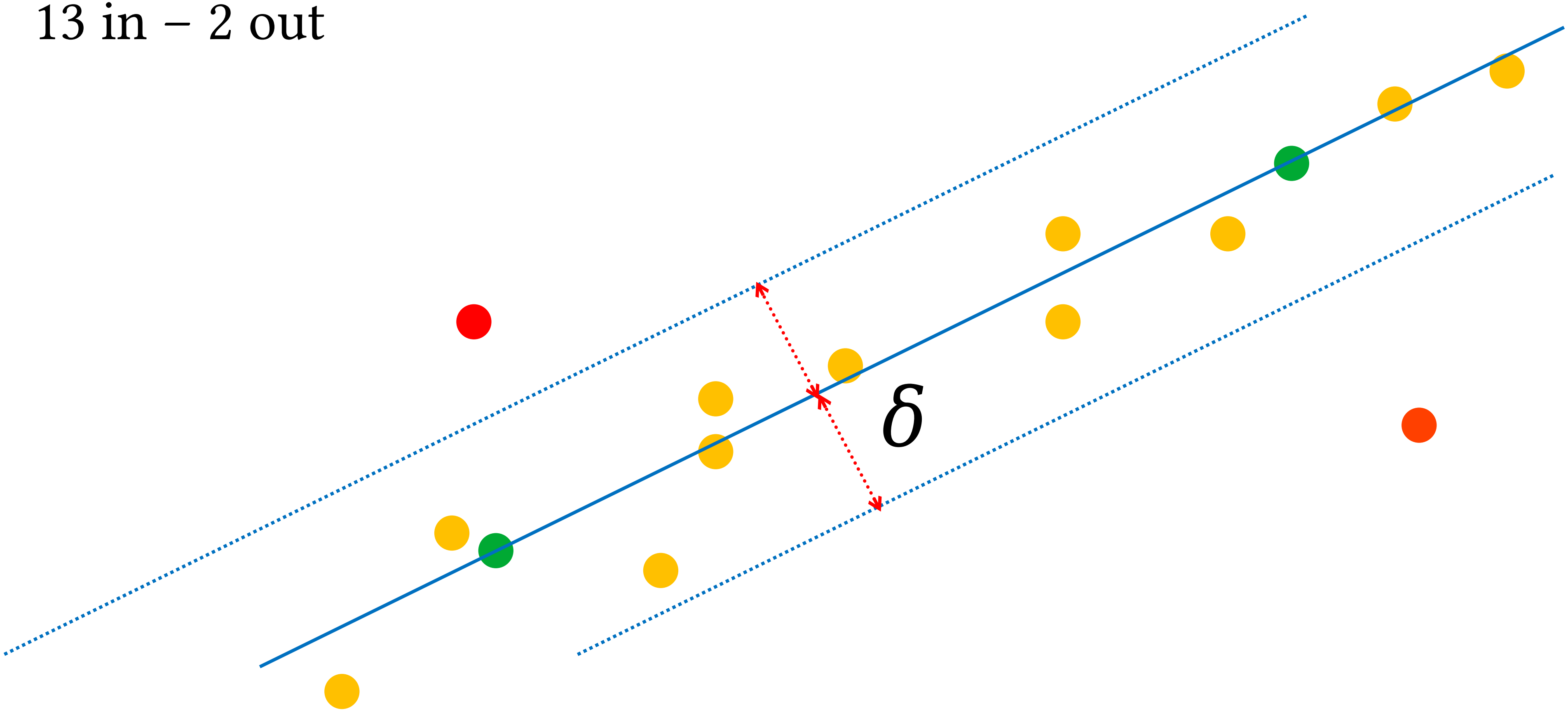
RANSAC line fitting

- Repeat until a “good” result is obtained



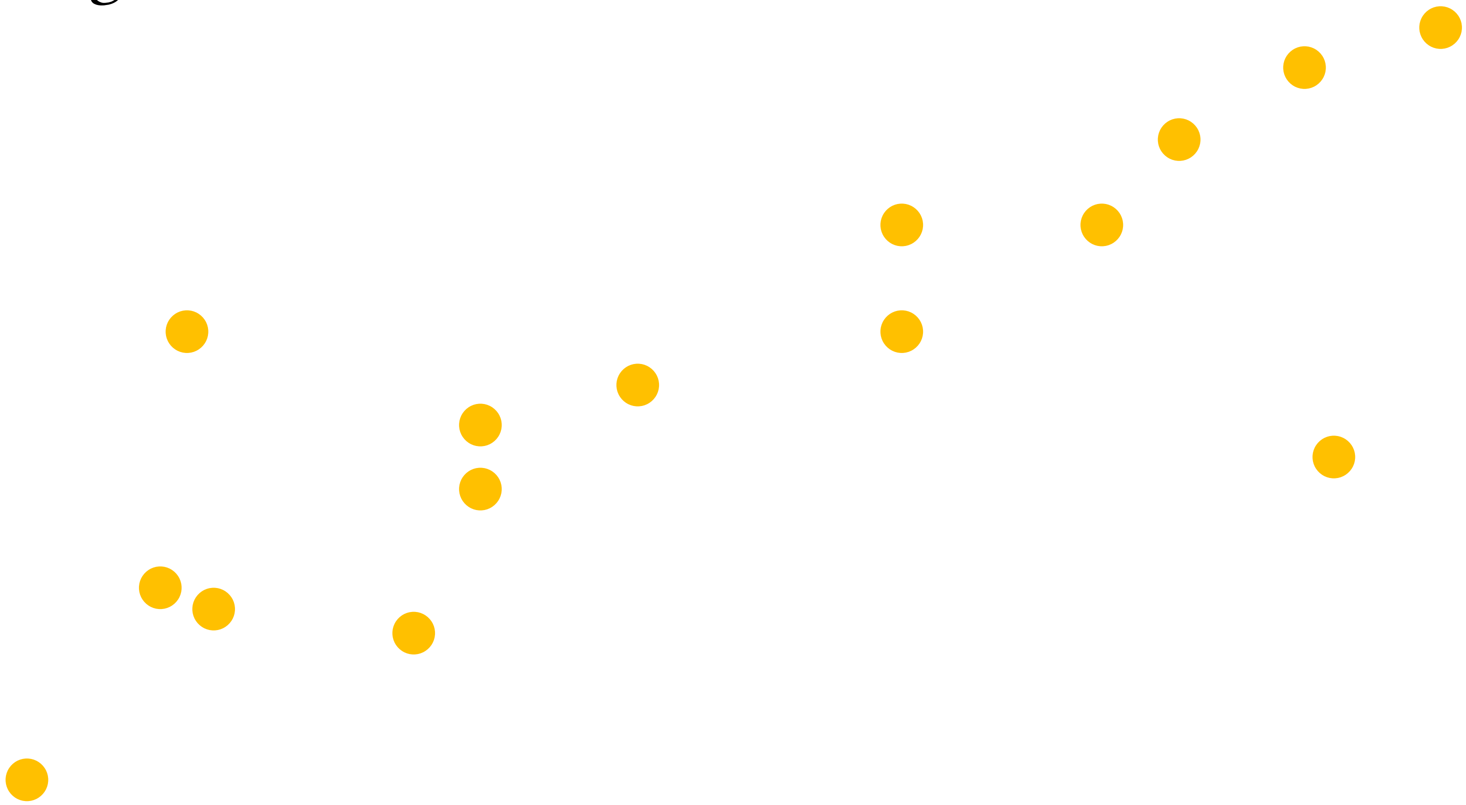
RANSAC line fitting

- 13 in – 2 out



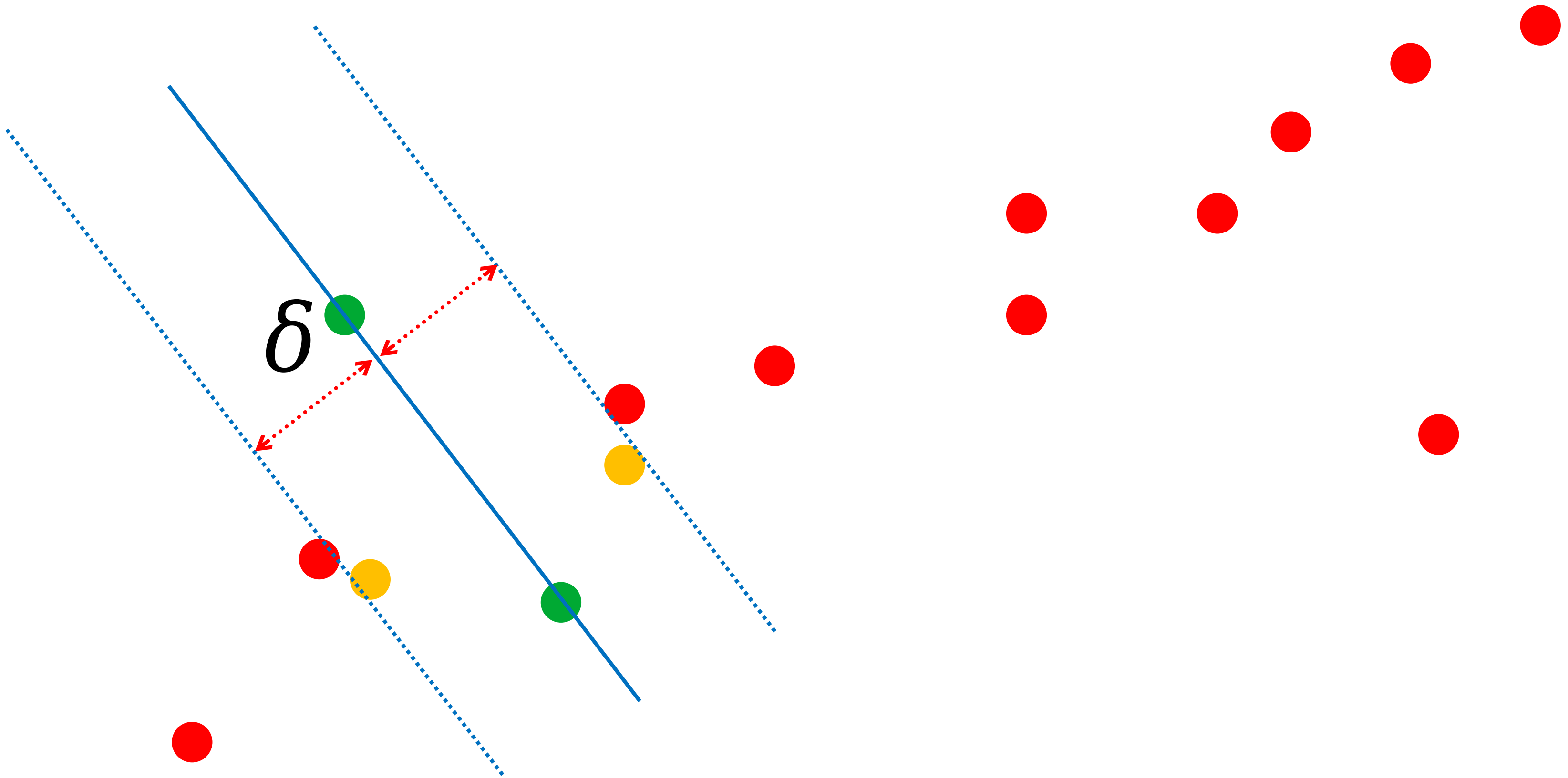
RANSAC line fitting

- Repeat until a “good” result is obtained



RANSAC line fitting

- 4 in…



How many iterations?

- Let's just consider the straight line example
- We have n=12 points
- In order to test a line we need s=2 points

$$\binom{n}{s} = \frac{n!}{(n-s)!\,s!} = 66$$

How many iterations?

- Given:
  - e $\longrightarrow$ probability a given point is an outlier
  - s $\longrightarrow$ number of points in a subset
  - N $\longrightarrow$ number of subsets (the unknown)
  - p $\longrightarrow$ probability to have at least 1 good subset

$$p = 1 - \left(1 - (1 - e)^s\right)^N$$

How many iterations?

- Given:
  - e $\longrightarrow$ probability a given point is an outlier
  - s $\longrightarrow$ number of points in a subset
  - N $\longrightarrow$ number of subsets (the unknown)
  - p $\longrightarrow$ probability to have at least 1 good subset

$$p = 1 - \left(1 - \left(\underbrace{1 - e}_{\text{Inlier Probability}}\right)^{s}\right)^{N}$$

How many iterations?

UNIVERSITÀ DI PARMA

- Given:
  - e → probability a given point is an outlier
  - s → number of points in a subset
  - N → number of subsets (the unknown)
  - p → probability to have at least 1 good subset

$$p = 1 - \left(1 - \underbrace{(1-e)^s}_{\text{All Inlier Probability in the Subset}}\right)^N$$

How many iterations?

- Given:
  - e $\longrightarrow$ probability a given point is an outlier
  - s $\longrightarrow$ number of points in a subset
  - N $\longrightarrow$ number of subsets (the unknown)
  - p $\longrightarrow$ probability to have at least 1 good subset

$$p = 1 - \left(1 - (1-e)^s\right)^N$$

One or more Outliers Probability in the Subset

How many iterations?

UNIVERSITÀ DI PARMA

- Given:
  - e $\longrightarrow$ probability a given point is an outlier
  - s $\longrightarrow$ number of points in a subset
  - N $\longrightarrow$ number of subsets (the unknown)
  - p $\longrightarrow$ probability to have at least 1 good subset

$$p = 1 - \left(1 - (1 - e)^s\right)^N$$

Probability that N subsets always contain outliers

How many iterations?

- Given:
  - e → probability a given point is an outlier
  - s → number of points in a subset
  - N → number of subsets (the unknown)
  - p → probability to have at least 1 good subset

$$p = 1 - \left(1 - (1 - e)^s\right)^N$$

Probability to have at least one uncontaminated subset

How many iterations?

UNIVERSITÀ DI PARMA

- N can be then computed as f(p,e)
  - s is known given the model

$$\left(1-\left(1-e\right)^{s}\right)^{N}=p-1 \quad \rightarrow \quad N=\frac{\log\left(1-p\right)}{\log\left(1-\left(1-e\right)^{s}\right)}$$

How many iterations?

- Assuming we need a "strong" probability (i.e. p=0.99)

$$N = \frac{\log(1-p)}{\log(1-(1-e)^s)}$$

	proportion of outliers e						
s	5%	10%	20%	25%	30%	40%	50%
2	2	3	5	6	7	11	17
3	3	4	7	9	11	19	35
4	3	5	9	13	17	34	72
5	4	6	12	17	26	57	146
6	4	7	16	24	37	97	293
7	4	8	20	33	54	163	588
8	5	9	26	44	78	272	1177

How many iterations (N)?

- Again let's consider the straight line example
- We have n=12 points
- We also estimate a ~20% outliers
- In order to test a line we need s=2 points
- We want a p=0.99

$$N = 5 \ll 66$$

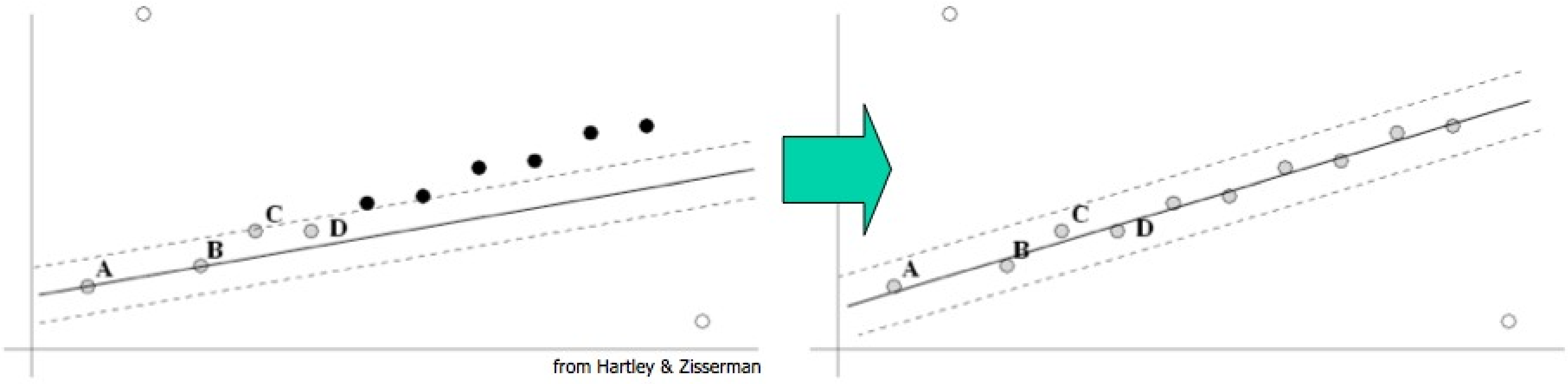
How many iterations?

- Therefore it is not necessary to test all subsets
- We can simply randomly choose N
  - N is usually << all possible combinations
- We can further reduce N
  - Stop iterations when a “sufficient” number of inliers is found

$$T = (1 - e) \cdot n$$

RANSAC: improvements

- Once a model is tested it can be refined considering all inliers
  - LS can be used again



RANSAC: pro&cons

- It often works!
- We need data assumptions
  - inliers/outliers probability
- When several outliers it may be slow...
  - No upper time
  - You can limit iterations $\longrightarrow$ far from optimal solution

UNIVERSITÀ DI PARMA

Model Fitting

Question time!