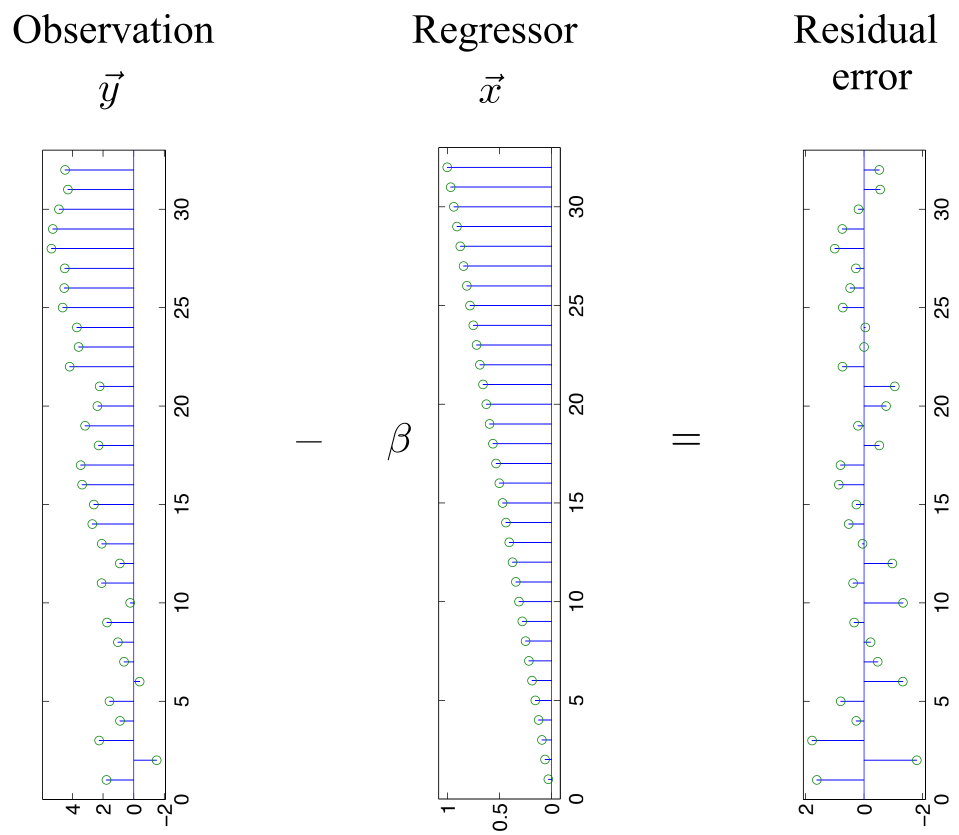
Mathematical Tools
for Neural and Cognitive Science

Fall semester, 2016

# Section 2: Least Squares

## Least Squares (outline)

- Standard regression: Fit data with weighted sum of regressors. Solution via calculus, orthogonality, SVD
- Choosing regressors, overfitting
- Robustness: weighted regression, iterative outlier trimming, robust error functions, iterative re-weighting
- Constrained regression: linear, quadratic
- Total Least Squares (TLS) regression, and Principal Components Analysis (PCA)

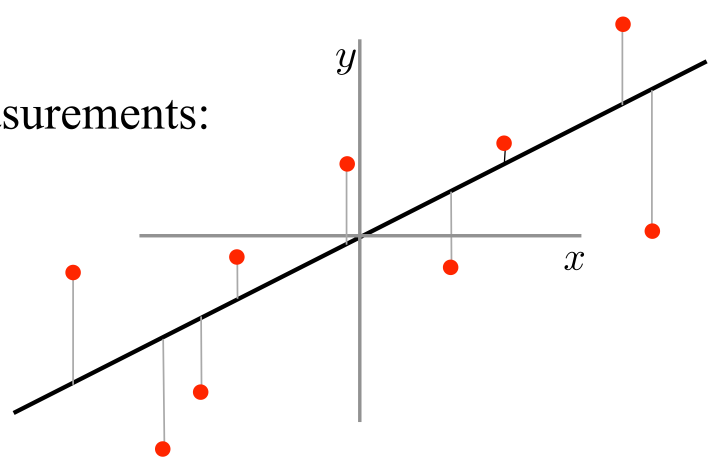


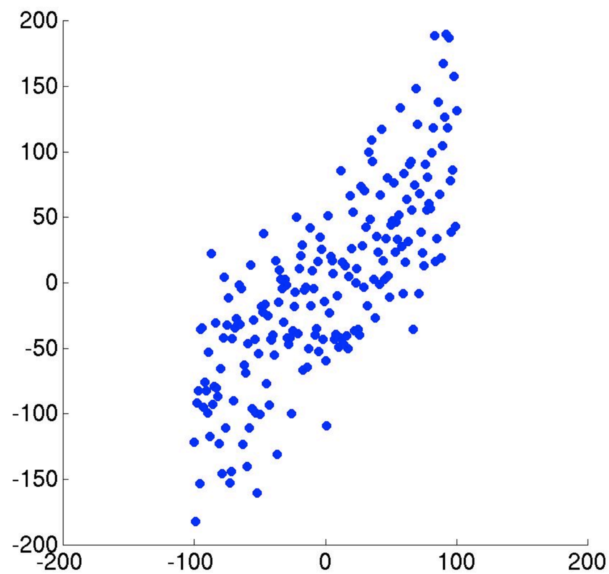

# Least squares regression:

"objective function"

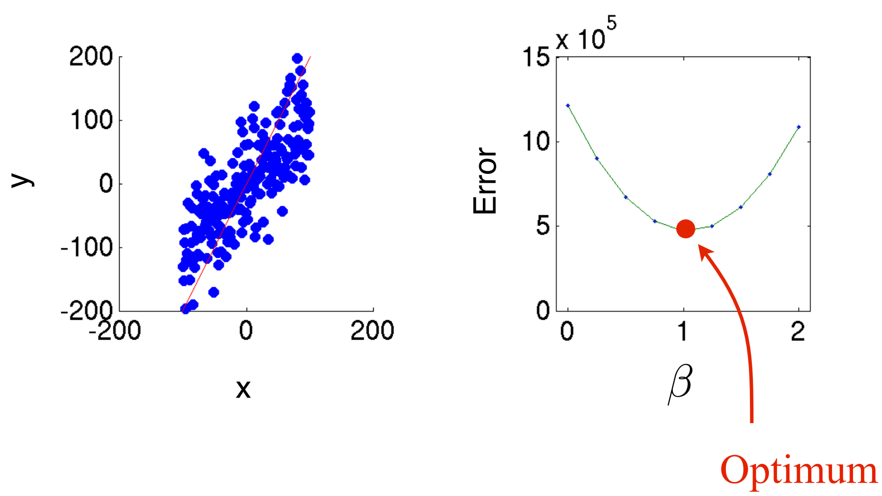
$$\min_{\beta} \sum_{n} (y_n - \beta x_n)^2$$

In the space of measurements:





200
150
100
50
0
-50
-100
-150
-200
-200
-100
0
100
200

y
x
Error
x 10^5
β
Optimum

$$\min_{\beta} \sum_{n} (y_n - \beta x_n)^2$$

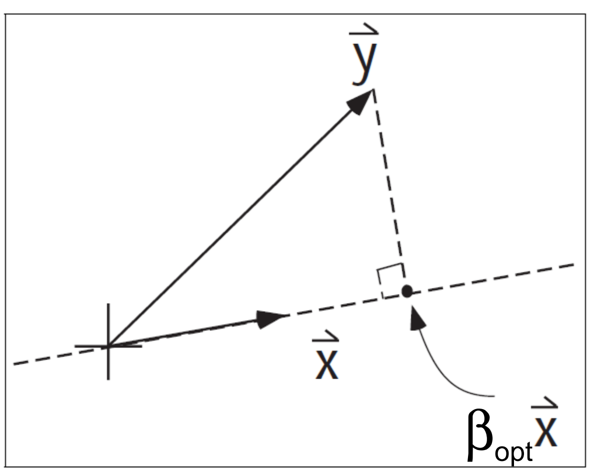
can solve this with calculus...

*[on board]*

... or linear algebra:

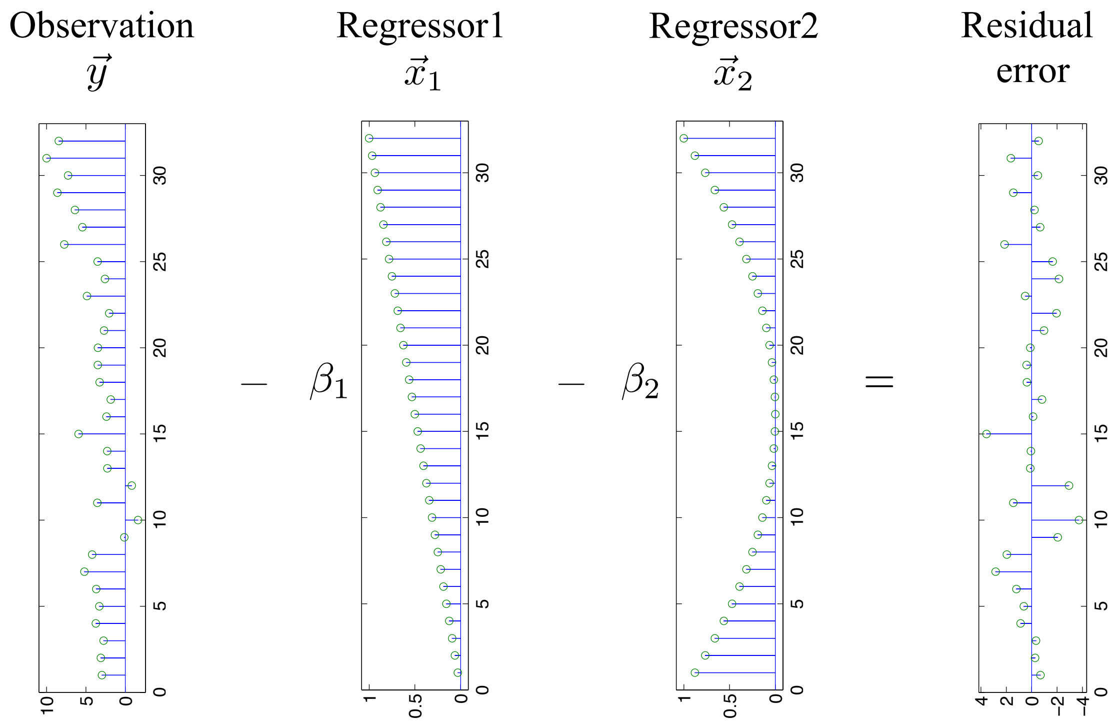
$$\min_{\beta} \sum_{n} (y_n - \beta x_n)^2 \quad = \min_{\beta} ||\vec{y} - \beta \vec{x}||^2$$

Geometry:
(note: this is *not* the measurement space of previous plots)



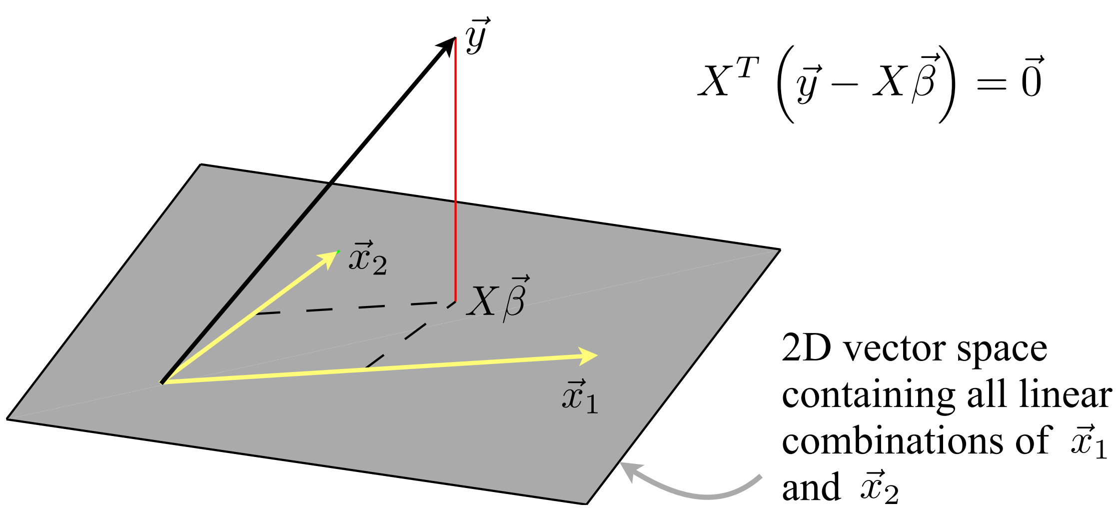
Multiple regression

$$\min_{\vec{\beta}} ||\vec{y} - \sum_k \beta_k \vec{x}_k||^2 = \min_{\vec{\beta}} ||\vec{y} - X\vec{\beta}||^2$$




# Solution via the Orthogonality Principle

$$X^T\left(\vec{y} - X\vec{\beta}\right) = \vec{0}$$




Alternatively, use SVD...

$$\min_{\vec\beta} ||\vec y - X\vec\beta||^2 = \min_{\vec\beta} ||\vec y - USV^T\vec\beta||^2$$
$$= \min_{\vec\beta} ||U^T\vec y - SV^T\vec\beta||^2$$
$$= \min_{\vec\beta^*} ||\vec y^* - S\vec\beta^*||^2$$
$$\text{where} \quad \vec y^* = U^T\vec y, \quad \vec\beta^* = V^T\vec\beta$$

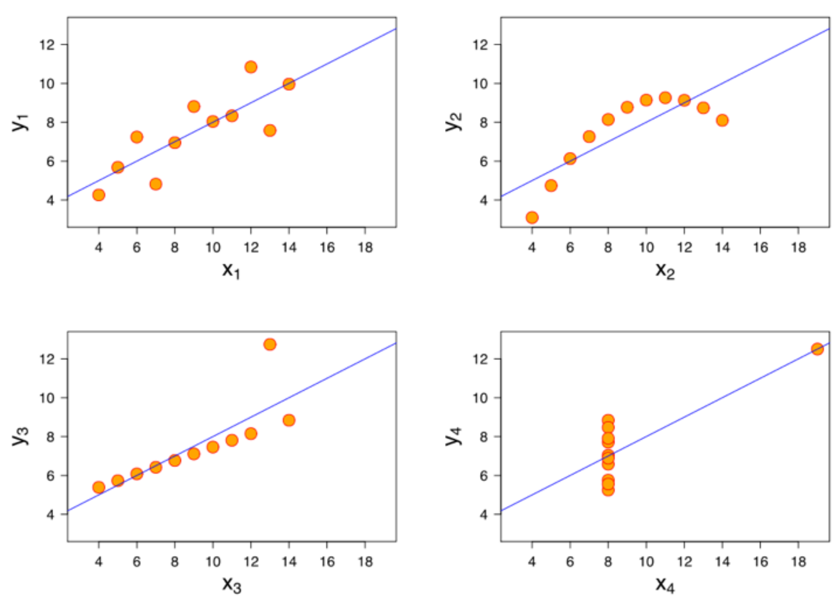
Solution: $\beta^*_{\mathrm{opt},k} = y^*_k / s_k, \quad$ for each $k$

or $\quad \vec\beta^*_{\mathrm{opt}} = S^{\#}\vec y^*$

*[on board: transformations, elliptical geometry]*

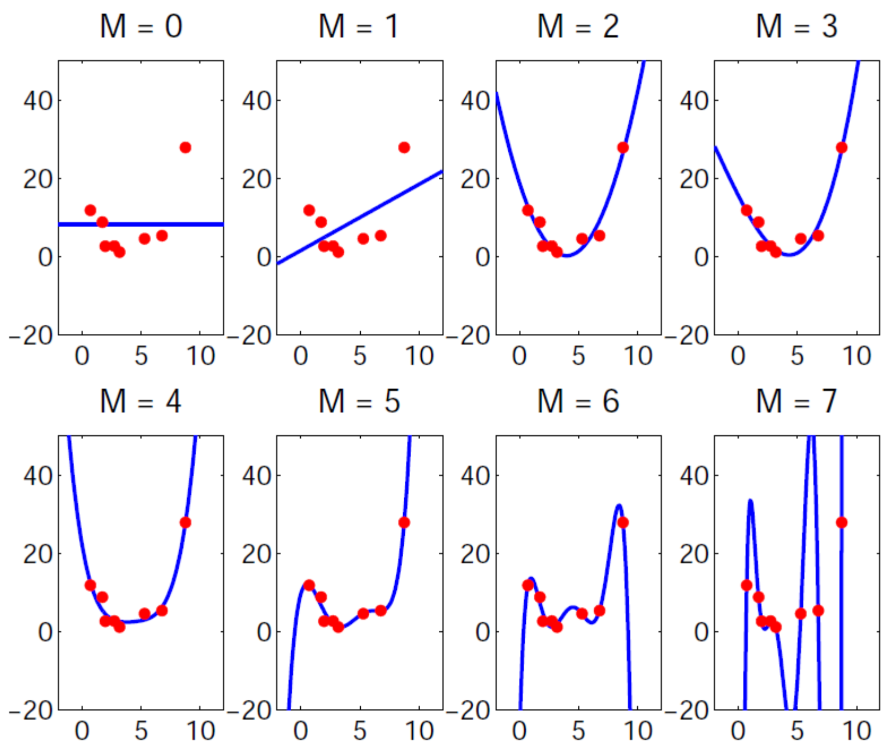
## Interpretation: what does it mean?

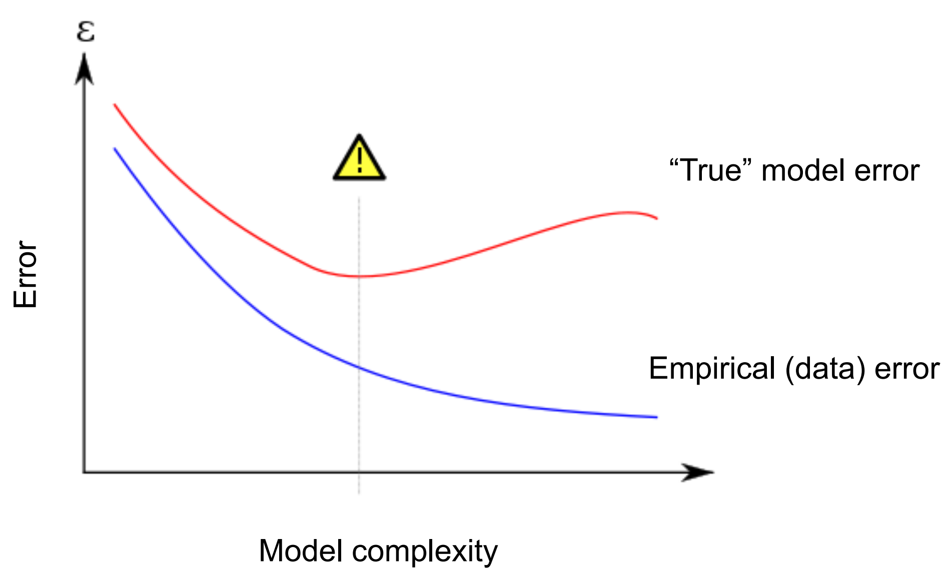
Note that these all give the same regression fit:




[Anscombe, 1973]

# Polynomial regression - how many terms?




(to be continued, when we get to "statistics"...)

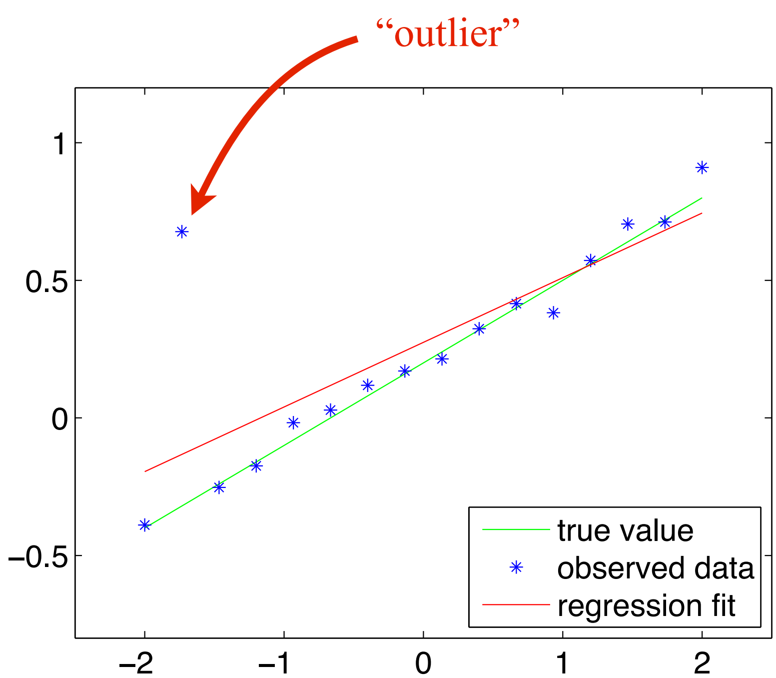
# Weighted Least Squares

$$\min_{\beta} \sum_{n} [w_n(y_n - \beta x_n)]^2$$

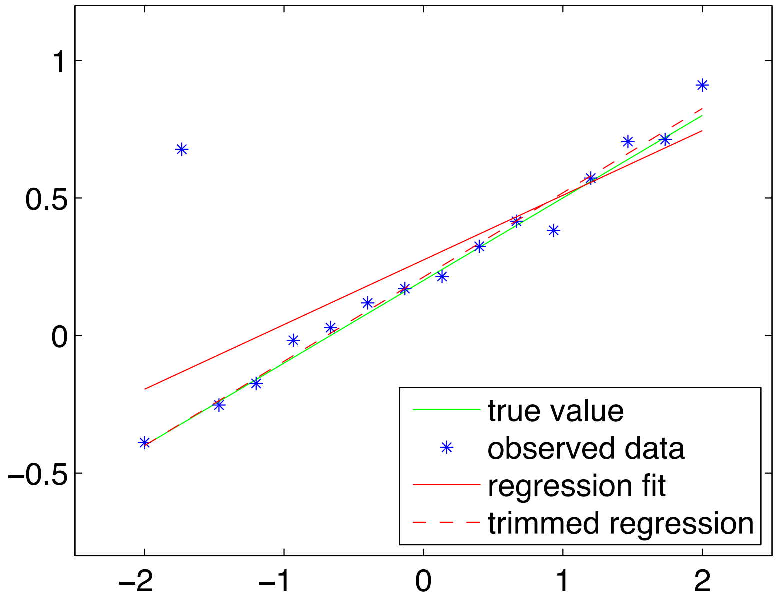
$$= \min_{\beta} ||W(\vec{y} - \beta\vec{x})||^2$$

diagonal matrix

Solution via simple extensions of basic regression solution

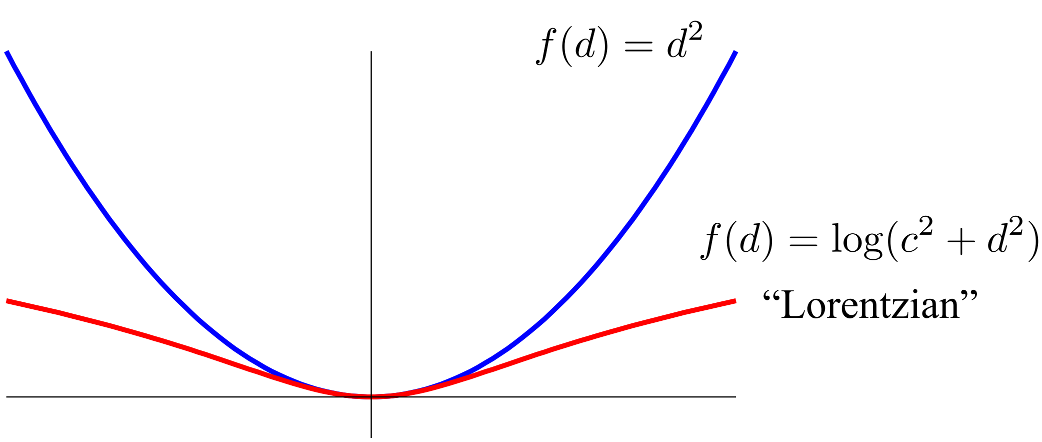



Solution 1: "trimming"...




When done iteratively (discard the outlier, re-fit, repeat), this is a so-called "greedy" method. When do you stop?

Solution 2: Use a "robust" error metric.
For example:




Note: generally can't obtain solution directly (i.e., requires an iterative optimization procedure).

In some cases, can use iteratively re-weighted least squares (IRLS)

# Constrained Least Squares

Linear constraint:

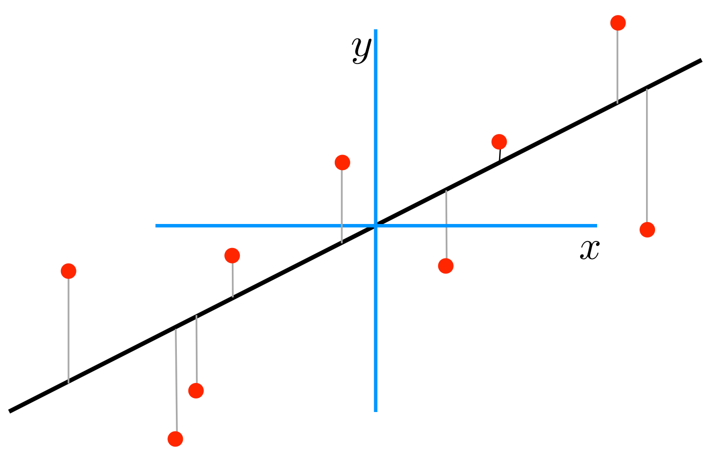
$$\min_{\vec{\beta}} ||\vec{y} - X\vec{\beta}||^2, \quad \text{where } \vec{c} \cdot \vec{\beta} = \alpha$$

Quadratic constraint:

$$\min_{\vec{\beta}} ||\vec{y} - X\vec{\beta}||^2, \quad \text{where } ||C\vec{\beta}||^2 = \alpha$$

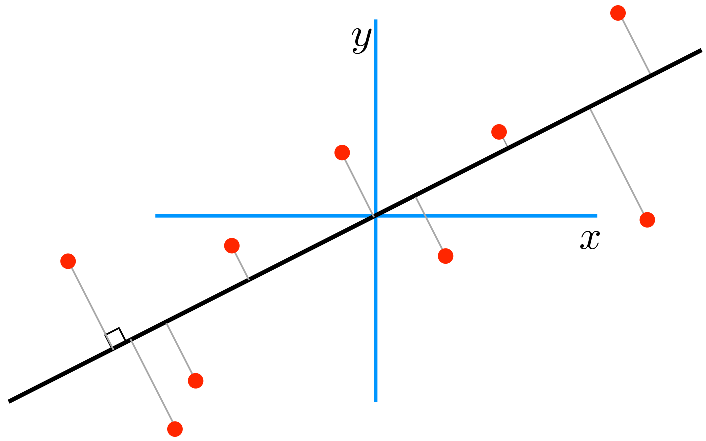
Both can be solved exactly using linear algebra... *[on board]*

Standard regression error: squared error of the "dependent" variable




$$\min_{\beta} ||\vec{y} - \beta\vec{x}||^2$$

Total Least Squares regression: Error is squared distance from the fitted line...




$$\min_{\hat{u}} ||D\hat{u}||^2, \quad \text{where } ||\hat{u}||^2 = 1$$

Solution via SVD

Elliptical geometry (again)

Partition of total variance into orthogonal components

Symmetric matrix $M^T M$, eigenvectors/eigenvalues

Principal Components Analysis (PCA)

*[on board: SVD solution, elliptical geometry, partition of variance]*

# Eigenvectors/eigenvalues

Define symmetric matrix:

$$\begin{aligned} C &= M^T M \\ &= (USV^T)^T (USV^T) \\ &= VS^T U^T U S V^T \\ &= V(S^T S)V^T \end{aligned}$$

If $\vec{v}_k$ is the $k$th column of $V$ then:

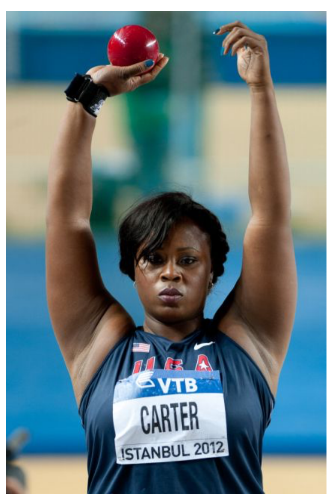
$$\begin{aligned} C\vec{v}_k &= V(S^T S)V^T \vec{v}_k \\ &= V(S^T S)\hat{e}_k \\ &= s_k^2 V \hat{e}_k \\ &= s_k^2 \vec{v}_k \end{aligned}$$

And, for arbitrary vectors $\vec{x}$ :

$$C\vec{x} = \sum_k s_k^2 (\vec{v}_k^T \vec{x}) \vec{v}_k$$

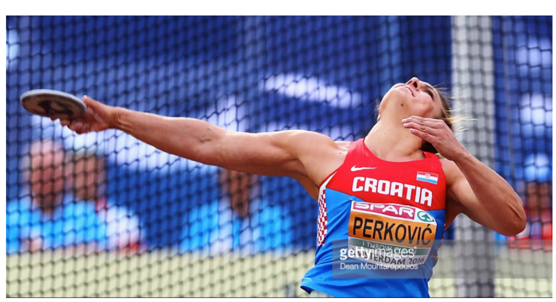
3D geometry:
Javelin, Discus, Shotput

Olympic gold medalists
(Rio, 2016)




Michelle Carter (USA)


Thomas Röhler (Germany)




Sandra Perković (Croatia)