



Lecture 28

Least Squares

Slides created by John DeNero (denero@berkeley.edu) and Ani Adhikari (adhikari@berkeley.edu)
Contributions by Fahad Kamran (fhdkmrn@berkeley.edu) and Vinitra Swamy (vinitra@berkeley.edu)

Announcements

Linear Regression (Review)

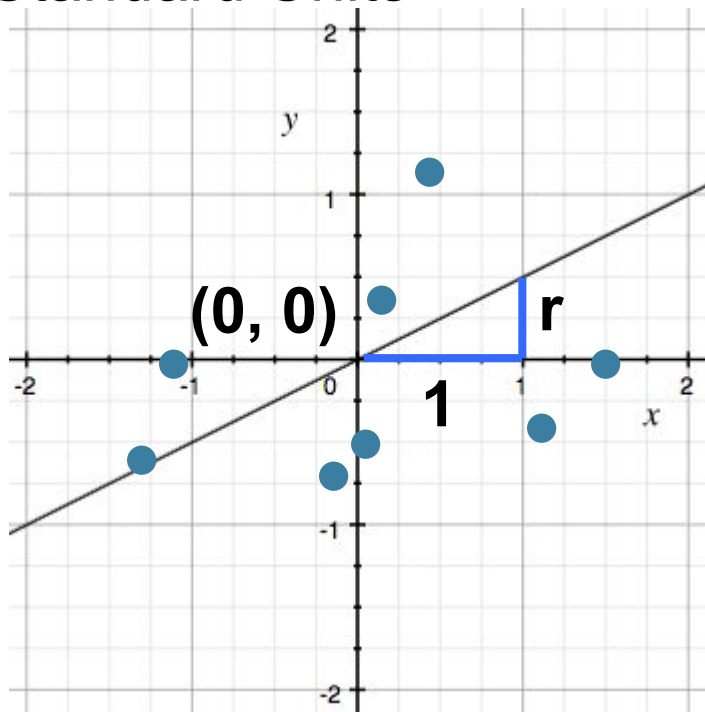
Regression Estimate

To find the regression estimate of y :

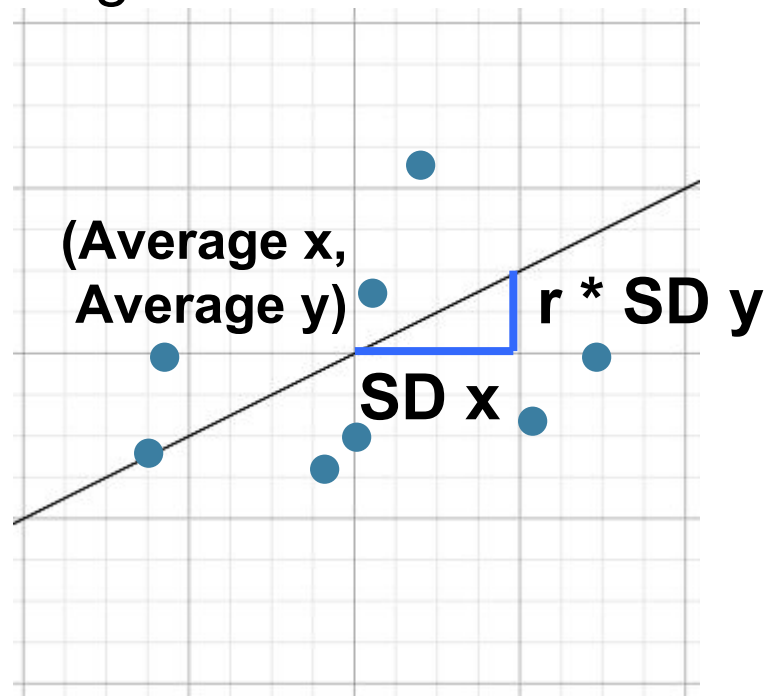
- Convert the given x to standard units
 - Multiply by r
 - That's the regression estimate of y , but:
 - It's in standard units
 - So convert it back to the original units of y
-

Regression Line

Standard Units



Original Units



Slope and Intercept

estimate of $y = \text{slope} * x + \text{intercept}$

$$\text{slope of the regression line} = r \cdot \frac{\text{SD of } y}{\text{SD of } x}$$

$$\text{intercept of the regression line} = \text{average of } y - \text{slope} \cdot \text{average of } x$$

Discussion Question

A course has a midterm (average 70; standard deviation 10) and a really hard final (average 50; standard deviation 12)

If the scatter diagram comparing midterm & final scores for students has a typical oval shape with correlation 0.75, then...

What do you expect the average final score would be for students who scored 90 on the midterm?

How about 60 on the midterm?

(Demo)

Least Squares

Error in Estimation

- **error = actual value – estimate**
- Typically, some errors are positive and some negative
- To measure the rough size of the errors
 - **square** the **errors** to eliminate cancellation
 - take the **mean** of the squared errors
 - take the square **root** to fix the units
 - **root mean square error** (rmse)

(Demo)

Least Squares Line

- Minimizes the root mean squared error (rmse) among all **lines**
- Equivalently, minimizes the mean squared error (mse) among all lines
- Names:
 - “Best fit” line
 - Least squares line
 - Regression line

(Demo)

Numerical Optimization

- Numerical minimization is approximate but effective
- Lots of machine learning uses numerical minimization
- If the function `mse(a, b)` returns the mse of estimation using the line “estimate = ax + b”,
 - then `minimize(mse)` returns array `[a0, b0]`
 - `a0` is the slope and `b0` the intercept of the line that minimizes the mse among lines with arbitrary slope `a` and arbitrary intercept `b` (that is, among all lines)

(Demo)
