FoDA : Linear Regression

L11 : explanatory & dependent variables

Data

labeled data

(X, y) actual observations

$X \in \mathbb{R}^{n \times d} \qquad y \in \mathbb{R}^n$

$$(X, y) = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$$

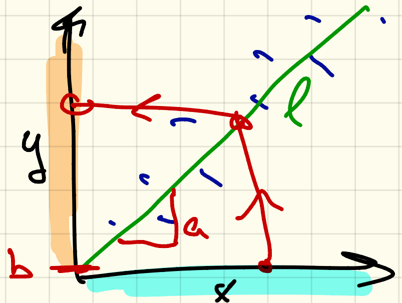
X = explanatory variable

y = dependent variable

fit line

slope, intercept

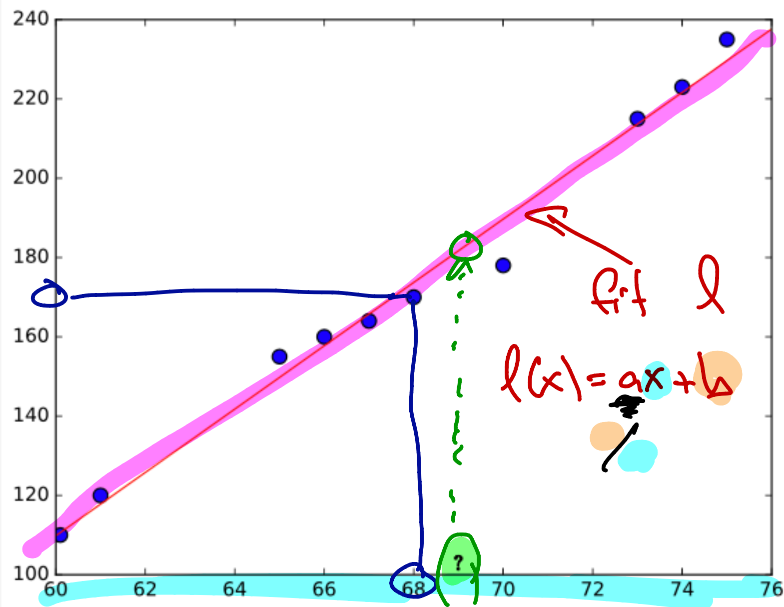
$$\hat{y} = \ell_{a,b}(x) = ax + b$$



- Measure Error based on Prediction
  (want to use on data we don't have yet!)

- explanatory x dependent y

height (in)	weight (lbs)
66	160
68	170
60	110
70	178
65	155
61	120
74	223
73	215
75	235
67	164
69	?



How to Measure Error?

residual

$\hat{y} = \ell(x)$ \hat{y_i}

ith data point $r_i = y_i - \hat{y}_i = y_i - \ell(x_i)$

true label

predicted label

model

$M_\alpha(\ )$

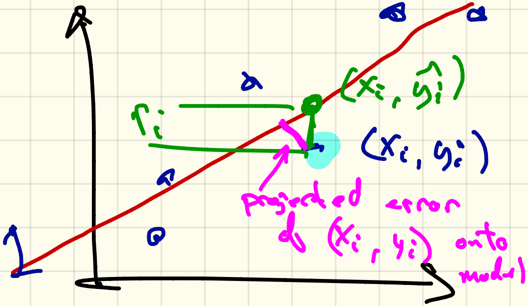
Sum of Squared Errors

$$SSE((X,y),\ell) = \sum_{(x_i,y_i)\in(X,y)} (r_i)^2$$

$y_i = 3$

$\hat{y}_i = 4$

$r_i = 3-4 = -1$



Sum of Squared Errors

$$SSE((X, y), \ell) = \sum_{i=1}^{n} (r_i)^2 = \sum_{i=1}^{n} (y_i - \underbrace{\ell(x_i)}_{\hat{y}_i})^2$$

(X, y): data; ℓ: model; n: # data points; y_i: true label; $\ell(x_i)$: prediction label

why?

- squaring $\rightarrow$ makes non-negative
  why not absvalue $|r_i|$
- Norm (L_2-norm) $\|r\|_2^2$ $\quad r = (r_1, r_2, \ldots r_n)$

common notation

value v_i

predicted value $\hat{v}_i$

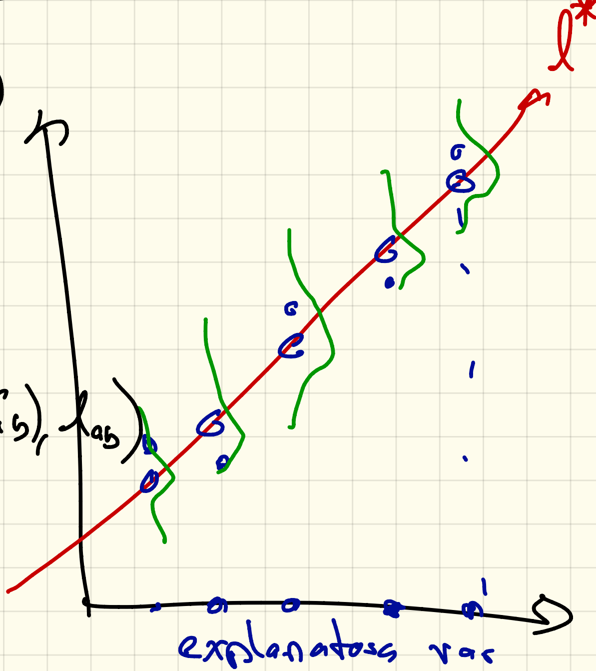
- Start w/ Bayesian Inference

  Assume Normal Noise on y_i

  $$N(l^*(x_i), \sigma^2)$$

  ↳ Negative Log-Likelihood

  ↳ $SSE((X, y), l)$



- Easy to Solve

  for optimal

  $$(a^*, b^*) = \underset{a,b}{\arg\min}\ SSE((X, y), l_{ab})$$

  → "closed form"

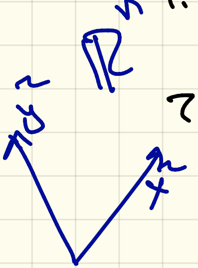
  → gradient descent

  ↳ convex

How to Solve for

$$a^*, b^* = \underset{a,b\in\mathbb{R}}{\arg\min}\ SSE((x,y), l_{a,b})$$

$$l_{a,b}(x) = ax + b$$



1. average $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ $\quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$
2. set "center" $\tilde{x} = (x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x})$

   $\tilde{y} = (y_1 - \bar{y}, y_2 - \bar{y}, \dots, y_n - \bar{y})$
3. $a^* = \dfrac{\langle \tilde{y}, \tilde{x} \rangle}{\|\tilde{x}\|^2} = \dfrac{\|\tilde{y}\|}{\|\tilde{x}\|}\cos\theta$

$b^* = \bar{y} - a\bar{x}$