Weighted Least Squares

In the MLR model,

$$Y = X\beta + \varepsilon,$$

if $var(\varepsilon) \neq \sigma^2 I$, we could instead assume

$$var(\varepsilon) = \Sigma = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n^2 \end{pmatrix}.$$

Then we can try to transform our data to control heteroskedasticity. The idea is to use the weights inversely proportional to the variance, $w_i = 1/\sigma_i^2$.

In matrix form, define $W = diag(w_1, \dots, w_n) = diag(1/\sigma_1^2, \dots, 1/\sigma_n^2)$.

Then transform:

$$\tilde{Y} = \underline{W^{1/2} Y}$$

$$\tilde{X} = \underline{W^{1/2} X} \quad \& \text{ minimize } \|\tilde{Y} - \tilde{X}\beta\|^2.$$

$$\tilde{\varepsilon} = W^{1/2}\varepsilon$$

The WLS sol is:

$$\hat{\beta} = (\tilde{X}^T\tilde{X})^{-1}\tilde{X}^T\tilde{Y}$$

$$= ((W^{1/2}X)^T(W^{1/2}X))^{-1}(W^{1/2}X)^T(W^{1/2}Y)$$

$$= (X^T W^{1/2} W^{1/2} X)^{-1} X^T W^{1/2} W^{1/2} Y$$

$$= (X^T \underline{W} X)^{-1} X^T \underline{W} Y$$

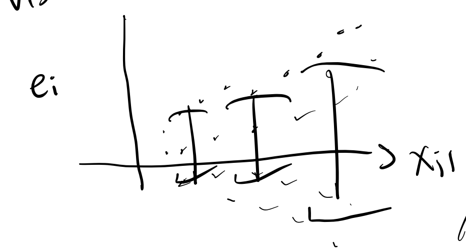
What is w_i?

Idea: $w_i = \frac{1}{\sigma_i^2}$

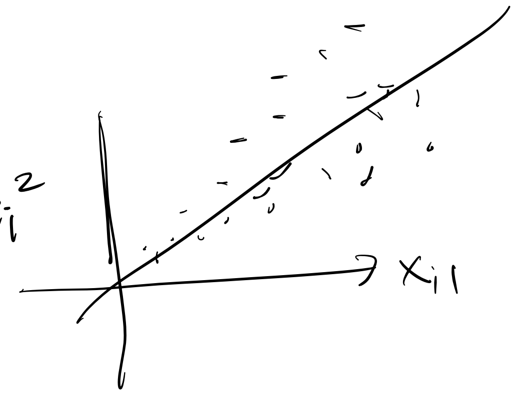
If a datapoint is noisier, weight it less.

I need to estimate σ_i^2

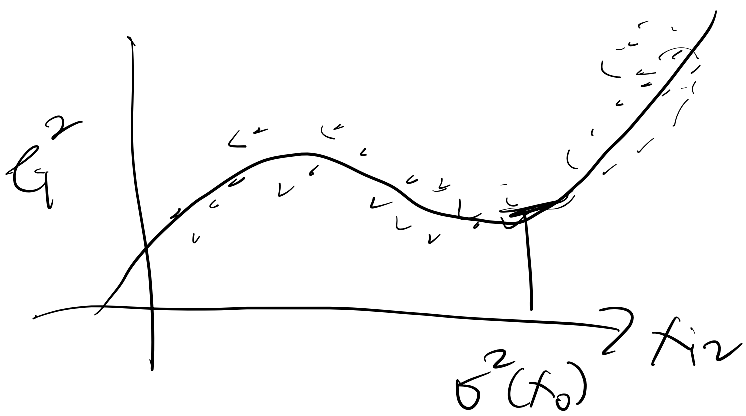
visually:



$S_i = e_i^2$



$\hat{S}_i = \hat{\alpha}_0 + \hat{\alpha}_1 x_{i1}$



ASIDE

Dependence in (X_i, Y_i)

X_{11}	Y_{11}
X_{12}	Y_{12}
X_{13}	Y_{13}
X_{21}	Y_{21}
X_{22}	Y_{22}
X_{23}	Y_{23}

block dependence

$$\mathrm{Var}(Y) = \mathrm{Var}(\varepsilon) = \begin{bmatrix} \Sigma_1 & & & \\ & \Sigma_2 & & \\ & & \ddots & \\ & & & \Sigma_n \end{bmatrix} = W$$

repeated measures/longitudinal data

Multicollinearity:

Consider: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$

but $X_2 = 2.54 X_1$.

"True" Model:

$$Y = 10 + \underline{2} X_1 + \varepsilon$$

$$X_1 = \frac{X_2}{2.54}$$

$$X_2 = 2.54\, X_1$$

Also "True" Model:

$$Y = 10 + 3X_1 - X_1 + \varepsilon$$

$$= 10 + \underline{3}X_1 - \underline{\frac{1}{2.54}} X_2 + \varepsilon$$

$\beta_0 = 10$
$\beta_1 = 2$
$\beta_2 = 0$

equivalent

vs.

$\beta_0 = 10$
$\beta_1 = 3$
$\beta_2 = -1/2.54$.

non-unique βs b/c of degenerate X!

Design Matrix:

$$X = [\underline{\mathbb{1}_n} \;\; X_1 \;\; X_2]$$

rank$(X) = 2$

$$\underset{3\times 3}{\underline{X^T X}}$$

$\text{rank}(X^T X) = \underline{2}$

$(X^T X)^{-1}$ does not exist... :(

$$\boxed{\hat{\beta} = (X^T X)^{-1} X^T Y}$$

not unique

$$Var(\hat{\beta}) = \underline{\sigma^2 (X^T X)^{-1}}$$

$$X^T X = V D V^T$$

V is eigenvectors

D is eigenvalues

$$D = \begin{bmatrix} d_1 & & \\ & d_2 & \\ & & \underset{\approx}{0} \end{bmatrix}$$

b/c rank(X) = 2.

$$Var(\hat{\beta}) = (VDV^T)^{-1} = (V^T)^{-1} D^{-1} V^{-1}$$

$$= (V^{-1})^T D^{-1} V^{-1}$$

$$= (V^{-1})^T \begin{bmatrix} 1/d_1 & & \\ & 1/d_2 & \\ & & 1/0 \end{bmatrix} V^{-1}$$

(!!!)

$\Rightarrow$ $\hat{\beta}$ is super unstable