Weekly excercise, 38.

Show that:

$\mathbb{E}[(y-\tilde{y})^2 = \text{Bias}[\tilde{y}] + \text{var}[\tilde{y}] + \sigma^2$

where;

$\text{Bias}[\tilde{y}] = \mathbb{E}[(y - \mathbb{E}[\tilde{y}])^2]$

and

$\text{var}[\tilde{y}] = \mathbb{E}[(\tilde{y} - \mathbb{E}[\tilde{y}])^2] = \frac{1}{n}\sum_i (\tilde{y}_i - \mathbb{E}[\tilde{y}])^2$

We have our <u>model</u>

$y = f(x) + \epsilon \sim N(0, \sigma^2)$

and

$\tilde{y} = X\beta$: Model prediction

design matrix

So how to decompose $\mathbb{E}[(y-\tilde{y})^2]$?

Start by expanding the squared error term:

$(y-\tilde{y})^2 = (f(x) + \epsilon - \tilde{y})^2$

can be rewritten as:

$(f(x) - \tilde{y} + \epsilon)^2 = (f(x) - \tilde{y})^2 + 2(f(x) - \tilde{y})\epsilon + \epsilon^2.$

Now we take the expectation of this:

$\mathbb{E}[(y-y^2)] = \mathbb{E}[(f(x) - \tilde{y})^2 + \mathbb{E}[2(f(x) - \tilde{y})\epsilon] + \mathbb{E}[\epsilon^2]$

We can now simplyfy!

- $\mathbb{E}[2(f(x)-\tilde{y})\epsilon] = 0$ because $\mathbb{E}[\epsilon]=0$ and independent of $\tilde{y}$
- $\mathbb{E}[\epsilon^2]=\sigma^2$, since $\epsilon \sim N(0,\sigma^2)$

Simplyfies to:

$$\mathbb{E}[(y-\tilde{y})^2] = \mathbb{E}[(f(x)-\tilde{y})^2] + \overbrace{\sigma^2}^{\text{Noise}}$$

We now decompose the first expression $\mathbb{E}[(f(x)-\tilde{y})^2]$

$$\mathbb{E}[(f(x)-\tilde{y})^2] = \mathbb{E}[(f(x)-\mathbb{E}[\tilde{y}]+\mathbb{E}[\tilde{y}]-\tilde{y})^2]$$

Expand this:

$$= \mathbb{E}[(f(x)-\mathbb{E}[\tilde{y}]^2)] + \mathbb{E}[(\mathbb{E}[\tilde{y}]-\tilde{y})^2] + 2\mathbb{E}[(f(x)-\mathbb{E}[\tilde{y}])\underbrace{(\mathbb{E}[\tilde{y}]-\tilde{y})}_{0}]$$

The last term vanishes because $\mathbb{E}[\tilde{y}]-\tilde{y}$ has zero expectation.

We are left with:

$$=(f(x)-\mathbb{E}[\tilde{y}])^2 + \mathbb{E}[(\mathbb{E}[\tilde{y}]-\tilde{y})^2]$$

The term $(f(x)-\mathbb{E}[\tilde{y}])^2$ is the Bias

$$\text{Bias}[\tilde{y}] = (f(x)-\mathbb{E}[\tilde{y}])^2 = \mathbb{E}[(f(x)-\mathbb{E}[\tilde{y}])^2]$$

The term $\mathbb{E}[(\tilde{y}-\mathbb{E}[\tilde{y}])^2]$ is the variance.

$$\text{Var}[\tilde{y}] = \mathbb{E}[(\tilde{y}-\mathbb{E}[\tilde{y}])^2]$$

Putting it all together:

$$\mathbb{E}[y-\tilde{y}]^2 = \underbrace{\text{Bias}[\tilde{y}]}_{\text{Bias term.}} + \underbrace{\text{Var}[\tilde{y}]}_{\text{Variance term}} + \underbrace{\sigma^2}_{\text{Noise}}$$