Problem 2.22.

When there is noise in the data $E_{our}(g^{(D)})$

$$= E_{x,y}\left[(g^{(D)}(x) - y(x))^2 \right]$$

where $y(x) = f(x) + \epsilon$. If ϵ is $N(0,\sigma)$. show that bias variance decomposition becomes

$$E_D\left[E_{our}(g^{(D)}) \right] = \sigma^2 + bias + var$$

$$E_D\left[E_{our}(g^{(D)}) \right]$$

$$= E_\epsilon E_x E_D\left[(y(x) - g^{(D)}(x))^2 \right]$$

$$= E_\epsilon E_x E_D\left[(f(x) + \epsilon - g^{(D)}(x))^2 \right]$$

$$= E_\epsilon E_x E_D\left[f(x)^2 + \epsilon^2 + g^{(D)}(x)^2 + 2f(x)\epsilon - 2\epsilon g^{(D)}(x) - 2f(x)g^{(D)}(x) \right]$$

$$= E_x E_D\Big[E_\epsilon[f(x)^2] + E_\epsilon[\epsilon^2] + E_\epsilon[g^{(D)}(x)^2] + 2f(x)E_\epsilon[\epsilon] - 2g^{(D)}(x)E_\epsilon[\epsilon] - 2E_\epsilon[f(x)g^{(D)}(x)] \Big]$$

$E_\epsilon[\epsilon] = 0$ since noise is zero mean

Also $f(x)$ is independent of ϵ

$$= E_x E_D\left[f(x)^2 + var(\epsilon) - (E_\epsilon[\epsilon])^2 + g^{(D)}(x)^2 - 2f(x)g^{(D)}(x) \right]$$

$\because var(\epsilon) = E(\epsilon^2) - (E[\epsilon])^2$

$$= E_x\left[E_D[f(x)^2] + \sigma^2 + E_D[g^{(D)}(x)^2] - 2f(x)E_D[g^{(D)}(x)] \right]$$

$$= E_x\left[f(x)^2 + \sigma^2 + var(g^{(D)}(x)) + E_D[g^{(D)}(x)]^2 - 2f(x)E_D[g^{(D)}(x)] \right]$$

$$= E_x\left[\sigma^2 + var(g^{(D)}(x)) + (f(x) - E_D[g^{(D)}(x)])^2 \right]$$

$$= E_x[\sigma^2] + E_x[var(g^{(D)}(x))] + E_x[Bias]$$

$$= \sigma^2 + E_x[var(g^{(D)}(x))] + E_x[Bias]$$

$$= \sigma^2 + bias + variance$$

//