Problem 2 :

Let Y be a random variable. Show that

$$E(Y) = \text{argmin}_c E[(Y-c)^2]$$

Solution :

$$\underset{\text{L.H.S.}}{E(Y)} = \underset{\text{R.H.S.}}{\text{argmin}_c E[(Y-c)^2]}$$

→ consider R.H.S, Let $c^* = \text{argmin}_c E[(Y-c)^2]$

→ $(y-c)^2$ is a convex function with a single minimum (Global minimum) with parameter c

→ c^* can be found by equating the first derivative with respect to c → 0

$\{(a-b)^2 = a^2+b^2-2ab$

$$\Rightarrow \frac{\partial(E[(Y-c)^2])}{\partial c} = 0$$

$$\Rightarrow \frac{\partial(E[Y^2+c^2-2Yc])}{\partial c} = 0$$

$$\Rightarrow \frac{\partial(E[Y^2]+E[c^2]-2E[Yc])}{\partial c} = 0$$

$$\Rightarrow \frac{\partial E[Y^2]}{\partial c} + \frac{\partial E[c^2]}{\partial c} - \frac{\partial(2E[Yc])}{\partial c} = 0$$

$$\Rightarrow 0 + 2c - 2E[Y] = 0$$

$$\Rightarrow c = E[Y] = \text{L.H.S}$$

So $\boxed{E[Y] = \text{argmin}_c E[(Y-c)^2]}$ Proof

PRIYANKA UPADHYAY

Problem 3

Prove that Bias-Variance tradeoff with irreducible error.
Please Note that you should prove both equalities.

$$E\left[\{y_0-\hat{f}(x_0)\}^2\right]$$
$$=E\left[\{\hat{f}(x_0)-E(\hat{f}(x_0))\}^2\right]+\left[E\{\hat{f}(x_0)\}-f(x_0)\right]^2$$
$$+\operatorname{var}(\varepsilon)$$
$$=\operatorname{Var}(\hat{f}(x_0))+\left[\operatorname{Bias}(\hat{f}(x_0))\right]^2+\operatorname{Var}(\varepsilon)$$

Solution

As we know $y_0=f(x_0)+\varepsilon$

where ε is random number with expected value $\hat{\varepsilon}=E[\varepsilon]=0$ and variance $E[(\varepsilon-\hat{\varepsilon})^2]=E[\varepsilon^2]$
$=\sigma^2$
$=\operatorname{var}(\varepsilon)$

So,
$$E\left[\{y_0-\hat{f}(x_0)\}^2\right] \quad -(1)$$
$$\Rightarrow E\left[\{\underbrace{f(x_0)+\varepsilon}_{y_0 \text{ value}}-\hat{f}(x_0)\}^2\right]$$
$$\Rightarrow E\left[\{\underbrace{f(x_0)-\hat{f}(x_0)}_{a}+\underbrace{\varepsilon}_{b}\}^2\right] \qquad \because (a+b)^2=a^2+b^2+2ab$$
$$\Rightarrow E\left[\{f(x_0)-\hat{f}(x_0)\}^2\right]+E(\varepsilon)^2+2E\left[\{f(x_0)-\hat{f}(x_0)\}\varepsilon\right]$$

as ε is an Independent Random Number then
$$2E\left[\{f(x_0)-\hat{f}(x_0)\}\varepsilon\right]$$
$$=2E\left[\{f(x_0)-\hat{f}(x_0)\}\right]\underbrace{E[\varepsilon]}_{0}\Rightarrow 0$$

So $E[\{f(x_0)-\hat{f}(x_0)\}^2] + E[\varepsilon]^2 + 0$

as we know that $E[\varepsilon]^2 = \mathrm{Var}(\varepsilon)$

$\therefore E[\{f(x_0)-\hat{f}(x_0)\}^2] + \mathrm{var}(\varepsilon) \rightarrow$ (2)

= Adding and Subtracting $E(\hat{f}(x_0))$

$$\Rightarrow E[\{\underbrace{f(x_0)-E(\hat{f}(x_0))}_{a} + \underbrace{E(\hat{f}(x_0))-\hat{f}(x_0)}_{b}\}^2] + \mathrm{var}(\varepsilon) \quad -(3)$$

$$\Rightarrow E[\{\hat{f}(x_0)-E(\hat{f}(x_0))\}^2] + E[\{E(\hat{f}(x_0))-\hat{f}(x_0)\}^2]$$
$$+ 2E[\{f(x_0)-E(\hat{f}(x_0))\}\{E(\hat{f}(x_0))-\hat{f}(x_0)\}]$$
$$+ \mathrm{var}(\varepsilon)$$

$$\Rightarrow E[\{\hat{f}(x_0)-E(\hat{f}(x_0))\}^2] + E[\{E(\hat{f}(x_0))-\hat{f}(x_0)\}^2]$$
$$\underbrace{+2E[f(x_0).\cancel{E(\hat{f}(x_0))} - f(x_0)\hat{f}(x_0) - E^2(\hat{f}(x_0)) + \cancel{E(\hat{f}(x_0))}.\hat{f}(x_0)]}_{0}$$
$$+ \mathrm{Var}(\varepsilon)$$

$$\Rightarrow \boxed{E[\{\hat{f}(x_0)-E(\hat{f}(x_0))\}^2] + E[\{E(\hat{f}(x_0))-\hat{f}(x_0)\}^2] + \mathrm{Var}(\varepsilon)}$$

Proof 1

and as we know that from Book

$$Var(\hat{f}(x_0)) = E[\hat{f}(x_0) - E(\hat{f}(x_0))]^2$$

and $Bias(\hat{f}(x_0)) = E(f(x_0) - y_0)$

and $y_0 = f(x_0) + \underset{0}{\varepsilon}$ $(\because$ so $y_0 = f(x_0)$

$$\therefore \underbrace{E[\{\hat{f}(x_0) - E(\hat{f}(x_0))\}^2]}_{Var(\hat{f}(x_0)} + [\{\underbrace{E(\hat{f}(x_0)) - y_0}_{Bias(\hat{f}(x_0))}\}^2] + Var(\varepsilon)$$

$$\therefore \Rightarrow \boxed{Var(\hat{f}(x_0)) + [Bias(\hat{f}(x_0))]^2 + \underbrace{Var(\varepsilon)}_{\text{irreducible error}}}$$

Proof 2