An alternative proof for the bias-variance trade-off proof

Given any two random variables Y and Z and its expectations EY and EZ,

$$Y - Z = (Y - EY) + (EY - EZ) - (Z - EZ).$$

Recall that

$$(a + b + c)^2 = a^2 + b^2 + c^2 + 2ab + 2ac + 2bc.$$

For $a = Y - EY, b = EY - EZ, c = -(Z - EZ)$, we can write

$$\begin{aligned} & (Y - Z)^2 \\ = \; & (Y - EY)^2 + (EY - EZ)^2 + (Z - EZ)^2 \\ + \; & 2(Y - EY)(EY - EY) - 2(Y - EY)(Z - EZ) - 2(EY - EZ)(Z - EZ). \end{aligned}$$

Taking expectations on both sides, because

$$E(Y - EY) = 0, E(Z - EZ) = 0,$$

and assume $Y - EY$ and $Z - EZ$ are uncorrelated, we have

$$E(Y - Z)^2 = E(Y - EY)^2 + (EY - EZ)^2 + Var(Z).$$

Recall $EY = h_\theta(X)$, and for the feature variable X, let Z be the predictor $f_\theta(X)$. If we assume the noise term $Y - h_\theta(X)$ is independent of X, which implies the uncorrelatedness of the noise term and $Z = f_\theta(X)$, we recover the identity given in the lecture today (April 4, 2017) by Professor Gonzalez:

$$E(Y - f_\theta(X))^2 = Noise\ variance + bias^2 + var.of f_\theta(X).$$