

Décomposition biais-variance

$$EQM = E[(\theta - \hat{\theta})^2] = E(\theta^2) + E(\hat{\theta}^2) - 2E(\theta \hat{\theta})$$

avec $\theta = f(x) + \varepsilon$ et $\hat{\theta} = \hat{f}(x)$

$$= \underbrace{E[(f(x) + \varepsilon)^2]}_{\downarrow} + E[f(x)^2] - 2 \underbrace{E[(f(x) + \varepsilon) \hat{f}(x)]}$$
$$\begin{aligned} & 2E[f(x)\hat{f}(x) + \varepsilon\hat{f}(x)] \\ & 2E[f(x)\hat{f}(x)] + 2E(\varepsilon\hat{f}(x)) \\ & 2E[f(x)\hat{f}(x)] + 2[E(\hat{f}(x))E(\varepsilon) - 2\text{COV}(\varepsilon, \hat{f}(x))] \\ & 2E[f(x)\hat{f}(x)] + 2E(\hat{f}(x))E(\varepsilon) - 4\text{COV}(\varepsilon, \hat{f}(x)) \end{aligned}$$

$$\begin{aligned} & E[f(x)^2 + \varepsilon^2 + 2f(x)\varepsilon] \\ & E[f(x)^2] + E(\varepsilon^2) + 2E(f(x))E(\varepsilon) - 4\text{COV}(f(x), \varepsilon) \end{aligned}$$

$$\begin{aligned} * &= E(f(x)^2) + E(\hat{f}(x)^2) - 2E(f(x)\hat{f}(x)) = E((f(x) - \hat{f}(x))^2) \\ &= E[(f(x) - \hat{f}(x))^2] \\ &= \text{Var}(f(x) - \hat{f}(x)) + \underbrace{[E(f(x) - \hat{f}(x))]^2} \end{aligned}$$

\downarrow
 $f(x)$ est déterministe
 \downarrow
 $\text{Var}(\hat{f}(x))$

Biais de $\hat{f}(x)$ au carré

Donc

$$EQM = \text{Var}(\hat{f}(x)) + [\text{biais}(f(x))]^2 + \dots$$

$$\dots + E(\varepsilon^2) + 2E(f(x))E(\varepsilon) - 4\text{COV}(f(x), \varepsilon) + \dots$$

$$\dots - 2E(\hat{f}(x))E(\varepsilon) + 4\text{COV}(\hat{f}(x), \varepsilon)$$

Si $E(\varepsilon) = 0$

$$\begin{aligned} & \rightarrow E(\varepsilon^2) = \text{Var}(\varepsilon) + [E(\varepsilon)]^2 = \text{Var}(\varepsilon) \\ & \rightarrow 2E(f(x))E(\varepsilon) = 0 \\ & \rightarrow 2E(\hat{f}(x))E(\varepsilon) = 0 \end{aligned}$$

Donc

$$EQM = \text{var}(\hat{f}(x)) + [\text{Biais}(\hat{f}(x))]^2 + \text{var}(\varepsilon^2) + \dots$$

$$\dots + 4 \text{COV}(\hat{f}(x), \varepsilon) - 4 \text{COV}(f(x), \varepsilon)$$

Questions ...

Quelle est la covariance entre $f(x)$ et ε ?

$f(x)$ est une probabilité ...

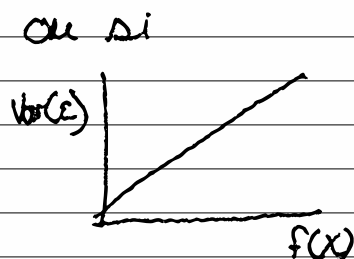
sa variance est $p(1-p)$ ou $\frac{p(1-p)}{n}$

Si $\text{COV}(f(x), \varepsilon)$ est grand... l'est aussi

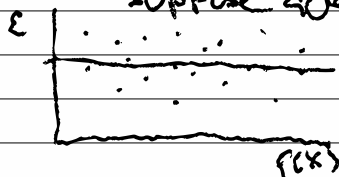
Qu'est ce qui se passe si...



comme avec Y binaire?



En régression linéaire, on suppose que...



Est-ce qu'on fait le même genre d'hypothèse avec Y binaire?