BIAS (SESGO) vs VARIANZA

(a.) *Vemos que* $\hat{\theta}_b = (1+m)\hat{\theta}_u$ *y* $E\left[\hat{\theta}_b\right] = (1+m)E\left[\hat{\theta}_u\right] = (1+m)\theta$. *El sesgo cuadrático es de la forma siguiente:*

$$bias^2\left(\hat{\theta}_b\right) = \left|\theta - E\left[\hat{\theta}_b\right]\right|^2 = \left|\theta - (1+m)\theta\right|^2 = m^2\theta^2$$

(b.) *Para el término de varianza:*

$$\mathrm{var}\left(\hat{\theta}_b\right) = E\left[\left|\hat{\theta}_b - E\left[\hat{\theta}_b\right]\right|^2\right] = E\left[\left|(1+m)\hat{\theta}_u - (1+m)\theta\right|^2\right] = (1+m)^2 E\left[\left|\hat{\theta}_u - E\left[\hat{\theta}_u\right]\right|^2\right] =$$
$$= (1+m)^2\,\mathrm{var}\left(\hat{\theta}_b\right)$$

(c.) *El nuevo término de MSE es la suma de la varianza y el sesgo cuadrático:*

$$MSE\left(\hat{\theta}_b\right) = E\left[\left|\theta - \hat{\theta}_b\right|^2\right] = \left|\theta - E\left[\hat{\theta}_b\right]\right|^2 + \mathrm{var}\left(\hat{\theta}_u\right) = m^2\theta^2 + (1+m)^2\,\mathrm{var}\left(\hat{\theta}_u\right)$$

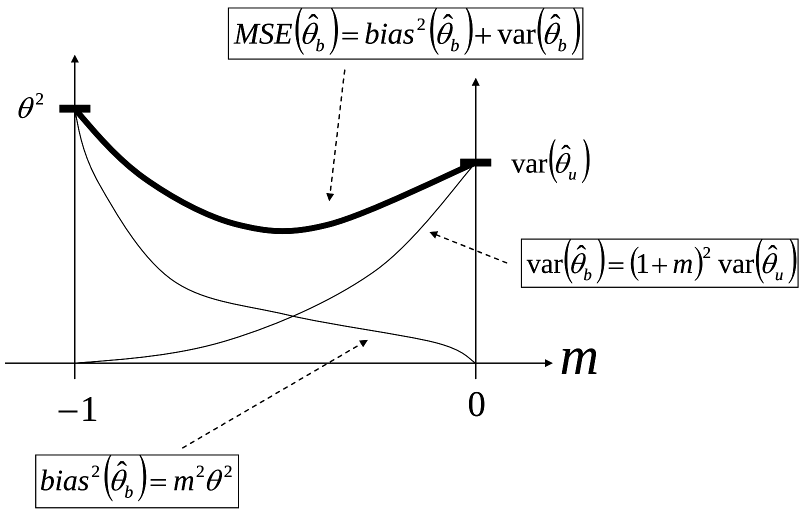
El mínimo se obtiene de la forma habitual:

$$\frac{\partial}{\partial m} MSE\left(\hat{\theta}_b\right) = \frac{\partial}{\partial m}\left(m^2\theta^2 + (1+m)^2\,\mathrm{var}\left(\hat{\theta}_u\right)\right) = 2m\theta^2 + 2(1+m)\mathrm{var}\left(\hat{\theta}_u\right) = 0$$

El valor óptimo de 'm' se alcanza para:

$$m = -\frac{\mathrm{var}\left(\hat{\theta}_u\right)}{\theta^2 + \mathrm{var}\left(\hat{\theta}_u\right)} = -\frac{1}{1+\rho} \quad \textit{como } \rho = \frac{\theta^2}{\mathrm{var}\left(\hat{\theta}_u\right)} \geq 0 \Rightarrow -1 \leq m \leq 0$$

(d.)



(e.) *Tenemos que:*

$$\hat{\theta}_b = (1+m)\hat{\theta}_u = \left(1 - \frac{1}{1+\rho}\right)\hat{\theta}_u = \frac{\rho}{1+\rho}\hat{\theta}_u = \frac{\theta^2}{\theta^2 + \mathrm{var}\left(\hat{\theta}_u\right)}\hat{\theta}_u$$

$$MSE\left(\hat{\theta}_b\right) = m^2\theta^2 + (1+m)^2\,\mathrm{var}\left(\hat{\theta}_u\right) = \frac{\rho}{1+\rho}\mathrm{var}\left(\hat{\theta}_u\right)$$

Para la distribución exponencial anterior: $\rho = N$ *y:*

$$MSE\left(\hat{\theta}_b\right) = \frac{\rho}{1+\rho}\mathrm{var}\left(\hat{\theta}_u\right) = \frac{N}{1+N}\mathrm{var}\left(\hat{\theta}_u\right) < \mathrm{var}\left(\hat{\theta}_u\right) = MSE\left(\hat{\theta}_u\right)$$

Especialmente de interés para tamaños de observación N *pequeños.*

(f.) La mejora en el MSE queda como la diferencia:

$$MSE(\hat{\theta}_b) - MSE(\hat{\theta}_u) = m^2\theta^2 + (1+m)^2 \operatorname{var}(\hat{\theta}_u) - \operatorname{var}(\hat{\theta}_u) =$$
$$= m^2\theta^2 + (1+m)^2 V - V < 0$$

Un punto importante es que la expresión anterior <u>*siempre es máxima*</u> *para* $|\theta| = \theta_o$ *en el intervalo* $|\theta| \le \theta_o$.

Luego:

$$m(2V + m(\theta^2 + V)) < 0 \quad como\ m < 0 \Rightarrow 2V + m(\theta^2 + V) > 0$$

Por tanto:

$$m > -\frac{2V}{V+\theta^2} \Rightarrow (1+m) > \frac{\theta^2 - V}{\theta^2 + V}$$

(g.) La mejora en el MSE siempre es máxima para $|\theta| = \theta_o$ *en el intervalo* $|\theta| \le \theta_o$.

$$MSE(\hat{\theta}_b) - MSE(\hat{\theta}_u) = m^2\theta_o^2 + (1+m)^2 V - V < 0$$

Luego:

$$\frac{\partial}{\partial m}\left(MSE(\hat{\theta}_b) - MSE(\hat{\theta}_u)\right) = \frac{\partial}{\partial m}\left(m^2\theta_o^2 + (1+m)^2 V - V\right) = 0$$

Finalmente:

$$m = -\frac{V}{V+\theta_o^2} \Rightarrow \hat{\theta}_b = \frac{\theta_o^2}{V+\theta_o^2}\hat{\theta}_u$$

(h.) Vemos que para un rango de parámetros $|\theta| \le \theta_o$ *arbitrariamente grande* $\theta_o \to \infty$, *tenemos de la solución en (g.):*

$$\lim_{\theta_o \to \infty} \hat{\theta}_b = \lim_{\theta_o \to \infty} \frac{\theta_o^2}{V+\theta_o^2}\hat{\theta}_u \to \hat{\theta}_u$$

y no es posible aportar una solución al problema.