Example of biased estimator better than MVUE

$$\hat{\theta}_b = (1+a)\cdot\hat{\theta}_{MVUE}$$

$$E[\hat{\theta}_b] = (1+a)E[\hat{\theta}_{MVUE}] = (1+a)\theta_0$$

$$MSE(\hat{\theta}_b) = \underbrace{E\left[\left(\hat{\theta}_b - E[\hat{\theta}_b]\right)^2\right]}_{\text{variance (1st term)}} + \underbrace{\left(E[\hat{\theta}_b] - \theta_0\right)^2}_{\text{bias}^2\text{ (2nd term).}}$$

$$(2nd) = \left[(1+a)\theta_0 - \theta_0\right]^2 = a^2\theta_0^2$$

$$(1st) = E\left[\left((1+a)\hat{\theta}_{MVUE} - (1+a)\theta_0\right)^2\right] =$$

$$= (1+a)^2 \, E\left[\left(\theta_{MVUE} - \theta_0\right)^2\right] =$$

$$= (1+a)^2 \, MSE\left(\hat{\theta}_{MVUE}\right).$$

$$\underbrace{MSE\left(\hat{\theta}_b\right)}_{y} = (1+a)^2 \underbrace{MSE\left(\hat{\theta}_{MVUE}\right)}_{x} + a^2\theta_0^2$$

We want: $y < x$ i.e $y - x < 0$

$$y - x = (1+a)^2 x + a^2\theta_0^2 - x$$

$$= \cancel{x} + a^2 x + 2ax + a^2\theta_0^2 - \cancel{x} = a\left[a(x+\theta_0^2) + 2x\right]$$

Want < 0.

$a \geq 0$ No

$a < 0$ and $a(x+\theta_0^2) + 2x > 0 \implies a > -\dfrac{2x}{x+\theta_0^2}$

$$\boxed{-\frac{2x}{x+\theta_0^2} < a < 0}$$

I got
a biased
estimator
better than the best
unbiased one.

Show that: $\boxed{|1+a| < 1.}$

equivalent to: $-1 < 1+a < 1$ $\Rightarrow \begin{cases} a < 0 \checkmark \\ a > -2 \checkmark \end{cases}$

$\hat{\theta}_b = (1+a)\hat{\theta}_{MVUE}$.

Norm shrinks

Biased estimator is better than the unbiased

Regularisation ✓

Optimal value: $\frac{dy}{da} = 0 = 2(1+a)x + 2a\theta_0^2$

$\Rightarrow \boxed{a_* = \frac{x}{x+\theta_0^2}}$ $\rightarrow$ Don't know.

Ridge regression – toy problem.

$$\left[\sum_{n=1}^{N}(x_n x_n)+\lambda\right]\hat{\theta}_b=\sum_{n=1}^{N}y_n x_n.$$

where each $x_n=1$.

$$(N+\lambda)\hat{\theta}_b=\sum_{n=1}^{N}y_n \Rightarrow \hat{\theta}_b=\frac{\sum_n y_n}{N+\lambda}=\frac{N}{N+\lambda}\bar{y}=\frac{N}{N+\lambda}\hat{\theta}_{LS}.$$

$$E[\hat{\theta}_b]=\frac{N}{N+\lambda}E[\hat{\theta}_{LS}]=\frac{N}{N+\lambda}\theta_0.$$

$$MSE(\hat{\theta}_b)=E\left[(\hat{\theta}_b-\theta_0)^2\right]=E\left[\left(\frac{N}{N+\lambda}\bar{y}-\theta_0\right)^2\right]$$

Put: $\frac{N}{N+\lambda} = z$

$$MSE = z^2 E[(\bar{y})^2] - 2z E[\bar{y}]\theta_0 + \theta_0^2 = z^2 E[(\bar{y})^2] - 2z\theta_0^2 + \theta_0^2$$

$$\sigma^2[\bar{y}] = E[(\bar{y})^2] - (E[\bar{y}])^2$$

$\sigma^2[\bar{y}] \to \sigma_n^2/N$, $E[\bar{y}] \to \theta_0$

$$E[(\bar{y})^2] = \frac{\sigma_n^2}{N} + \theta_0^2$$

$$MSE(\hat{\theta}_b) = z^2\left[\frac{\sigma_n^2}{N} + \theta_0^2\right] - 2z\theta_0^2 + \theta_0^2$$

Seek the minimum of MSE:

$$\frac{d\,MSE(\hat{\theta}_b)}{dz} = 0 \rightarrow z_* = \frac{\theta_0^2}{\frac{\sigma_n^2}{N} + \theta_0^2} = \frac{1}{\frac{\sigma_n^2}{N\theta_0^2} + 1}.$$

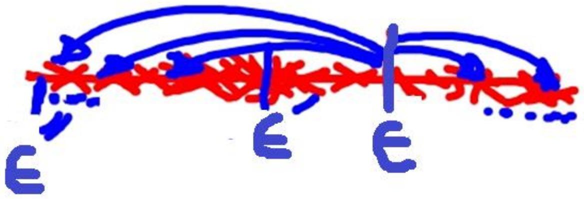
$$z = \frac{N}{N+\lambda} = \frac{1}{1 + \frac{\lambda}{N}} \qquad \lambda_* = \frac{\sigma_n^2}{\theta_0^2}.$$

$$MSE(\hat{\theta}_b^*) = \frac{\sigma_n^2}{N} \frac{1}{1 + \frac{\sigma_n^2}{N\theta_0^2}} < \frac{\sigma_n^2}{N}.$$

$p(y)$

$$G = E\left[(y-\varepsilon)^2\right]$$

$$G = \int (y-\varepsilon)^2 p(y)\,dy =$$



$$= \int y^2 p(y)\,dy - 2\varepsilon \int y\, p(y)\,dy$$

$$+ \varepsilon^2 \underbrace{\int p(y)\,dy}_{1} \,. =$$

$$= E\left[y^2\right] - 2\varepsilon \cdot E[y] + \varepsilon^2$$

$$\frac{dG}{dE} = -2E[y] + 2E = 0.$$

$$\boxed{E = E[y].}$$