

Example of biased estimator
better than MVUE

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

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

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

$$(2nd) = [(1+d)\theta_0 - \theta_0]^2 = d^2\theta_0^2$$

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

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

$$= (1+d)^2 \text{MSE}(\hat{\theta}_{MVUE}).$$

$$\underbrace{\text{MSE}(\hat{\theta}_b)}_y = (1+d)^2 \underbrace{\text{MSE}(\hat{\theta}_{MVUE})}_x + d^2\theta_0^2$$

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

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

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

Want < 0 .

$a \geq 0$ No

$$a < 0 \text{ and } a(x + \theta_0^2) + 2x > 0 \Rightarrow a > -\frac{2x}{x + \theta_0^2}$$

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

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

Show that: $|1+a| < 1$.

equivalent to: $-1 < 1+a < 1$

$$\Rightarrow \begin{cases} a < 0 \\ a > -2 \end{cases}$$

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

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 $a_* = \frac{x}{x + \theta_0^2}$ \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 $x_n = 1$.

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

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

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

$$\text{Put: } \frac{N}{N+1} = z$$

$$\begin{aligned} \text{MSE} &= z^2 E[(\bar{y})^2] - \\ &- 2z E[\bar{y}] \theta_0 + \theta_0^2 = z^2 E[(\bar{y})^2] - 2z\theta_0^2 + \theta_0^2 \end{aligned}$$

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

$\swarrow \frac{\sigma_n^2}{N}$ $\swarrow \theta_0^2$

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

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

Seek the minimum of MSE:

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

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

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

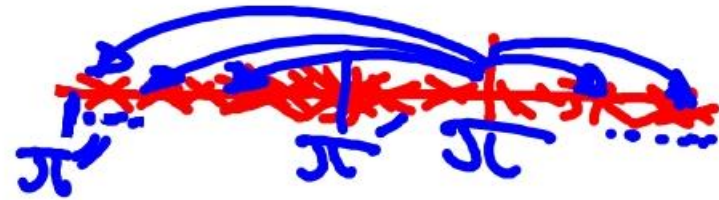
$p(y)$

$$G = E[(y - \pi)^2]$$

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

$$= \int y^2 p(y) dy - 2\pi \int y p(y) dy + \pi^2 \int p(y) dy =$$

$$= E[y^2] - 2\pi E[y] + \pi^2$$



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

$$\pi = E[y].$$