Labs3

Jakub Kuciński

2022-04-22

Task 1

First let's calculate MSE on single coordinate:

$$
\begin{aligned}
\mathbb{E}\|\hat{\beta}_i - \beta_i\|^2 &= \mathbb{E}\|((X'X + \lambda I)^{-1} X'Y)_i - \beta_i\|^2 = && (1)\\
&= \mathbb{E}\|((1+\lambda)^{-1} X'Y)_i - \beta_i\|^2 = && (2)\\
&= \mathbb{E}\|((1+\lambda)^{-1} X'(X\beta + \epsilon))_i - \beta_i\|^2 = && (3)\\
&= \mathbb{E}\|((1+\lambda)^{-1} X'(X\beta + \epsilon))_i - \beta_i\|^2 = && (4)\\
&= \mathbb{E}\|((1+\lambda)^{-1} (X'X\beta + X'\epsilon))_i - \beta_i\|^2 = && (5)\\
&= \mathbb{E}\|((1+\lambda)^{-1} (\beta + Z))_i - \beta_i\|^2 = && (6)\\
&= \mathbb{E}\left\|\frac{Z_i}{\lambda+1} - \frac{\lambda\beta_i}{\lambda+1}\right\|^2 = && (7)\\
&= \mathbb{E}\left[\frac{1}{(\lambda+1)^2} Z_i^2\right] - 2\mathbb{E}\left[\frac{\lambda}{(\lambda+1)^2} Z_i\beta_i\right] + \mathbb{E}\left[\frac{\lambda^2}{(\lambda+1)^2}\beta_i^2\right] = && (8)\\
&= \frac{1}{(\lambda+1)^2}\sigma^2 - 2\cdot 0 + \frac{\lambda^2}{(\lambda+1)^2}\beta_i^2 = && (9)\\
&= \frac{1}{(\lambda+1)^2}\sigma^2 + \frac{\lambda^2}{(\lambda+1)^2}\beta_i^2 && (10)
\end{aligned}
$$

Now moving to the vector norm we get:

$$\mathbb{E}\|\hat{\beta} - \beta\|^2 = \frac{p}{(\lambda+1)^2}\sigma^2 + \frac{\lambda^2}{(\lambda+1)^2}\|\beta\|^2$$

We can find minimum of MSE by calculating the derivative of MSE with respect to λ:

$$\frac{\partial}{\partial\lambda}\mathbb{E}\|\hat{\beta} - \beta\|^2 = \frac{\partial}{\partial\lambda}\left(\frac{p}{(\lambda+1)^2}\sigma^2 + \frac{\lambda^2}{(\lambda+1)^2}\|\beta\|^2\right) = \frac{-2p}{(\lambda+1)^3}\sigma^2 + \frac{2\lambda}{(\lambda+1)^3}\|\beta\|^2 = 0$$

Now we can get the lambda that minimizes MSE:

$$\lambda = \frac{p\sigma^2}{\|\beta\|^2}$$

For $k = 20, 100, 200$ we get optimal values of lambda equal to: 3.877551, 0.7755102, 0.3877551.

The expression for bias can also be easily obtained:

$$\text{bias}(\hat{\beta}_i) = \mathbb{E}[\hat{\beta}_i - \beta_i] = \mathbb{E}\left[\frac{Z_i}{\lambda+1} - \frac{\lambda\beta_i}{\lambda+1}\right] = \mathbb{E}\left[\frac{Z_i}{\lambda+1}\right] - \mathbb{E}\left[\frac{\lambda\beta_i}{\lambda+1}\right] = 0 - \frac{\lambda\beta_i}{\lambda+1} = -\frac{\lambda}{\lambda+1}\beta_i$$

For $\beta_i = 0$ bias is actually zero. For $\beta_i = 3.5$ and optimal values of λ for $k = 20, 100, 200$ we obtain expected biases: -2.7824268, -1.5287356, -0.9779412