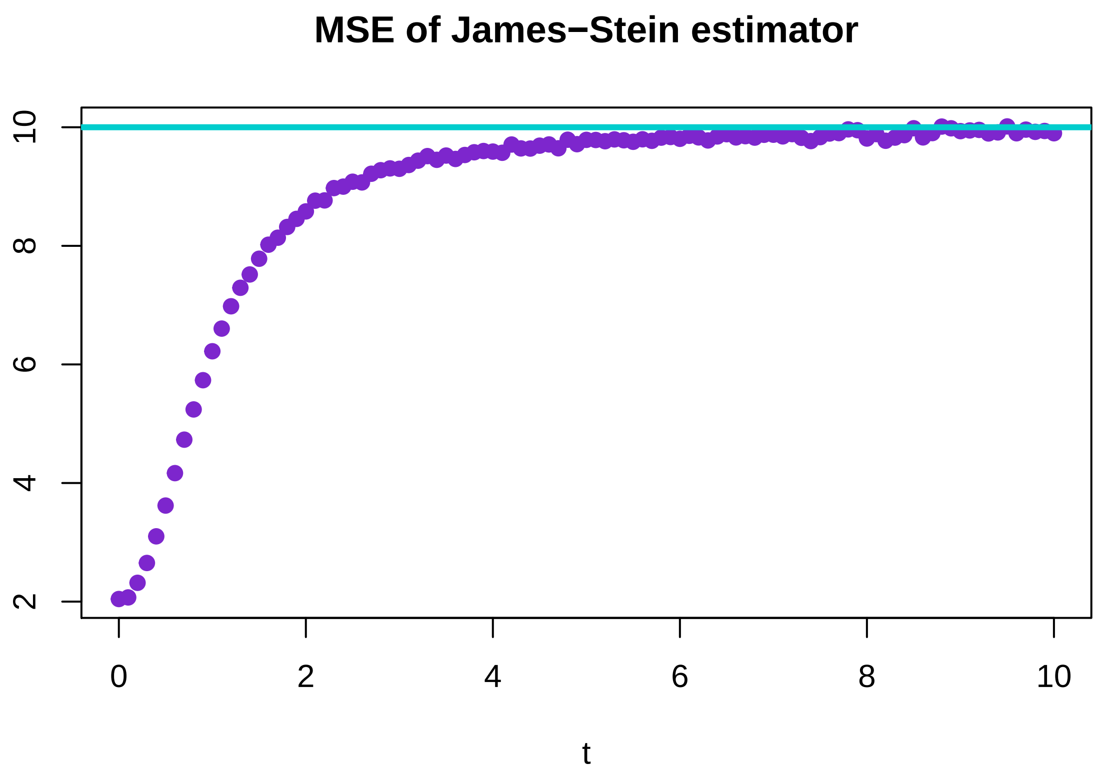
Laboratorium 5

Aleksander Milach

17/06/2020

Exercise 1



For values of parameter t - the means of a random vector Z, the mean squared error of James - Stein estimator is always just below the value of that parameter for Z. For values of t closer to 0, the difference is getting larger, the closer to 0 we are.

Exercise 2

Point 1)

We are going to proof a following inequality

$$P\left(\exists i \notin supp(\beta) : \hat{\beta}_i(\lambda_0) \neq 0\right) \leq \alpha,$$

where $\lambda_0 = \phi^{-1}\left(\dfrac{1 + \sqrt[p]{1-\alpha}}{2}\right)$.

Let us notice first, that since when X is an orthogonal matrix, $\hat{\beta}^{OLS} = X^T Y = X^T(X\beta + \epsilon) = X^T X\beta + X^T X\epsilon$, if $\beta_i = 0$ namely $i \notin supp(\beta)$, $\hat{\beta}_i^{OLS} \sim N(0,1)$. Now we can use a property of quantile function of normal distribution for that variable

$$P\left(|\hat{\beta}_i^{OLS}| \leq \phi^{-1}\left(\frac{1+\sqrt[p]{1-\alpha}}{2}\right) : i \notin supp(\beta)\right) = \sqrt[p]{1-\alpha},$$

which is equivalent to

$$P\left(|\hat{\beta}_i^{OLS}| \leq \lambda_0 : i \notin supp(\beta)\right) = \sqrt[p]{1-\alpha}.$$

Now we can use a property of LASSO estimator, which by its construction is equal to 0, when the absolute value of OLS estimator is smaller than the value of the tuning parameter, namely

$$P\left(\hat{\beta}_i(\lambda_0) = 0 : i \notin supp(\beta)\right) = P\left(|\hat{\beta}_i^{OLS}| \leq \lambda_0 : i \notin supp(\beta)\right) = \sqrt[p]{1-\alpha},$$

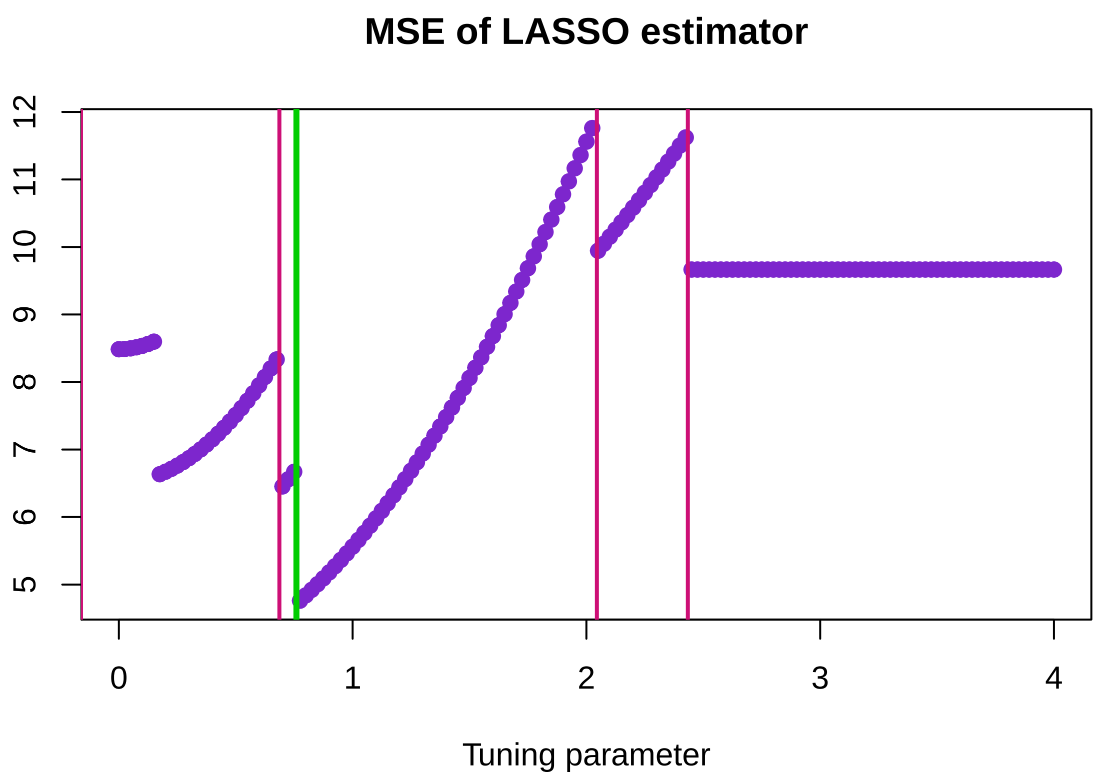
which gives

$$P\left(\hat{\beta}_i(\lambda_0) = 0 : \forall_i \notin supp(\beta)\right) \geq \left(\sqrt[p]{1-\alpha}\right)^p = 1-\alpha,$$

thus by calculating the probability of the opposite event

$$P\left(\exists i \notin supp(\beta) : \hat{\beta}_i(\lambda_0) \neq 0\right) \leq \alpha.$$

Point 2)



From the plot we can notice, that MSE is minimal just after the graph crosses the green line corresponding to one of the components of OLS estimator for β. If we calculate mean squared error for that value of λ we obtain, that minimal MSE is reached for $\lambda_1 = 0.7596$ and equals 4.7142.

Point 3)

After simulating a lot of values for ϵ we receive, that the mean squared prediction error for λ_0 equals 7.0016, and for λ_1 equals 2.9832. As we expect, the MSE is much lower, when the tuning parameter has a value of λ_1.

Point 4)

After simulating a lot of values for ϵ we receive, that the family wise error rate for λ_0 equals 0.019, and for λ_1 equals 0.551. As we expect, the FWER is below $\alpha = 0.05$, when the tuning parameter has a value of λ_0, however increases greatly for the tuning parameter value λ_1.