Mini-course Machine Learning in Empirical Economic Research

Andreas Dzemski[1]

[1]University of Gothenburg

May 17, 2019

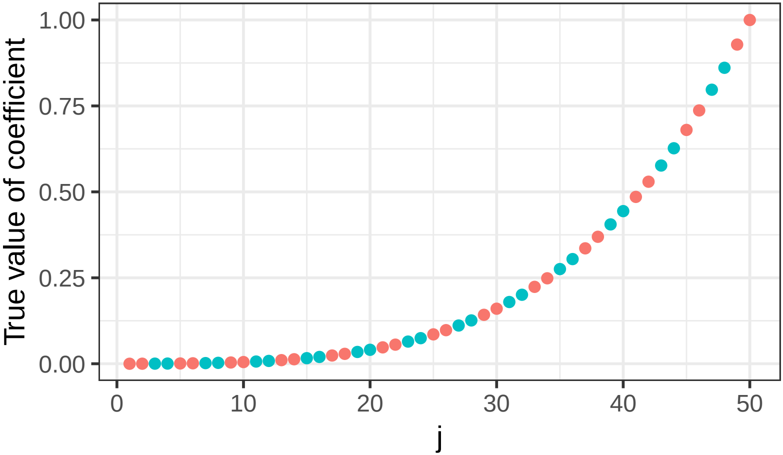

Figure: True values of coefficients

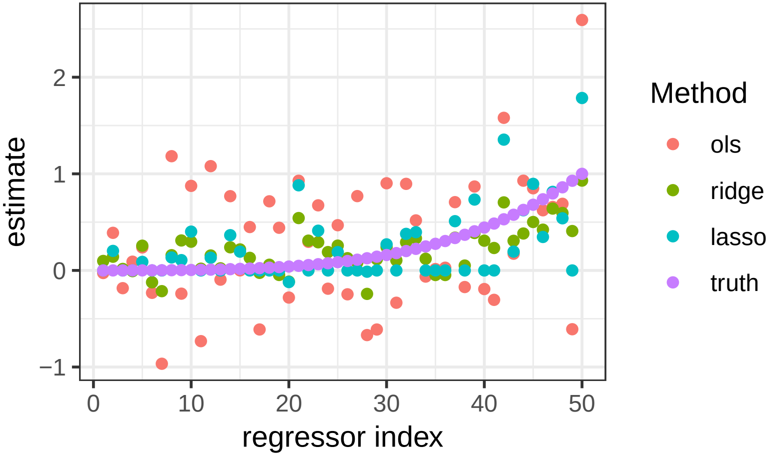

Figure: Estimation results

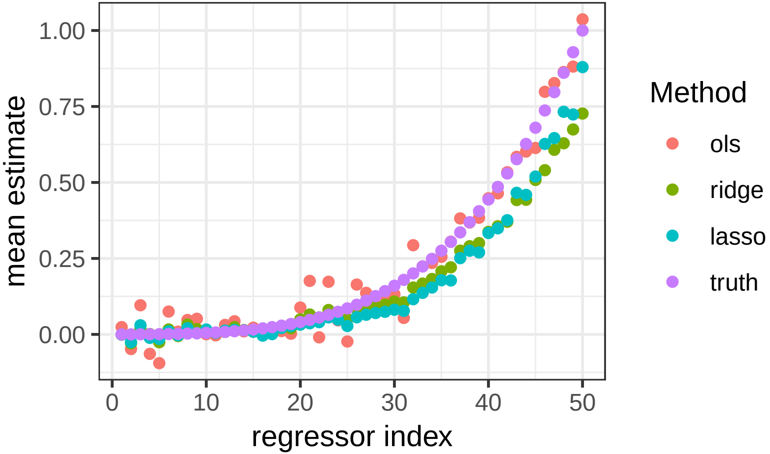

Figure: Expected estimates (average over 200 simulations)

OLS is terrible for prediction

	method	mse
1	ols	27.91
2	ridge	2.73
3	lasso	2.76

Table: Mean-squared-error $MSE(f)$

$$
\begin{aligned}
MSE(f) =& \mathbb{E} \int_x \left(\hat{f}(x) - f(x) \right)^2 \, dF(x) \\
=& \int_x \left(\mathbb{E}\hat{f}(x) - f(x) \right)^2 \, dF(x) + \int_x \mathbb{E} \left(\hat{f}(x) - \mathbb{E}\hat{f}(x) \right)^2 \, dF(x) \\
=& \int_x bias^2 \left(f(x) \right) \, dF(x) + \int_x \mathrm{var} \left(\hat{f}(x) \right) \, dF(x).
\end{aligned}
$$

Under Gauss-Markov assumptions OLS is unbiased

$$\int_x bias^2\left(f(x)\right)\, dF(x) = \int_x (\mathbb{E}(\hat{\beta}'x) - \beta'x)\, dF(x)$$
$$= \int_x \big(\underbrace{\mathbb{E}[\hat{\beta} - \beta]}_{=0}\big)'x\, dF(x) = 0$$

- For prediction unbiasedness is not a desirable property.
- Not surprising that Ridge performs best (James-Stein estimator, Empirical Bayes theory)

Lasso selects variables

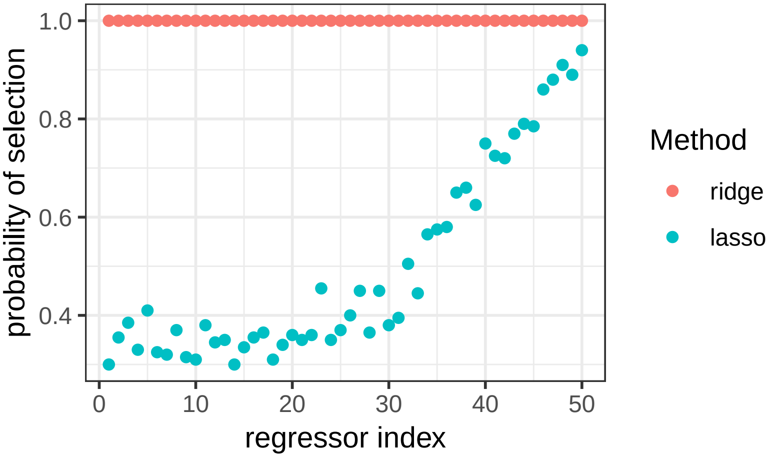

Figure: Probability of including variables (average over 200 simulations)

A sparse design

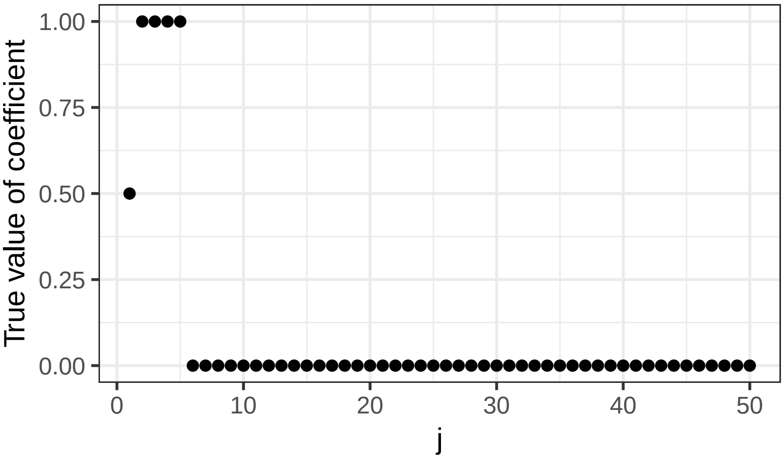

Figure: True values of coefficients

Lasso detects many of the zero coefficients

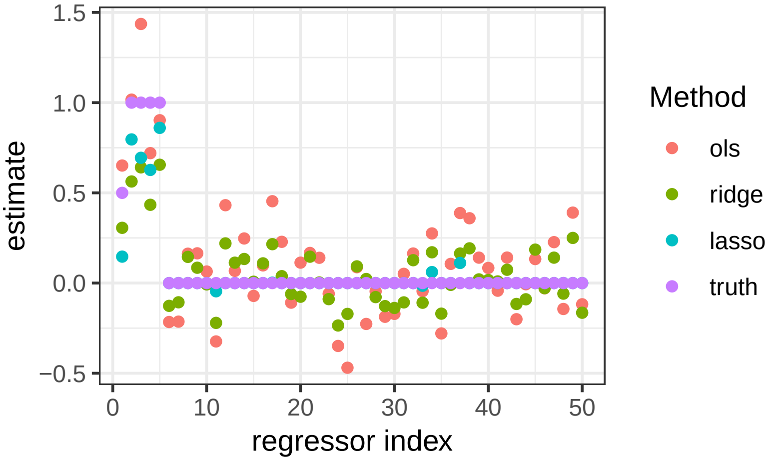

Figure: Estimation results

Lasso is good at selecting the true model

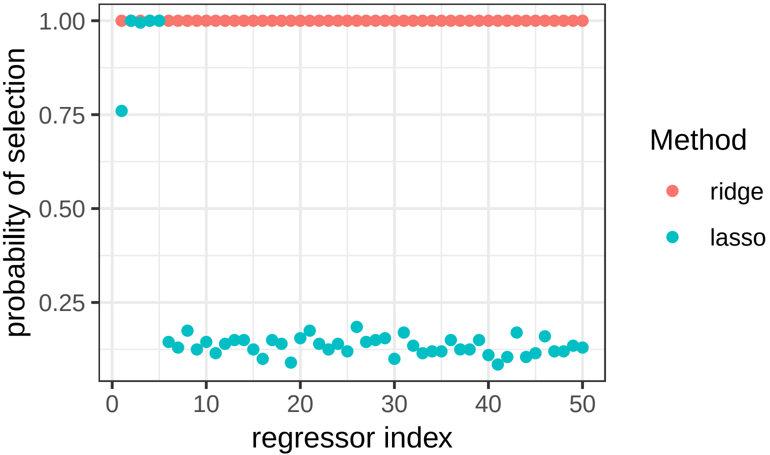

Figure: Probability of including variables (average over 200 simulations)

But still shrinkage

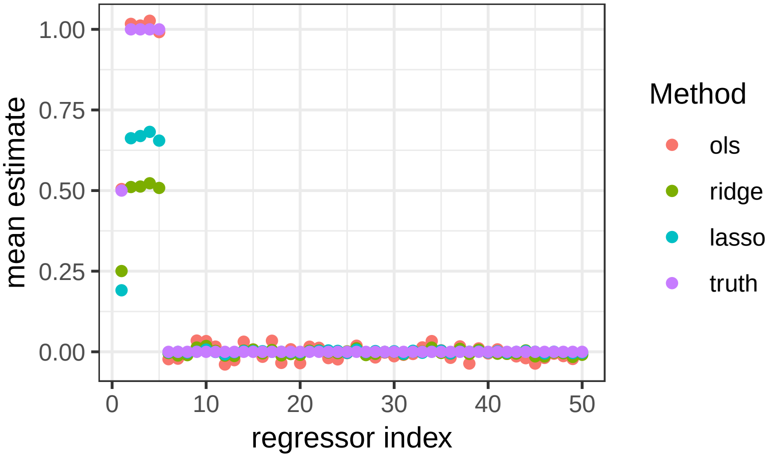

Figure: Expected estimates (average over 200 simulations)