Probabilities of Selecting Each Model

Zheng Yuan

October 2019

1 Simulation Result

Probabilities(Based on 1,000 Simulations) of Selecting Each Model					
Coefficients	**Model**	**Category**	**LOOCV**	**LOOBIC**	**BIC**
$\beta = (2, 0, 0, 4, 0)'$	1,4	Optimal	0.596	0.991	0.956
	1,2,4	II	0.108	0.002	0.015
	1,3,4	II	0.107	0.004	0.014
	1,4,5	II	0.116	0.003	0.015
	1,2,3,4	II	0.023	0.000	0.000
	1,2,4,5	II	0.022	0.000	0.000
	1,3,4,5	II	0.023	0.000	0.000
	1,2,3,4,5	II	0.005	0.000	0.000
$\beta = (2, 0, 0, 4, 8)'$	1,4,5	Optimal	0.712	0.994	0.970
	1,3,5	I	0.000	0.000	0.000
	1,2,5	I	0.000	0.000	0.000
	1,2,4,5	II	0.133	0.003	0.016
	1,3,4,5	II	0.131	0.003	0.014
	1,2,3,4,5	II	0.024	0.000	0.000
$\beta = (2, 9, 0, 4, 8)'$	1,4,5	I	0.000	0.000	0.000
	1,2,4,5	Optimal	0.837	0.996	0.996
	1,3,4,5	I	0.000	0.000	0.000
	1,2,3,4,5	II	0.163	0.004	0.004
$\beta = (2, 9, 6, 4, 8)'$	1,2,3,5	I	0.000	0.000	0.000
	1,2,4,5	I	0.000	0.000	0.000
	1,3,4,5	I	0.000	0.000	0.000
	1,2,3,4,5	Optimal	1.000	1.000	1.000

1.1 Notations

In the above table,

- $\beta = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$ is the coefficients vector for the true model, where β_1 is the intercept term, i.e., $x_1 = 1$ is this framework
- "Model" is the variables included in the true model, for example, "1,4" represents that the true model is $y = \beta_1 x_1 + \beta_4 x_4$
- Category I: At least one nonzero component of β is not in β_α
- Category II: β_α contains all nonzero components of β

1.2 Summary

The result in the table above can be interpreted as follows:

- 1.In terms of the probability of selecting the optimal model, **the LOOBIC has the best performance among the three methods under consideration**, except for the case where the largest model (the model with all the $\beta_k \neq 0$) is the optimal model. The LOOBIC is slightly better than the BIC in all the cases.
- 2.The probability of selecting a model from Category I (incorrect model) is negligible for all three methods in all cases under consideration.
- 3.As expected, the LOOCV tends to select unnecessarily large models. **The probability of selecting the optimal model by applying the LOOCV can be very low** (e.g.,$\leq$ 0.6). Specifically, the more zero components the β has, the worse performance the LOOCV has. On the other hand, **the performance of the LOOBIC is much stable and better in the cases where the optimal model is not the largest model, even though the optimal model is sparse**.