ps5_writeup

December 7, 2023

1 BA 222 PS 5

Ethan Chang

1. How many variables does Lasso pick, and which ones?

- Lasso picked three variables: Tonnage, cabins, and passenger_density.

2. What is the fitted equation of the OLS model with those variables?

$$\text{crew} = 0.051188 + 0.011988 * \text{Tonnage} + 0.662980 * \text{cabins} + 0.025922 * \text{passenger\_density}$$

3. What is the R^2 of the OLS model with those variables? How does that compare to the OLS model with all of the variables and dummies?

- The R^2 of the OLS model with those variables is 0.913. This is slightly lower than the R^2 of the OLS model with all of the variables and dummies, which is 0.947.

4. Include a screenshot of your OLS model using the variables picked by Lasso.

[1]:
```
from IPython.display import Image
Image(filename='ols_lasso.png', width=500)
```

[1]:

OLS Regression Results

Dep. Variable:	crew	R-squared:	0.913
Model:	OLS	Adj. R-squared:	0.911
Method:	Least Squares	F-statistic:	537.0
Date:	Thu, 07 Dec 2023	Prob (F-statistic):	2.60e-81
Time:	13:21:41	Log-Likelihood:	-229.10
No. Observations:	158	AIC:	466.2
Df Residuals:	154	BIC:	478.5
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.0512	0.602	0.085	0.932	-1.138	1.240
Tonnage	0.0120	0.009	1.282	0.202	-0.006	0.030
cabins	0.6630	0.080	8.245	0.000	0.504	0.822
passenger_density	0.0259	0.013	1.970	0.051	-7.83e-05	0.052

Omnibus:	119.832	Durbin-Watson:	1.752
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1407.645
Skew:	2.626	Prob(JB):	2.16e-306
Kurtosis:	16.647	Cond. No.	645.

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.