

## Analysis

|                    | ols      | ridge    | lasso    | pcr      | plsr     |
|--------------------|----------|----------|----------|----------|----------|
| (Intercept)        | 0.00000  | 0.00000  | 0.00000  | 0.00000  | 0.00000  |
| Income             | -0.59817 | -0.56871 | -0.55166 | -0.59887 | -0.59817 |
| Limit              | 0.95844  | 0.71866  | 0.92505  | 0.67141  | 0.95844  |
| Rating             | 0.38248  | 0.59306  | 0.36787  | 0.67064  | 0.38248  |
| Cards              | 0.05286  | 0.04425  | 0.04500  | 0.04043  | 0.05286  |
| Age                | -0.02303 | -0.02538 | -0.01666 | -0.02327 | -0.02303 |
| Education          | -0.00747 | -0.00588 | 0.00000  | -0.00600 | -0.00747 |
| GenderFemale       | -0.01159 | -0.01068 | 0.00000  | -0.01165 | -0.01159 |
| StudentYes         | 0.27815  | 0.27318  | 0.26681  | 0.27636  | 0.27815  |
| MarriedYes         | -0.00905 | -0.01103 | 0.00000  | -0.01116 | -0.00905 |
| EthnicityAsian     | 0.01595  | 0.01638  | 0.00000  | 0.01741  | 0.01595  |
| EthnicityCaucasian | 0.01101  | 0.01101  | 0.00000  | 0.01119  | 0.01101  |

Table 1: Regression Coefficients for 5 Regression Methods

Comparing the regression coefficients of five regression methods, we can see that the coefficients of Income, Cards, Age, Education, GenderFemale, StudentYes, MarriedYes, EthnicityAsian and EthnicityCaucasian are relatively close for each regression method while the coefficients of Limit and Rating are relatively differed from each other. We can also see that the coefficients of Education, GenderFemale, MarriedYes, EthnicityAsian, and EthnicityCaucasian are very small and close to zero.

First, comparing the ols regression coefficients with the ridge regression coefficients, we can see that the ridge regression coefficients are generally smaller than the ols regression coefficients. This is because ridge regression method shrinks the coefficients of predictor variables and makes the coefficients closer to the true ones.

Second, comparing the lasso regression coefficients with the ridge regression coefficients, we can see that the lasso regression has several coefficients as zero. This is because lasso performs variable selection and yields model that involves only a subset of the variables. Lasso zeros out the coefficients of collinear variables. The advantage of lasso regression model over ridge regression model is that lasso regression method does both the parameter shrinkage and the variable selection and ridge regression will include all predictors in the final model and create a challenge in model interpretation.

|       | MSE     |
|-------|---------|
| ols   | 0.05179 |
| ridge | 0.05259 |
| lasso | 0.05154 |
| pcr   | 0.05200 |
| plsr  | 0.05179 |

Table 2: MSE of 5 Regression Methods