Disclaimer: NOTE: ***These materials are ONLY for revision purposes and does not in any way violate Fiverr Terms of Service***

-revision materials for linear regression

Given the linear regression equation:

$$y = \beta_0 + \beta_1 x + \epsilon$$

Assuming we have $\{(x_i, y_i), i = 1, \dots n\}$

To derive the least squares we do the following

Let

x_i	y_i
1	1
2	2
3	3
1	1
1	1

$$\Rightarrow y = \beta_1 x + (\beta_0 + \epsilon)$$

Given the linear regression equation:

$$y = \beta_0 + \beta_1 x + \epsilon$$

Assuming we have $\{(x_i, y_i), i = 1, \dots n\}$

To derive the least squares we do the following

Let

x_i	y_i
1	1
2	2
3	3
1	1
1	1

$$\Rightarrow y = \beta_1 x + (\beta_0 + \epsilon)$$

$$= \sum_{i=1}^{n} \left(y_i - \beta_0 - \beta_1 x - \epsilon - \sum_{j=1}^{p} \beta_j x_{ij}\right)$$

$$Y = \beta_0 + \beta_1 x + \dots \beta_p x_p + \epsilon$$

$$\Rightarrow = \sum^{n} (y_i - \beta_0 - \beta_1 x - \epsilon)^2$$

Disclaimer: NOTE: ***These materials are ONLY for revision purposes and does not in any way violate Fiverr Terms of Service***

-revision materials for linear regression

i	x_i	y_i
	1	1
	2	2
	3	3
	1	1
	1	1

$$\Rightarrow y = \beta_1 x + (\beta_0 + \varepsilon)$$

$$= \sum_{i=1}^{n} \left(y_i - \beta_0 - \beta_1 x - \varepsilon - \sum_{j=1}^{p} \beta_j x_{ij}\right)^2$$

$$Y = \beta_0 + \beta_1 x + \dots \beta_p x_p + \varepsilon$$

$$\Rightarrow = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x - \varepsilon)^2$$

(ii) Including a ridge penalty we have

$$\sum_{n=1}^{n} (y_i - \beta_0 - \beta_1 x - \varepsilon) + \lambda \sum_{i=1}^{p} \beta_i^2$$

Last term is the ridge penalty we are adding into a function

$$\beta_1^2 + \beta_0^2 + \varepsilon^2 \le t$$

$$= \sum_{i=1}^{n} \left(y_i - \beta_0 - \beta_1 x - \varepsilon - \sum_{j=1}^{p} \beta_j x_{ij}\right)^2$$

$$Y = \beta_0 + \beta_1 x + \dots \beta_p x_p + \varepsilon$$

$$\Rightarrow = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x - \varepsilon)^2$$

(ii) Including a ridge penalty we have

$$\sum_{n=1}^{n} (y_i - \beta_0 - \beta_1 x - \varepsilon) + \lambda \sum_{i=1}^{p} \beta_i^2$$

Last term is the ridge penalty we are adding into a function

$$\beta_1^2 + \beta_0^2 + \varepsilon^2 \le t$$

$$= \sum_{n=1}^{n} (y_i - \beta_0 - \beta_1 x - \varepsilon) - \sum_{j=1}^{\infty} \beta_j x_{ij})^2$$

$$= \sum_{n=1}^{\infty} (y_i - \beta_0 - \beta_1 x - \varepsilon)^2 + \lambda^2 \sum_{j=1}^{p} \beta_j^2$$

In General we have:

Disclaimer: NOTE: ***These materials are ONLY for revision purposes and does not in any way violate Fiverr Terms of Service***
-revision materials for linear regression

$$\Rightarrow = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x - \varepsilon)^2$$

(ii) Including a ridge penalty we have

$$\sum_{n=1}^{n} (y_i - \beta_0 - \beta_1 x - \varepsilon) + \lambda \sum_{i=1}^{p} \beta_i^2$$

Last term is the ridge penalty we are adding into a function.

$$\beta_1^2 + \beta_0^2 + \varepsilon^2 \leq t$$

$$= \sum_{n=1}^{n} (y_i - \beta_0 - \beta_1 x - \varepsilon) - \sum_{j=1}^{\infty} \beta_j x_{ij})^2$$

$$= \sum_{n=1}^{\infty} (y_i - \beta_0 - \beta_1 x - q)^2 + \lambda^2 \sum_{j=1}^{p} \beta_j^2$$

In General we have;

$$h_\theta(x_i) = \beta_0 + \beta_1 x_1 + \varepsilon + \lambda \sum_{j=1}^{p} \beta_i^2$$

(c) Assuming we extend the model to

$$y = \beta_0 + \beta_1 x + \beta_2 x + \varepsilon$$

Where $a = x!$

Reason:

- The estimator implies that among the possible unbiased estimators, the OLS achieves the minimum variance which gets a non-unique value for the OLS estimator.

- The obtained estimates are highly unstable despite being un[illegible] thus making the solution to have a non-unique value.

Disclaimer: NOTE: ***These materials are ONLY for revision purposes and does not in any way violate Fiverr Terms of Service***
-revision materials for linear regression

(c) Assuming we extend the model to

$$y = \beta_0 + \beta_1 x + \beta_2 z + \varepsilon$$

where $z = x!$

Reasons

- The estimator implies that among the possible unbiased estimators, the OLS achieves the minimum variance which gets a non-unique value for the OLS estimator.

- The obtained estimates are highly unstable despite being unbiased thus making the solution to have a non-unique value.

(d) Now adding penalties to both β_1 and β_2

$$= \sum_{i=1}^{n} \left(y_i - \beta_0 x - \sum_{j=1}^{p} \beta_j x_{ij}\right)^2 + \lambda\left[\sum \beta_1^2 x + \sum \beta_2^2 z\right]$$

$$y = \beta_0 + \beta_1 x + \beta_2 z + \varepsilon$$

where $z = x!$

Reasons

- The estimator implies that among the possible unbiased estimators, the OLS achieves the minimum variance which gets a non-unique value for the OLS estimator.

- The obtained estimates are highly unstable despite being unbiased thus making the solution to have a non-unique value.

(d) Now adding penalties to both β_1 and β_2

$$= \sum_{i=1}^{n} \left(y_i - \beta_0 x - \sum_{j=1}^{p} \beta_j x_{ij}\right)^2 + \lambda\left[\sum \beta_1^2 x + \sum \beta_2^2 z\right]$$

$$= \sum_{i=1}^{n} y_i - \beta_0 x + \lambda \sum_{i=1}^{n} (\beta_1^2 + \beta_2^2) - \left(\sum_{j=1}^{p} \beta_j x_{ij}\right)^2$$

$$= \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x - \beta_2 x - \varepsilon) + \lambda \sum_{j=1}^{p} \beta_j^2$$

Disclaimer: NOTE: ***These materials are ONLY for revision purposes and does not in any way violate Fiverr Terms of Service***

-revision materials for linear regression

- The estimator implies that among the ... estimators, the OLS achieves the minimum variance which gets a non-unique value for the OLS estimator.

- The obtained estimates are highly unstable despite being unbiased thus making the solution to have a non-unique value.

(d) Now adding penalties to both B_1 and B_2

$$= \sum_{i=1}^{n} \left(y_i - B_0 x - \sum_{j=1}^{p} B_j x_{ij}\right)^2 + \lambda \left[\sum B_1^2 x + \sum B_2^2 x\right]$$

$$= \sum_{i=1}^{n} y_i - B_0 x + \lambda \sum_{i=1}^{n} (B_1^2 + B_2^2) - \left(\sum_{j=1}^{p} B_j x_{ij}\right)^2$$

$$= \sum_{i=1}^{n} (y_i - B_0 - B_1 x - B_2 x - \varepsilon) + \lambda \sum_{j=1}^{p} B_j^2$$

e) Discussion

- These estimates and predictors are in the different ~~set~~ scale parametrization.

(d) Now adding penalties to both B_1 and B_2

$$= \sum_{i=1}^{n} \left(y_i - B_0 x - \sum_{j=1}^{p} B_j x_{ij}\right)^2 + \lambda \left[\sum B_1^2 x + \sum B_2^2 x\right]$$

$$= \sum_{i=1}^{n} y_i - B_0 x + \lambda \sum_{i=1}^{n} (B_1^2 + B_2^2) - \left(\sum_{j=1}^{p} B_j x_{ij}\right)^2$$

$$= \sum_{i=1}^{n} (y_i - B_0 - B_1 x - B_2 x - \varepsilon) + \lambda \sum_{j=1}^{p} B_j^2$$

e) Discussion

- These estimates and predictors are in the different ~~set~~ scale parametrization.

- This implies that the penalization $\lambda \sum_{i=1}^{p} B_i^2$ would have different impact in the predictors.
- Thus we always make calculations with use of standardized version of the obtained predictors.
- So in particular when there is a high collinearity (high corr... between the predictors) or the ...

Disclaimer: NOTE: ***These materials are ONLY for revision purposes and does not in any way violate Fiverr Terms of Service***

-revision materials for linear regression

$$= \sum_{i=1}^{n} y_i - \beta_0 x + \lambda \sum_{j=1}^{n} (\beta_1^2 + \beta_2^2) - \left(\sum_{j=1}^{p} \beta_j x_{ij}\right)^2$$

$$= \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x - \beta_2 x - \varepsilon) + \lambda \sum_{j=1}^{p} \beta_j^2$$

e) Discussion

- These estimates and predictors are in the different ~~set~~ scale parametrization.
- This implies that the penalization $\lambda \sum_{j=1}^{p} \beta_j^2$ would have different impact in the predictors.
- Thus we always make calculations with use of standardized version of the obtained predictors.
- So in particular when there is a high collinearity (high corr. between the predictors) or the number of predictors (p) is of similar magnitude to the number of observation (n).
- Note that λ is a tuning parameter that is controlling the [illegible] of penalisation.