FREE ONLINE EDUCATION swayam शिक्षित भारत, उन्नत भारत

NPTEL

MHRD

Introduction to Machine Learning

- Prof. Balaraman Ravindran | IIT Madras

Problem Solving Session (Week-2)

Shreya Bansal

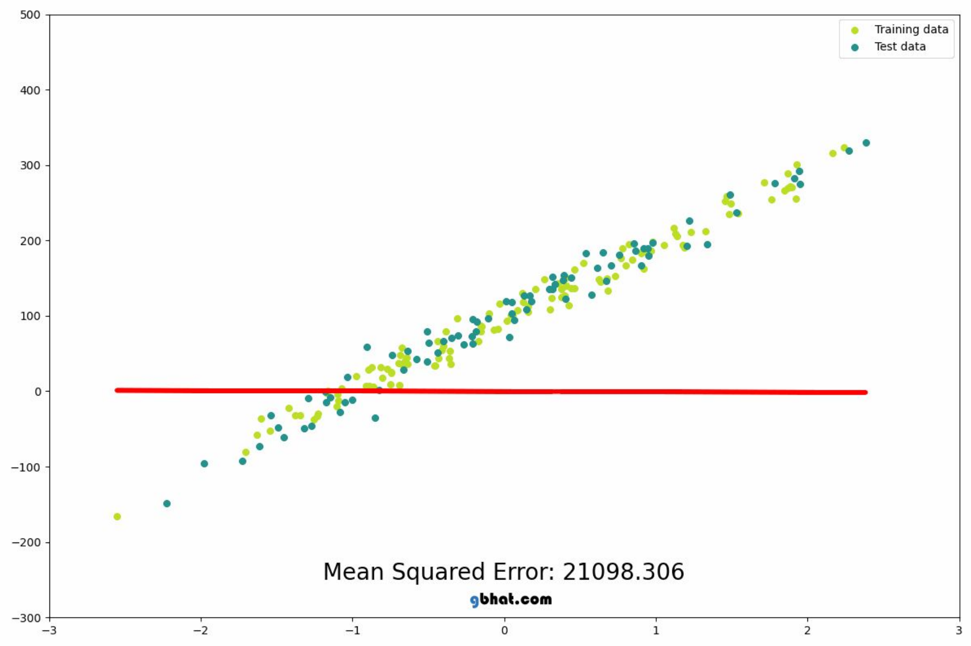
PMRF PhD Scholar
IIT Ropar

Week-2 Contents

1. Linear Regression
2. Multiple Regression
3. Subset Selection
4. Ridge and Lasso Regression
5. Principal Component Regression
6. Partial Least Square

Linear Regression

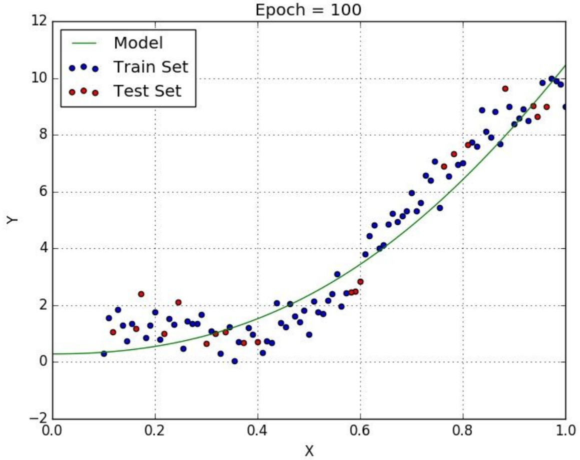
- Predicts a dependent variable using one independent variable
- Equation: $y = \beta_0 + \beta_1 x + \varepsilon$
- Solved using Ordinary Least Squares (OLS)

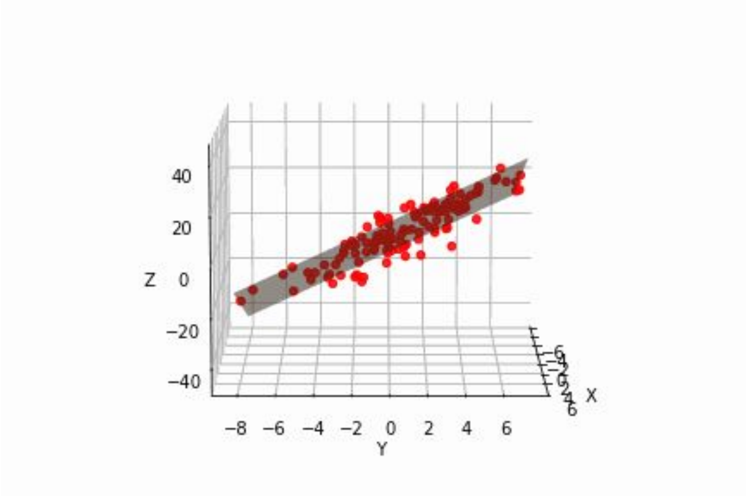


Source

Multiple Regression

- **Extends Linear Regression to multiple predictors**
- **Equation: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots +$ β□x□ $+ \varepsilon$**





Source: left, right

Regression

- Models the relationship between dependent and independent variables by fitting a linear equation:

  $$Y = X\beta + \varepsilon$$

  where:

  - Y is the output (dependent variable)

  - X is the input (independent variables)

  - β are the model parameters

  - ε is the error term

Least Squares Error Function

- The goal of linear regression is to minimize the squared error:
- Error =$(Y - X\beta)^2 = (Y - X\beta)^T (Y - X\beta)$ [As $P^2 = P^TP$]
- This represents the sum of squared differences between actual and predicted values.

Expanding the Error Function

- Using the matrix identity

$$(A - B)^T (A - B) = A^T A - A^T B - B^T A + B^T B$$

- $(Y - X\beta)^T (Y - X\beta)$ expands to: $Y^T Y - Y^T (X\beta) - (X\beta)^T Y + \beta^T X^T X \beta$
  - $Y^T (X\beta)$ is scalar so equal to its transpose

$$Y^T Y - (X\beta)^T Y - (X\beta)^T Y + \beta^T X^T X \beta$$

$$Y^T Y - 2\beta^T X^T Y + \beta^T X^T X \beta$$

- This is a quadratic function in β.

Finding Optimal β

- To minimize the error function, take its derivative w.r.t. β and set it to zero:

$$\partial/\partial\beta\ (Y^T Y - 2\beta^T X^T Y + \beta^T X^T X \beta) = 0$$

- Solving for β gives:

$$\beta = (X^T X)^{-1} X^T Y$$

Sample Question: Regression

x1	x2	x3	y
2	3	5	10
4	2	1	8
3	4	2	11
5	1	3	9
1	5	4	7

Sample Question: Linear Regression

x1	y
2	10
4	8
3	11
5	9
1	7

→

X (5 x 2)	
1	2
1	4
1	3
1	5
1	1

Y (5 X 1)
10
8
11
9
7

Step-1: Define X and Y

Sample Question: Linear Regression

X^T (2 x 5)	1	1	1	1	1
	2	4	3	5	1

X (5 x 2)	
1	2
1	4
1	3
1	5
1	1

a (2 x 2)	5	15
	15	55

a^{-1} (2 x 2)	1.1	-0.3
	-0.3	0.1

Step-2: Calculate $a^{-1} = (X^T X)^{-1}$

Sample Question: Linear Regression

X^T (2 x 5)	1	1	1	1	1
	2	4	3	5	1

Y (5 X 1)
10
8
11
9
7

→

b (2 x 1)	45
	137

Step-3: Calculate $b = X^T Y$

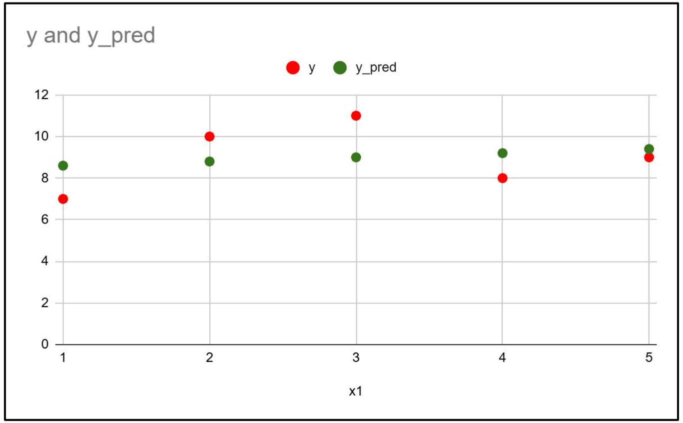
Sample Question: Linear Regression

a^{-1} (2 x 2)	1.1	-0.3
	-0.3	0.1

b (2 x 1)	45
	137

β (2 x 1)	8.4
	0.2

y= 8.4 + 0.2x



Step-4: Calculate $\beta = (X^T X)^{-1} X^T Y = a^{-1}b$

Sample Question: Regression

x1	x2	x3	y
2	3	5	10
4	2	1	8
3	4	2	11
5	1	3	9
1	5	4	7

Sample Question: Multiple Regression

X (5 x 4)			
1	2	3	5
1	4	2	1
1	3	4	2
1	5	1	3
1	1	5	4

Y (5 X 1)
10
8
11
9
7

Step-1: Define X and Y

Sample Question: Linear Regression

X^T (4 x 5)	1	1	1	1	1
	2	4	3	5	1
	3	2	4	1	5
	5	1	2	3	4

X (5 x 4)			
1	2	3	5
1	4	2	1
1	3	4	2
1	5	1	3
1	1	5	4

a (4 x 4)	5	15	15	15
	15	55	36	39
	15	36	55	48
	15	39	48	55

a^{-1} (4 x 4)	60.95	-9.1875	-7.5	-3.5625
	-9.1875	1.4218	1.125	30.51562
	-7.5	1.125	1	0.375
	-3.5625	0.5156	0.375	0.2968

Step-2: Calculate

$a^{-1} = (X^T X)^{-1}$

Sample Question: Linear Regression

X^T (4 x 5)	1	1	1	1	1
	2	4	3	5	1
	3	2	4	1	5
	5	1	2	3	4

b (4 x 1)	45
	137
	134
	135

Y (5 X 1)
10
8
11
9
7

Step-3: Calculate **b** = X^T Y

Sample Question: Linear Regression

a^{-1} (4 x 4)	60.95	-9.1875	-7.5	-3.5625
	-9.1875	1.4218	1.125	30.51562
	-7.5	1.125	1	0.375
	-3.5625	0.5156	0.375	0.2968

b (4 x 1)	45
	137
	134
	135

→

β (4 x 1)	-1.875
	1.718
	1.25
	0.6562

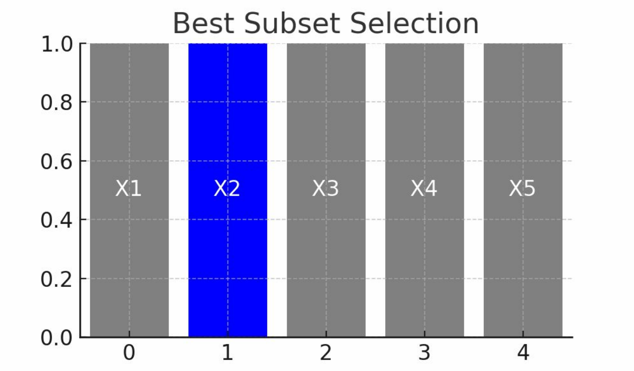
$$y = -1.875 + 1.718\, x_1 + 1.25\, x_2 + 0.6562\, x_3$$

Step-4: Calculate $\beta = (X^T X)^{-1} X^T Y = a^{-1}b$

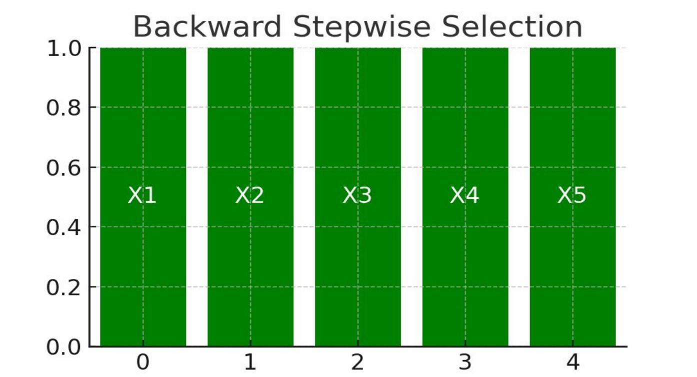
Sample Question: Regression (Output)

x1	x2	x3	y	y_pred	Error
2	3	5	10	8.592	1.9824
4	2	1	8	8.1532	0.023
3	4	2	11	9.5914	1.9841
5	1	3	9	9.9336	0.8716
1	5	4	7	8.7178	2.950
Average error = 1.562					

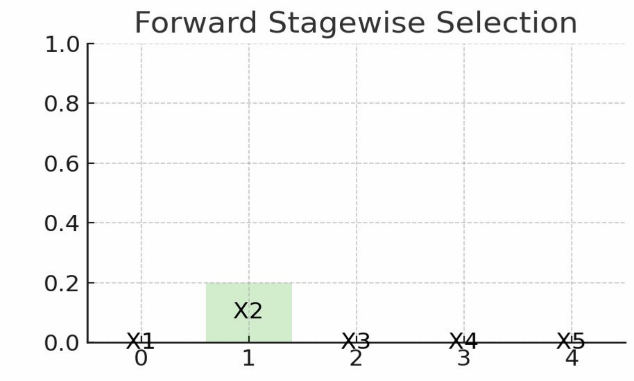
Subset Selection Methods



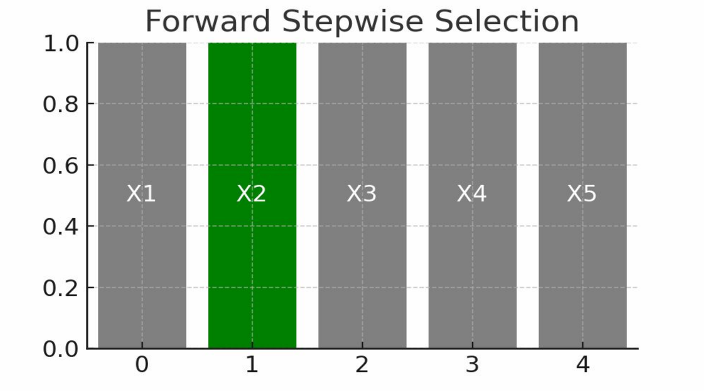
Best Subset Selection



Backward Stepwise Selection

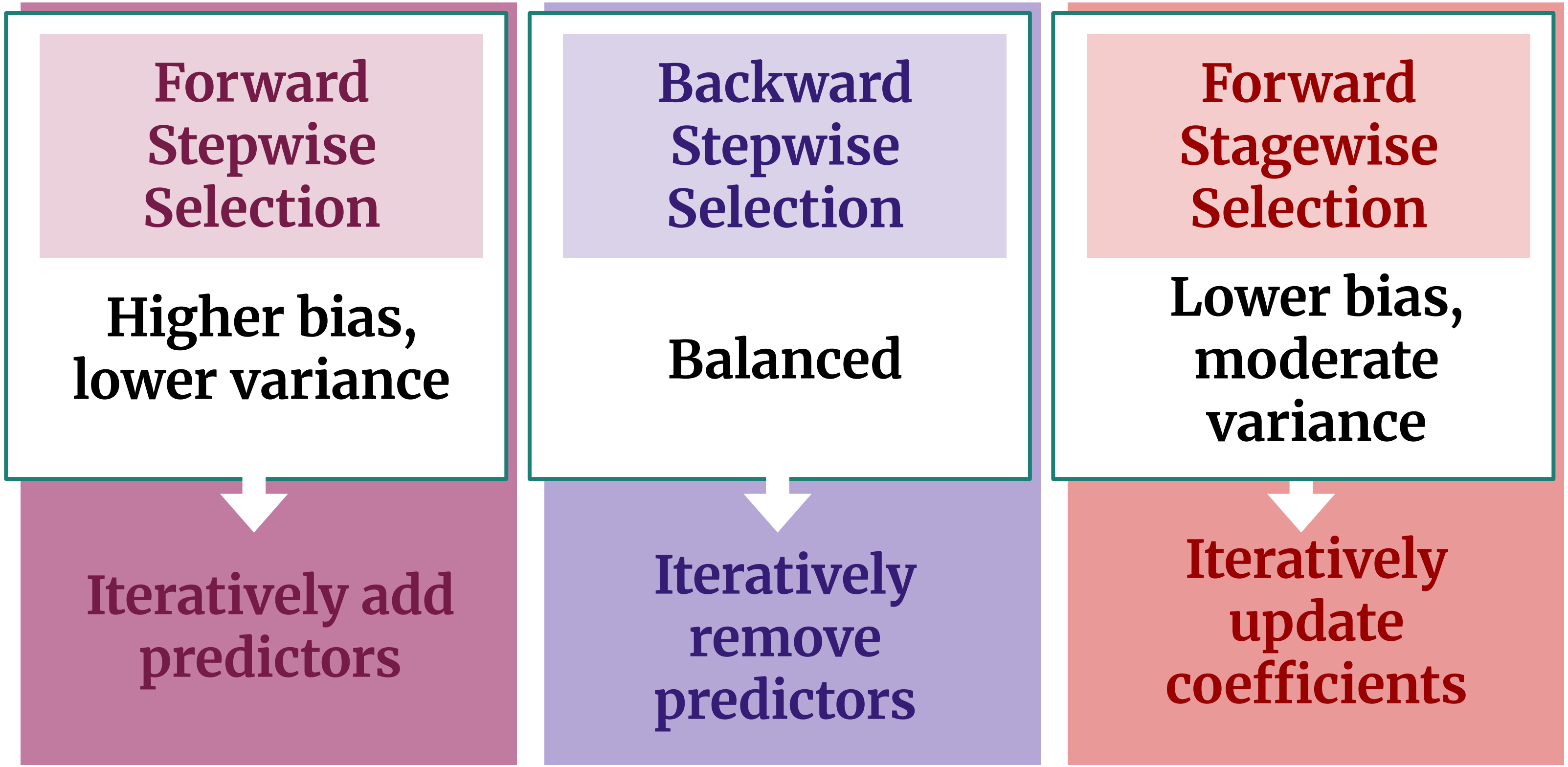


Forward Stagewise Selection



Forward Stepwise Selection

Subset Selection Methods



Forward Stepwise Selection

Higher bias, lower variance

Iteratively add predictors

Backward Stepwise Selection

Balanced

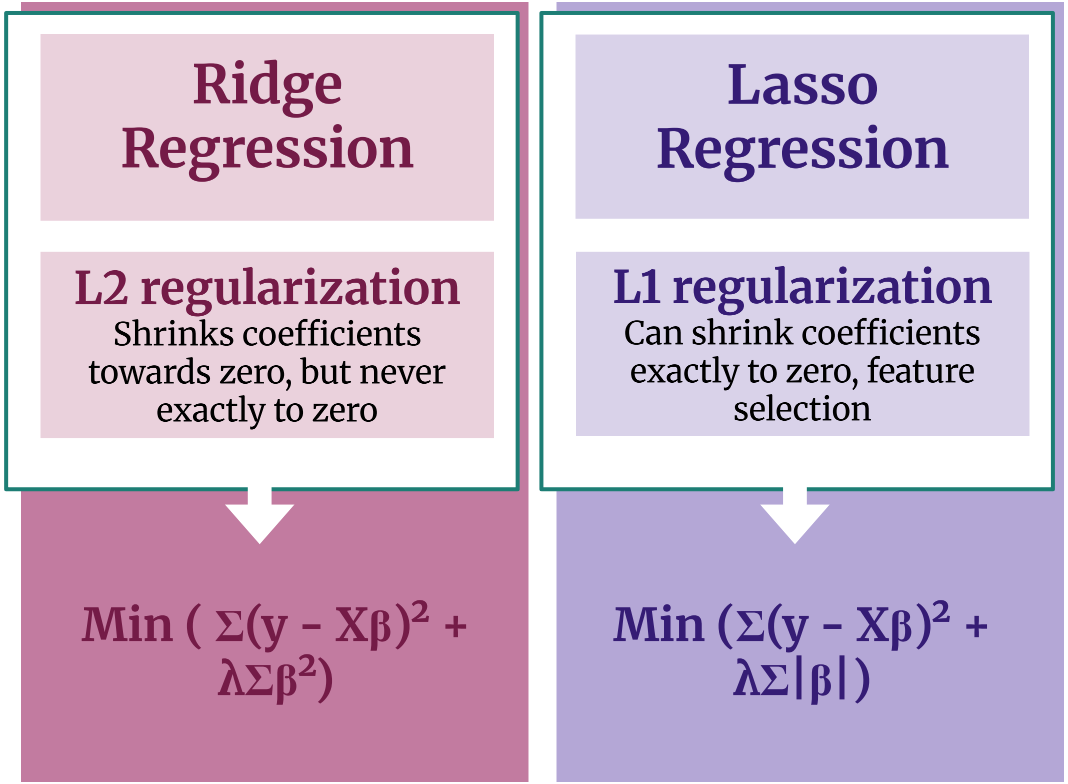
Iteratively remove predictors

Forward Stagewise Selection

Lower bias, moderate variance

Iteratively update coefficients

Ridge vs Lasso Regression



Ridge Regression

L2 regularization

Shrinks coefficients towards zero, but never exactly to zero

$\text{Min} \left(\Sigma(y - X\beta)^2 + \lambda\Sigma\beta^2 \right)$

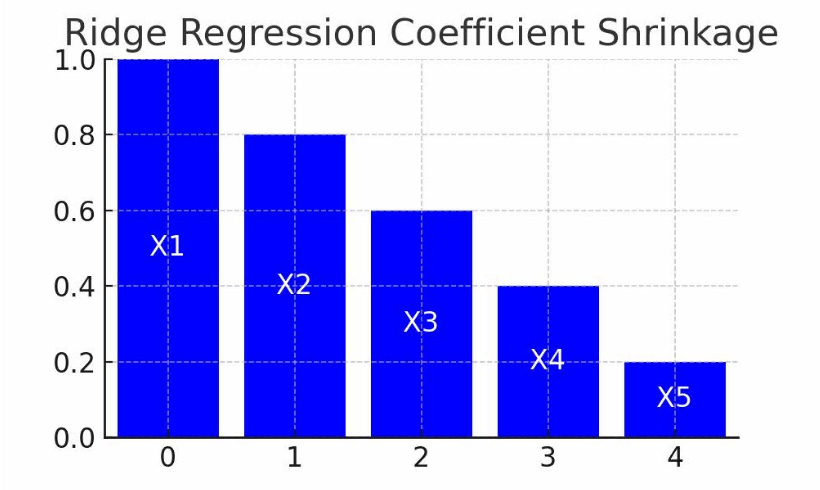
Lasso Regression

L1 regularization

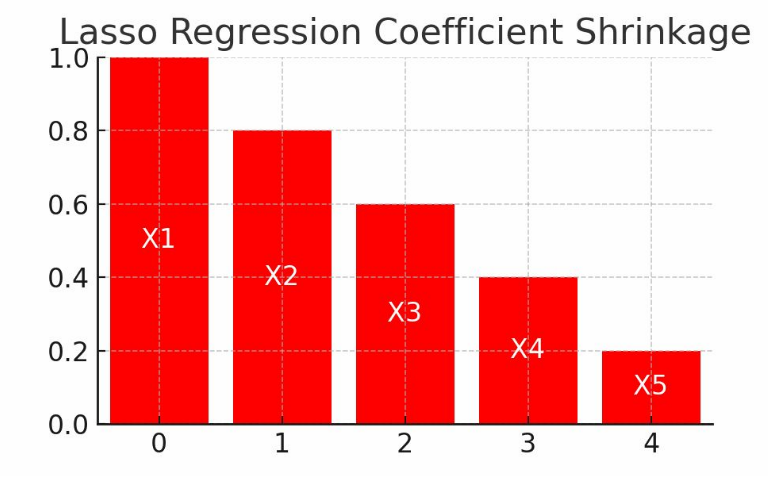
Can shrink coefficients exactly to zero, feature selection

$\text{Min} \left(\Sigma(y - X\beta)^2 + \lambda\Sigma|\beta| \right)$

Ridge vs Lasso Regression

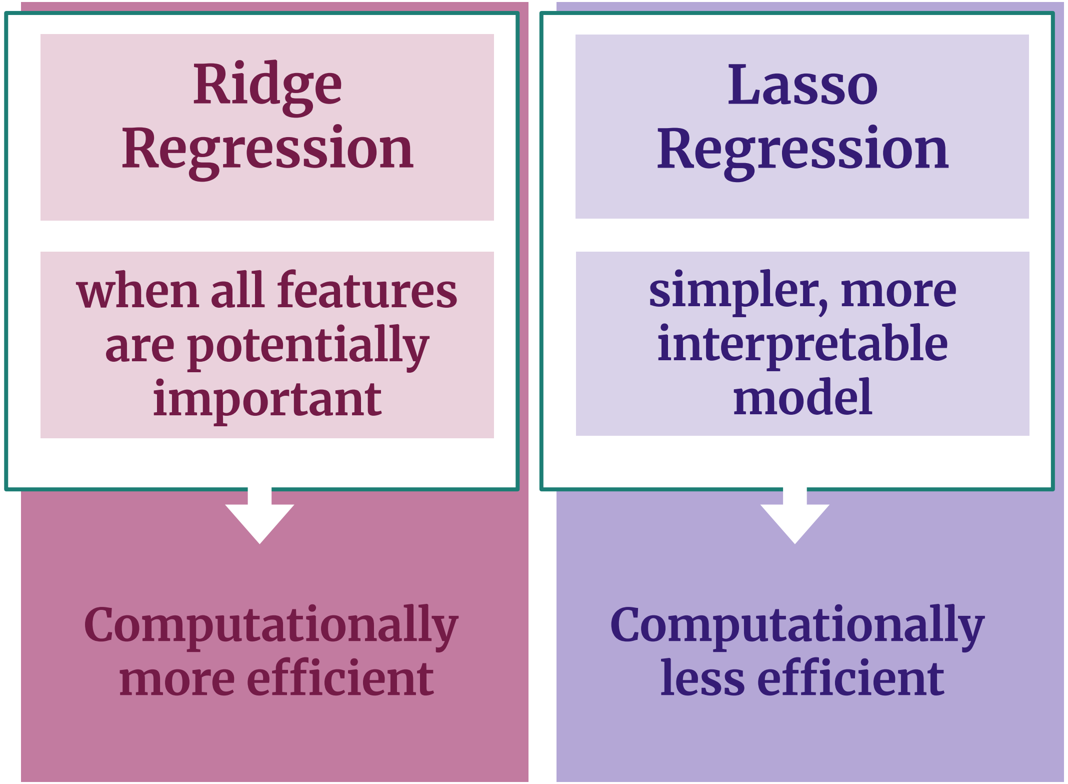


Ridge Regression Coefficient Shrinkage



Lasso Regression Coefficient Shrinkage

Choosing Between Ridge and Lasso



Ridge Regression

when all features are potentially important

Computationally more efficient

Lasso Regression

simpler, more interpretable model

Computationally less efficient

Sample Question: Ridge Regression

x1	x2	x3	y
2	3	5	10
4	2	1	8
3	4	2	11
5	1	3	9
1	5	4	7

Sample Question: Ridge Regression

X^T (4 x 5)	1	1	1	1	1
	2	4	3	5	1
	3	2	4	1	5
	5	1	2	3	4

X (5 x 4)			
1	2	3	5
1	4	2	1
1	3	4	2
1	5	1	3
1	1	5	4

$X^T X$ (4 x 4)	5	15	15	15
	15	55	36	39
	15	36	55	48
	15	39	48	55

I' (4 x 4)	0	0	0	0
	0	1	0	0
	0	0	1	0
	0	0	0	1

λ=1 (for simplicity)

a (4 x 4)	5	15	15	15
	15	56	36	39
	15	36	56	48
	15	39	48	56

a^{-1} (4 x 4) (Calculate)

Step-2: Calculate $a^{-1} = (X^T X + \lambda I')^{-1}$

Sample Question: Linear Regression

X^T (4 x 5)	1	1	1	1	1
	2	4	3	5	1
	3	2	4	1	5
	5	1	2	3	4

Y (5 X 1)
10
8
11
9
7

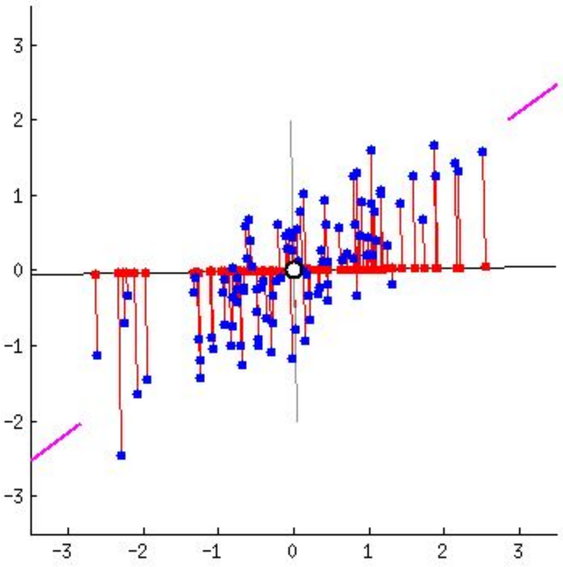
b (4 x 1)	45
	137
	134
	135

Step-3: Calculate $b = X^T Y$

Step-4: Calculate $\beta = (X^T X + \lambda I)^{-1} X^T Y = a^{-1}b$
(Calculate yourself)

Principal Component Regression (PCR)

- Uses PCA to reduce dimensionality before regression
- Addresses multicollinearity issues
- Regression is performed on principal components rather than original variables



Source

PCR Method

- **Standardize the data:** Center and scale both predictor (X) and response (Y)
- **Perform PCA on predictors:**
  - Compute the covariance matrix of X
  - Calculate eigenvectors and eigenvalues
  - Rank eigenvectors by descending eigenvalues
  - Select top k principal components (PCs)
- **Project data onto PCs:** Transform X into the new PC space.
- **Perform OLS regression:** Use the selected PCs as predictors for Y.
- **Transform coefficients back:** Convert PC coefficients to original variable space.

Sample Question: PC Regression

x1	x2	x3	y
2	3	5	10
4	2	1	8
3	4	2	11
5	1	3	9
1	5	4	7

Sample Question: PC Regression

Original Data	x1	x2	x3	y
	2	3	5	10
	4	2	1	8
	3	4	2	11
	5	1	3	9
	1	5	4	7
Mean	3	3	3	9
Std	1.58	1.58	1.58	1.58

Scaled Data			
z1	z2	z3	y
-0.63	0	1.26	4.42
0.63	-0.63	-1.26	3.16
0	0.63	-0.63	5.05
1.26	-1.26	0	3.79
-1.26	1.26	0.63	2.52

Step-1: Standardized data

Sample Question: PCR

X^T (3 x 5)	-0.63	0.63	0	1.26	-1.26
	0	-0.63	0.63	-1.26	1.26
	1.26	-1.26	-0.63	0	0.63

X (5 x 3)		
-0.63	0	1.26
0.63	-0.63	-1.26
0	0.63	-0.63
1.26	-1.26	0
-1.26	1.26	0.63

A (3 x 3)	1	−0.9	−0.6
	−0.9	1	0.3
	−0.6	0.3	1

Step-2: Calculate covariance matrix

$$A = (1/n-1)(X^T X)$$

$$n=5$$

Sample Question: Linear Regression

Step-3: Calculate Eigen value and eigen vector of A using solution of $|A-\lambda I|=0$

Solve:

$$-\lambda^3 + 3\lambda^2 - 1.74\lambda + 0.064=0$$

$$\lambda_1= 2.23,\ \lambda_2= 0.726 \text{ and } \lambda_3= 0.039$$

For eigenvector v, solve $(A-\lambda I)v=0$

Sample Question: Linear Regression

Step-4: Select top-K eigenvectors

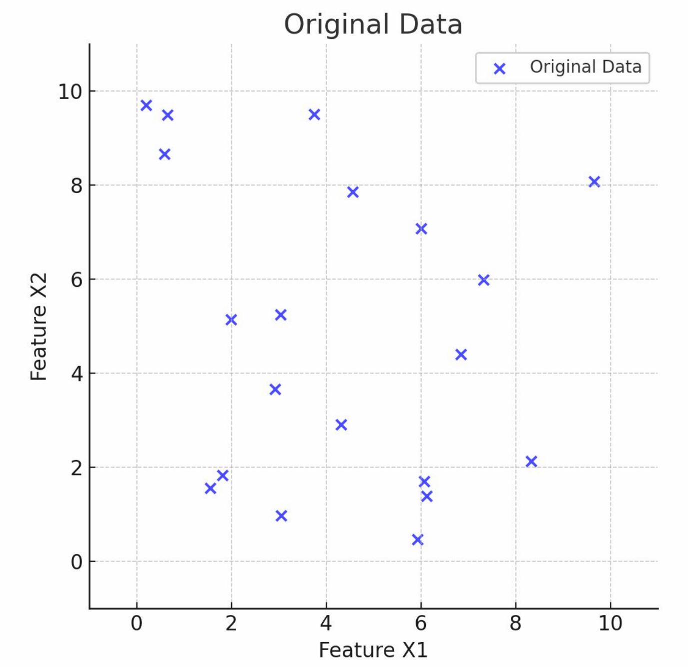
Step-5: Calculate $PC_i = X V_i$

Step-6: Apply Regression on PC

Step-7: Convert PC coefficients to original variable space using V_i

Partial Least Squares (PLS) Regression

- Similar to PCR but maximizes covariance between predictors and response
- Works well for high-dimensional data with correlated predictors



PLS Method

- Standardize the data (center and scale variables).
- Find components that maximize the covariance between X and Y spaces.
- Project the data onto these components.
- Perform regression using the projected data.

Assignment-2 (Cs-46- 2025) (Week-2)

Source

Question-1

01:00

In a linear regression model $y=\theta_0+\theta_1x_1+\theta_2x_2+...+\theta_px_p$, what is the purpose of adding an intercept term (θ_0)?

a) To increase the model’s complexity

b) To account for the effect of independent variables.

c) To adjust for the baseline level of the dependent variable when all predictors are zero.

d) To ensure the coefficients of the model are unbiased

Question-1- Correct answer

In a linear regression model $y=\theta_0+\theta_1x_1+\theta_2x_2+...+\theta_px_p$, what is the purpose of adding an intercept term (θ_0)?

a) To increase the model’s complexity
b) To account for the effect of independent variables.
c) To adjust for the baseline level of the dependent variable when all predictors are zero.
d) To ensure the coefficients of the model are unbiased

Correct options: (c)

Question-2

03:00

Which of the following is true about the cost function (objective function) used in linear regression?

a) It is non-convex.

b) It is always minimized at $\theta = 0$.

c) It measures the sum of squared differences between predicted and actual values.

d) It assumes the dependent variable is categorical.

Question-2 - Correct answer

Which of the following is true about the cost function (objective function) used in linear regression?

a) It is non-convex.

b) It is always minimized at θ = 0.

c) It measures the sum of squared differences between predicted and actual values.

d) It assumes the dependent variable is categorical.

Correct options: (c)

Question-3

01:00

Which of these would most likely indicate that Lasso regression is a better choice than Ridge regression?

a) All features are equally important
b) Features are highly correlated
c) Most features have small but non-zero impact
d) Only a few features are truly relevant

Question-3- Correct answer

Which of these would most likely indicate that Lasso regression is a better choice than Ridge regression?

a) All features are equally important

b) Features are highly correlated

c) Most features have small but non-zero impact

d) Only a few features are truly relevant

Correct options: (d)

Question-4

01:00

Which of the following conditions must hold for the least squares estimator in linear regression to be unbiased?

a) The independent variables must be normally distributed.

b) The relationship between predictors and the response must be non-linear.

c) The errors must have a mean of zero.

d) The sample size must be larger than the number of predictors.

Question-4 - Correct answer

Which of the following conditions must hold for the least squares estimator in linear regression to be unbiased?

a) The independent variables must be normally distributed.

b) The relationship between predictors and the response must be non-linear.

c) The errors must have a mean of zero.

d) The sample size must be larger than the number of predictors.

Correct options: (c)

Question-5

01:00

When performing linear regression, which of the following is most likely to cause overfitting?

a) Adding too many regularization terms.
b) Including irrelevant predictors in the model.
c) Increasing the sample size.
d) Using a smaller design matrix.

Question-5 - Correct answer

When performing linear regression, which of the following is most likely to cause overfitting?

a) **Adding too many regularization terms.**

b) **Including irrelevant predictors in the model.**

c) **Increasing the sample size.**

d) **Using a smaller design matrix.**

Correct options: (b)

Question-6

01:00

You have trained a complex regression model on a dataset. To reduce its complexity, you decide to apply Ridge regression, using a regularization parameter λ . How does the relationship between bias and variance change as λ becomes very large? Select the correct option

a) bias is low, variance is low.

b) bias is low, variance is high.

c) bias is high, variance is low.

d) bias is high, variance is high.

Question-6 - Correct answer

You have trained a complex regression model on a dataset. To reduce its complexity, you decide to apply Ridge regression, using a regularization parameterλ. How does the relationship between bias and variance change asλ becomes very large? Select the correct option

a) bias is low, variance is low.

b) bias is low, variance is high.

c) bias is high, variance is low.

d) bias is high, variance is high.

Correct options: (c)

Question-7

03:00

Given a training data set of 10,000 instances, with each input instance having 12 dimensions and each output instance having 3 dimensions, the dimensions of the design matrix used in applying linear regression to this data is

a) 10000 × 12

b) 10003 × 12

c) 10000 × 13

d) 10000 × 15

Question-7 - Correct answer

Given a training data set of 10,000 instances, with each input instance having 12 dimensions and each output instance having 3 dimensions, the dimensions of the design matrix used in applying linear regression to this data is

a) 10000 × 12

b) 10003 × 12

c) 10000 × 13

d) 10000 × 15

Correct options: (c)

Question-8

01:00

The linear regression model $y=a_0+a_1x_1+a_2x_2+...+a_px_p$ is to be fitted to a set of N training data points having P attributes each. Let X be N * (p+1) vectors of input values (augmented by 1‘s), Y be N x 1 vector of target values, and θ be (p+1)×1 vector of parameter values $(a_0,a_1,a_2,...,a_p)$. If the sum squared error is minimized for obtaining the optimal regression model, which of the following equation holds?

(a) $X^TX = XY$

(b) $X\theta = X^TY$

(c) $X^TX\theta = Y$

(d) $X^TX\theta = X^TY$

Question-8- Correct answer

The linear regression model $y=a_0+a_1x_1+a_2x_2+...+a_px_p$ is to be fitted to a set of N training data points having P attributes each. Let X be N * (p+1) vectors of input values (augmented by 1's), Y be N x 1 vector of target values, and θ be (p+1)×1 vector of parameter values $(a_0,a_1,a_2,...,a_p)$. If the sum squared error is minimized for obtaining the optimal regression model, which of the following equation holds?

(a) $X^TX = XY$

(b) $X\theta = X^TY$

(c) $X^TX\theta = Y$

(d) $X^TX\theta = X^TY$

Correct options: (d)

Question-9

01:00

Which of the following scenarios is most appropriate for using Partial Least Squares (PLS) regression instead of ordinary least squares (OLS)?

(a) When the predictors are uncorrelated and the number of samples is much larger than the number of predictors.

(b) When there is significant multicollinearity among predictors or the number of predictors exceeds the number of samples.

(c) When the response variable is categorical and the predictors are highly non-linear.

(d) When the primary goal is to interpret the relationship between predictors and response, rather than prediction accuracy.

Question-9- Correct answer

Which of the following scenarios is most appropriate for using Partial Least Squares (PLS) regression instead of ordinary least squares (OLS)?

(a) When the predictors are uncorrelated and the number of samples is much larger than the number of predictors.

(b) When there is significant multicollinearity among predictors or the number of predictors exceeds the number of samples.

(c) When the response variable is categorical and the predictors are highly non-linear.

(d) When the primary goal is to interpret the relationship between predictors and response, rather than prediction accuracy.

Correct options: (b)

Question-10

01:00

Consider forward selection, backward selection and best subset selection with respect to the same data set. Which of the following is true?

(a) Best subset selection can be computationally more expensive than forward selection

(b) Forward selection and backward selection always lead to the same result

(c) Best subset selection can be computationally less expensive than backward selection

(d) Best subset selection and forward selection are computationally equally expensive

(e) Both (b) and (d)

Question-10- Correct answer

Consider forward selection, backward selection and best subset selection with respect to the same data set. Which of the following is true?

(a) Best subset selection can be computationally more expensive than forward selection

(b) Forward selection and backward selection always lead to the same result

(c) Best subset selection can be computationally less expensive than backward selection

(d) Best subset selection and forward selection are computationally equally expensive

(e) Both (b) and (d)

Correct options: (a)

Assignment-2 (Cs-101- 2024) (Week-2)

Source

Question-1

01:00

State True or False: Typically, linear regression tend to underperform compared to k-nearest neighbor algorithms when dealing with high-dimensional input spaces.

a) True

b) False

Question-1- Correct answer

State True or False: Typically, linear regression tend to underperform compared to k-nearest neighbor algorithms when dealing with high-dimensional input spaces.

a) True

b) False

Correct options: (b)

Question-2

03:00

Given the following dataset, find the uni-variate regression function that best fits the dataset.

a) $f(x) = 1x +4$

b) $f(x) = 1x +5$

c) $f(x) = 1.5x +3$

d) $f(x) = 2x +1$

X	Y
2	5.5
3	6.5
4	9
10	18.5

Question-2 - Correct answer

Given the following dataset, find the uni-variate regression function that best fits the dataset.

a) $f(x) = 1x +4$

b) $f(x) = 1x +5$

c) $f(x) = 1.5x +3$

d) $f(x) = 2x +1$

Correct options: (c)

Question-3

01:00

Given a training data set of 500 instances, with each input instance having 6 dimensions and each output being a scalar value, the dimensions of the design matrix used in applying linear regression to this data is

a) 500 x 6

b) 500 x 7

c) 500 x 6^2

d) None

Question-3- Correct answer

Given a training data set of 500 instances, with each input instance having 6 dimensions and each output being a scalar value, the dimensions of the design matrix used in applying linear regression to this data is

a) 500 x 6
b) 500 x 7
c) 500 x 62
d) None

Correct options: (b)

Question-4

01:00

Assertion A: Binary encoding is usually preferred over One-hot encoding to represent categorical data (eg. colors, gender etc)

Reason R: Binary encoding is more memory efficient when compared to One-hot encoding

a) Both A and R are true and R is the correct explanation of A

b) Both A and R are true but R is not the correct explanation of A

c) A is true but R is false

d) A is false but R is true

Question-4 - Correct answer

Assertion A: Binary encoding is usually preferred over One-hot encoding to represent categorical data (eg. colors, gender etc)

Reason R: Binary encoding is more memory efficient when compared to One-hot encoding

a) Both A and R are true and R is the correct explanation of A

b) Both A and R are true but R is not the correct explanation of A

c) A is true but R is false

d) A is false but R is true

Correct options: (d)

Question-5

01:00

Select the TRUE statement

a) Subset selection methods are more likely to improve test error by only focussing on the most important features and by reducing variance in the fit.

b) Subset selection methods are more likely to improve train error by only focussing on the most important features and by reducing variance in the fit.

c) Subset selection methods are more likely to improve both test and train error by focussing on the most important features and by reducing variance in the fit.

d) Subset selection methods don’t help in performance gain in any way.

Question-5 - Correct answer

Select the TRUE statement

a) Subset selection methods are more likely to improve test error by only focussing on the most important features and by reducing variance in the fit.

b) Subset selection methods are more likely to improve train error by only focussing on the most important features and by reducing variance in the fit.

c) Subset selection methods are more likely to improve both test and train error by focussing on the most important features and by reducing variance in the fit.

d) Subset selection methods don’t help in performance gain in any way.

Correct options: (a)

Question-6

01:00

Rank the 3 subset selection methods in terms of computational efficiency

a) Forward stepwise selection, best subset selection, and forward stagewise regression.

b) Forward stepwise selection, forward stagewise regression and best subset selection.

c) Best subset selection, forward stagewise regression and forward stepwise selection.

d) Best subset selection, forward stepwise selection and forward stagewise regression.

Question-6 - Correct answer

Rank the 3 subset selection methods in terms of computational efficiency:

a) Forward stepwise selection, best subset selection, and forward stagewise regression.

b) Forward stepwise selection, forward stagewise regression and best subset selection.

c) Best subset selection, forward stagewise regression and forward stepwise selection.

d) Best subset selection, forward stepwise selection and forward stagewise regression.

Correct options: (b)

Question-7

03:00

Choose the TRUE statements from the following: (Multiple correct choice)

a) Ridge regression since it reduces the coefficients of all variables, makes the final fit a lot more interpretable.

b) Lasso regression since it doesn’t deal with a squared power is easier to optimize than ridge regression.

c) Ridge regression has a more stable optimization than lasso regression.

d) Lasso regression is better suited for interpretability than ridge regression.

Question-7 - Correct answer

Choose the TRUE statements from the following:

a) Ridge regression since it reduces the coefficients of all variables, makes the final fit a lot more interpretable.
b) Lasso regression since it doesn't deal with a squared power is easier to optimize than ridge regression.
c) Ridge regression has a more stable optimization than lasso regression.
d) Lasso regression is better suited for interpretability than ridge regression.

Correct options: (c) (d)

Question-8

01:00

Which of the following statements are TRUE? Let x_i be the i-th datapoint in a dataset of N points. Let v represent the first principal component of the dataset. (Multiple answer questions)

a) $v = \arg\max \sum_{i=1}^{N} (v^T x_i)^2 \text{ s.t. } |v| = 1|$

b) $v = \arg\min \sum_{i=1}^{N} (v^T x_i)^2 s.t. |v| = 1$

c) Scaling at the start of performing PCA is done just for better numerical stability and computational benefits but plays no role in determining the final principal components of a dataset.

d) The resultant vectors obtained when performing PCA on a dataset can vary based on the scale of the dataset.

Question-8- Correct answer

Which of the following statements are TRUE? Let x_i be the i-th datapoint in a dataset of N points. Let v represent the first principal component of the dataset. (Multiple answer questions)

a) $v = \arg\max \sum_{i=1}^{N} \left(v^T x_i\right)^2 \text{ s.t. } |v| = 1|$

b) $v = \arg\min \sum_{i=1}^{N} \left(v^T x_i\right)^2 s.t. |v| = 1$

c) Scaling at the start of performing PCA is done just for better numerical stability and computational benefits but plays no role in determining the final principal components of a dataset.

d) The resultant vectors obtained when performing PCA on a dataset can vary based on the scale of the dataset.

Correct options: (a) (d)

THANK YOU

Suggestions and Feedback

Next Session:

Tuesday:
10-Aug-2025
3:00 - 5:00 PM