

**Predictive Analytics - Regression and Classification**  
**Prof. Sourish Das**  
**Department of Mathematics**  
**Chennai Mathematical Institute**

**Lecture - 24**  
**III-Posed Problem and Regularisation, LASSO and Ridge**

Welcome to the part B of lecture 7 we are going to now talk about class of Ill-Posed Problems.


(Refer Slide Time: 00:21)


### Class of Ill-Posed Problems


▶ A class of problem is known as ill-posed problem - if either of the following feature exists

- ✓1. Unique solution does not exist
- ✓2. Unique solution exists - but computationally not feasible
- ✓3. Unique solution exists - but unreliable

- 1 Problem of variable selection in "large  $p$ , small  $n$ " setup considered as ill-posed problems
- 2 Problem of variable selection in large  $p$  is considered as ill-posed problems for model complexity.
- 3 Problem of multicollinearity also considered ill-posed problems.







A class of problem is known as ill-posed if either they fall into either of the 3 status or 3 features if you see 1 is there will be a unique solution does not exist for a problem, 2nd is unique solution exists - but computationally not feasible, you have a unique solution but

computationally it is not feasible computationally finding a solution is impossible and the 3rd possibility is unique solution exists, but the solution is highly unreliable.

(Refer Slide Time: 01:33)

CLASS OF ILL-POSED PROBLEMS

- ▶ A class of problem is known as ill-posed problem - if either of the following feature exists
  - ✓ 1. Unique solution does not exist
  - ✓ 2. Unique solution exists - but computationally not feasible
  - ✓ 3. Unique solution exists - but unreliable

Ex 1 Problem of variable selection in "large  $p$ , small  $n$ " setup considered as ill-posed problems  $X_{n \times p}$

Ex 2 Problem of variable selection in large  $p$  is considered as ill-posed problems for model complexity.

Ex 3 Problem of multicollinearity also considered ill-posed problems.

cm<sub>i</sub>

NPTEL

Now, I will give you 3 examples of all 3 cases, first example is this is the example one problem of variable selection for large  $p$  and small  $n$ . If you have a design matrix with  $n$  cross  $p$  and the number of column is more than number of samples then the number of feature is more than the number of samples, then what happens is  $X$  transpose  $X$  is not invertible as we have seen in the previous part of this lecture.



Since we do not have any unique solution, this considered as a ill-posed problem. Second example is the problem of variable selection particularly for very large  $p$ . So, you have  $p$  is so large that the model complexity become too high and searching through all the model complexity is become almost impossible. And the third problem is you have a solution, but

you can compute it, but your solutions are unreliable and multi collinearity is a good example of such problem. All these 3 type of problems are known as class of ill-posed problems.

(Refer Slide Time: 03:00)


### Class of Ill-Posed Problems 1

- ▶ Unique solution does not exist
- ▶ 1 Problem of variable selection in "large  $p$ , small  $n$ " setup considered as ill-posed problems
- ▶ Such problems are common in medical sciences.
- ▶ For example, in a study of the efficacy of treatment; suppose the study randomly chose to observe 100 patients. It means the sample size  $n$  is 100.
- ▶ Now scientist collects 1000 of test results from each patient, from regular glucose level to genetic marker, etc. means the number of features  $p$  is 1000.



First, we will talk about class of ill-posed first class of ill-posed problems where unique solution does not exist at all. Particularly problem for variable selection in large  $p$  small  $n$  set up considered as a ill-posed problem, such problems are very common in medical sciences.

(Refer Slide Time: 03:33)




▶ For example, in a study of the efficacy of treatment; suppose the study randomly chose to observe 100 patients. It means the sample size  $n$  is 100.  $n = 100$

▶ Now scientist collects 1000 of test results from each patient, from regular glucose level to genetic marker, etc. means the number of features  $p$  is 1000.  $p = 1000$

**Class of Ill-Posed Problems 1**

- ▶ Unique solution does not exist




1 Problem of variable selection in "large  $p$ , small  $n$ " setup considered as ill-posed problems



For example, in a study of the efficacy of a treatment, suppose the study randomly choose to observe 100 patients; it means the sample is of size  $n$  equal to 100, so  $n$  equal to 100. Now, scientist collects how run 1000 of tests and for each test on the patients say from regular glucose level to genetic market anything and everything that is possible and collect the results of those tests, those are the features of the; features of the problem.

Now  $p$  runs into 1000, so your  $n$  is 100 and  $p$  is 1000 and this is the very common scenario that you can find in the medical sciences.

(Refer Slide Time: 04:32)



▶ Unique solution does not exist *( $X^T X$ ) is not invertible*  
*Infinitely many solutions.*

1 Problem of variable selection in "large  $p$ , small  $n$ " setup considered as ill-posed problems

▶ Such problems are common in medical sciences.

▶ In such kind of problem, you have infinitely many solutions; in fact,  $\beta = 0$  is also a possible true solution.

▶ It means none of the features of your study has any significant effect on your target variable  $y$ , say efficacy. Certainly, it is not a desirable solution.

So, we will find it very often in such cases what happens? In such cases your  $X^T X$  is not invertible. So that means, you have infinitely many solutions; at least one solution, so that means in such cases you have infinitely many solutions; infinitely many solutions you have infinitely many solutions and  $\beta = 0$  is also a solution.

If  $\beta = 0$  is also a solution what is the interpretation of it? It means that none of the features you have spent so much money to collect the results of those you know glucose test and genetic marker etcetera, none of these features have any effect on the treatment of the target variable  $y$ . It does not have any effect. So, this is not a very desirable situation.

So, it is not the problem of the medical test there is no biological or medical problem, it is just a model the mathematical model that we are considering cannot address the issues,

cannot handle this. So, the model is inadequate in this scenario in this the model cannot address the large  $p$  small  $n$  problem.

(Refer Slide Time: 06:18)

The slide is titled "Class of Ill-Posed Problems 2". It contains the following text:

- ▶ Unique solution exists - but computationally not feasible
- 2 Problem of variable selection in large  $p$  is considered as ill-posed problems for model complexity.
- ▶ Suppose you are working in a credit rating group, where you are working with customer databases.
- ▶ The number of customers in the database is more than 100,000, and for each customer, you have 1000 features.



Handwritten notes in red ink below the last bullet point:  $n = 100000$  and  $p = 1000$ .

The slide also features the NPTEL logo in the top right corner, the CMJ logo in the bottom right corner, and a small video inset of a man speaking into a microphone.

So, this is a very common scenario that you will find in the medical sciences. Second problem is typically called problem 2 in the class of glucose problem unique solution exists, but computationally not feasible. For example, you are running a variable selection in large  $p$  considered the impulse problem for model complexity.

Suppose you are working in a credit rating group, where you are working with customer databases the number of customers in your database is more than 100,000. So, your  $n$  is 100,000 ok each customer you have 1000 of features. So,  $p$  is 1000 so theoretically you should be able to run a you know fit a model.

(Refer Slide Time: 07:18)



### Class of Ill-Posed Problems 2

- ▶ Unique solution exists - but computationally not feasible
- 2 Problem of variable selection in large  $p$  is considered as ill-posed problems for model complexity.
- ▶ For such large dataset, if you apply a stepwise feature selection algorithm; then it has to fit  $1 + \frac{p(p+1)}{2} = 500,501$  many models.  
*half-million model*
- ▶ It may take several days to complete the job.
- ▶ However, often time in the corporate environment you do not have several days and upper management wants the result by the end of the day.

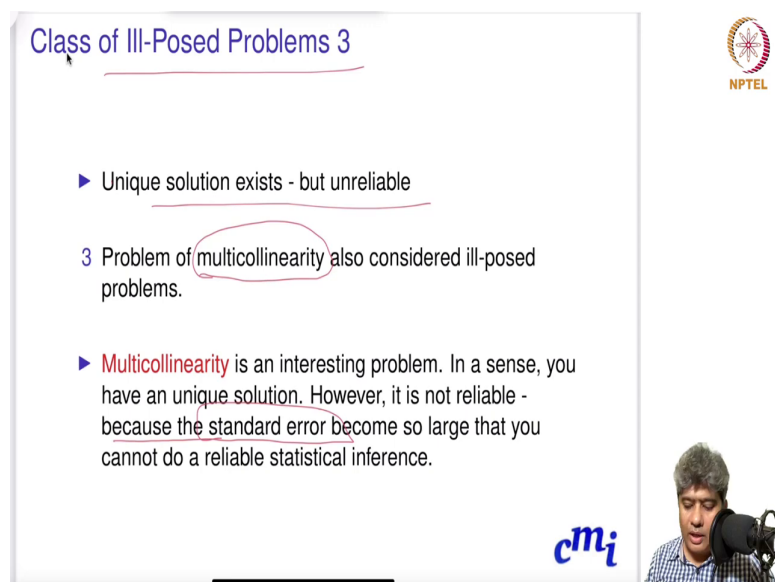
Now if you run a simple stepwise variable selection you have to fit 1 plus  $p$  into  $p$  plus 1 by 2 many models that will be 500,501 many models that is Half a million model. So, you have to fit Half million model to just run a simple stepwise variable selection, Half million model you have to consider forget about base subset selection just to fit a simple stepwise selection stepwise variable selection technique you have to fit Half million model and each model you have to fit with 100,1000 data points.

So, it is going to be a humongous computational facility you require, it may take several days to complete your job and I am talking about this is a very real life scenario you can very easily run into this kind of problems. However often time in corporate environment you do not have several days and upper management wants to result by the end of the day and these are the

scenario where theoretically you have a unique and good solution, but computationally finding the solution is not possible by the time.

You have a time budget your time budget is within 24 hours you have to finish this work, but you have to go through Half million models to find the best model and among the Half million models each model has to be trained using 100,1000 data points. So, it is simply not possible given the computational facility you it is not simply not possible to fit all these models.

(Refer Slide Time: 09:20)



The slide is titled "Class of Ill-Posed Problems 3". It contains the following text:

- ▶ Unique solution exists - but unreliable
- 3 Problem of multicollinearity also considered ill-posed problems.
- ▶ **Multicollinearity** is an interesting problem. In a sense, you have a unique solution. However, it is not reliable - because the standard error become so large that you cannot do a reliable statistical inference.

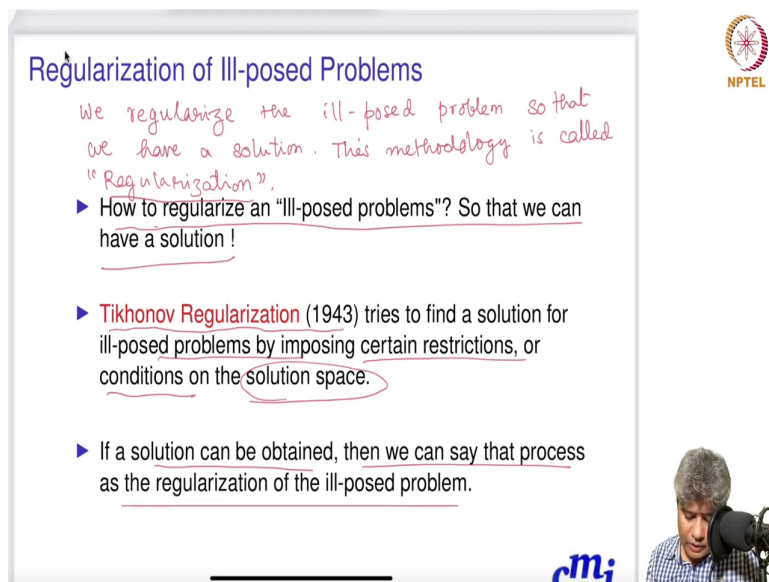
The slide also features the NPTEL logo in the top right corner, the cmj logo in the bottom right corner, and a small video inset of a man speaking into a microphone.

So, unique solution exists now the third problem is unique solution exists, but unreliable. So, the problem of multi colinearity is a third kind of problem where you have a unique solution, but it is unreliable. An interesting problem in the sense that you have a unique solution; however, it is not reliable so because and why it is not reliable? Because standard error



becomes, so large that it cannot do any reliable inference you cannot do any reliable statistical inference.

(Refer Slide Time: 10:04)



The slide is titled "Regularization of Ill-posed Problems". It features a handwritten note in red ink: "We regularize the ill-posed problem so that we have a solution. This methodology is called 'Regularization'." Below this, there are three bullet points:

- ▶ How to regularize an "ill-posed problems"? So that we can have a solution !
- ▶ Tikhonov Regularization (1943) tries to find a solution for ill-posed problems by imposing certain restrictions, or conditions on the solution space.
- ▶ If a solution can be obtained, then we can say that process as the regularization of the ill-posed problem.

The slide also includes the NPTEL logo in the top right corner and a small video inset of a man speaking in the bottom right corner.

So, how you solve it? So, essentially these problems does not have a solution as it is. If this is the scenario what you do? You regularize the problem you regularize the problem. So, that you have a solution and this methodology is called regularization. So, you regularize the ill-posed problem, so that we have a solution this methodology is called regularization ok.

The most popular is called the Tikhonov regularization in 1943 unlike the Tikhonov he was a Russian mathematician, he tries to find the solution for ill posed problems by imposing certain restriction or conditions on the solution space. If a solution cannot be obtained or if solution can be obtained then we can say the process has regularization process of the ill-posed problem.

(Refer Slide Time: 12:16)




Penalizing Objective Function  $OLS = \min_f RSS(f)$

- ▶ The class of functions is controlled by explicitly penalizing  $RSS(f)$  with a roughness penalty

$$PL_2 = PRSS(f; \lambda) = RSS(f) + \lambda P(f)$$

- ▶ The amount of penalty is controlled by  $\lambda \geq 0$ .
- ▶  $\lambda = 0$  means no-penalty
- ▶ Typically  $\lambda$  is estimated from data.

As we take  $f(\mathbf{X}) = \mathbf{X}\beta$

$$PL_2 = PRSS(\beta; \lambda) = RSS(\beta) + \lambda P(\beta)$$
$$= (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta) + \lambda P(\beta)$$


How you find a regularization you essentially put a penalize the objective function, you put a restriction on the parameter space of the objective function that is how you do it. So, how it works? So, the class of function is controlled by explicitly penalizing the residual sum of squares of  $f$  with roughness penalty. What is it?

So, typically what happens is how OLS is works? OLS try to OLS is try to minimize residual sum of squares of the target function that is how the OLS work ok, that solution is typically called OLS. So, penalize this we will still work with the RSS the residual sum of square, but on the functions on the objective function we will put a penal term and we call it penalize residual sum of squares ok and the amount of penalty will depend on the lambda.

Lambda is typically called the tuning parameter. So, if lambda equal to 0 means there is no penalty if you put lambda equal to 0 and if higher the lambda more the penalty and typically

lambda is estimated from the data. So, we take  $f^T X$  equal to  $X^T \beta$  generally and the residual sum of square of beta will be  $y^T y - X^T \beta$ . So, this is the residual sum of squares of beta and then we put a penalty term on the beta that is how we typically work.

(Refer Slide Time: 14:14)

### Penalizing Objective Function

- ▶ What about penalizing  $L_1$ -norm error? Can we penalize  $L_1$ -norm error?
- ▶ Yes we can. The model is:

$$PL_1 = \|y - X\beta\|_1 + \lambda P(f)$$

- ▶ For now we focus on  $L_2$ -norm error.

What about penalizing  $L_1$ -norm error? We can penalize  $L_1$ -norm error we can in this model, but for now we will focus on  $L_2$ -norm error only.

(Refer Slide Time: 14:34)

What penalty to choose?

▶ For the model,




$$PL_2^2(\beta) = (\mathbf{y} - \mathbf{X}\beta)^T(\mathbf{y} - \mathbf{X}\beta) + \lambda P(\beta),$$

one possible choice is  $L_2$ -norm penalty.

▶ That is

$$P(\beta) = (\beta - \beta_0)^T(\beta - \beta_0)$$

▶ Typical case  $\beta_0 = 0$  and the penalty looks like

$$P(\beta) = \beta^T \beta \equiv L_2 \text{ penalty}$$


So, we are not going to focus on L 1- norm error what penalty to choose that always a interesting question. So, the most possible choices or popular choices L 2-norm penalty that is  $P(\beta) = (\beta - \beta_0)^T(\beta - \beta_0)$  if you choose  $\beta_0 = 0$  that will be this is typically L 2 penalty this is typically L 2 penalty ok.

(Refer Slide Time: 15:14)

**Analysis with  $L_2$ -penalty**

▶ We want to minimize the  $L_2$ -penalized loss

$$PL_2^2(\beta) = (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta) + \lambda \beta^T \beta,$$

and we can obtain the Ridge solution as,

$$\hat{\beta}_{Ridge} = \operatorname{argmin}_{\beta} [(\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta) + \lambda \beta^T \beta]$$




▶ An equivalent way to write the ridge problem is

$$\hat{\beta}_{Ridge} = \operatorname{argmin}_{\beta} [(\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta)]$$

subject to  $\beta^T \beta \leq t$ ,  $t \rightarrow 0$

which makes explicit the size constraint on the parameters.

▶ There is a one-to-one correspondence between the parameters  $\lambda$  and  $t$ .



So, the analysis with  $L_2$  penalty is so we want to penalize these  $L_2$  penalized laws essentially  $\mathbf{y} - \mathbf{X}\beta$  transpose  $\mathbf{y} - \mathbf{X}\beta$  plus  $\lambda \beta$  transpose  $\beta$  and if we penalize this so basically that will give us the ridge solution. So, if we all we have to do just penalize just minimize this objective function, this is our objective function and we just minimize with respect to  $\beta$  and the solution is called ridge solution.

So, this the equivalent solution is equivalent construction is you just minimize the residual sum of square essentially, but subject to this constraint this also gives you the ridge solution same solution. Either you minimize this entire objective function with this sort of a Lagrange multiplier kind of effect you can see or you just solve it as a constraint optimization.

Here you minimize the residual sum of square subject to constraint  $\beta$  transpose  $\beta$  less than equal to  $T$ . So, it is something like that what is happening here if you look it carefully

what ticon of solution is doing is you put a budget on the total budget on the sort of a on the beta and if you put T goes to 0. If you put T goes to 0, then what happens is naturally many beta start become shrinking. If the beta start become shrinking beta start going to 0 they typically shrink towards 0 that is what happens.

(Refer Slide Time: 17:24)

### Ridge Regression

- ▶ Solving the following minimization problem,

$$\hat{\beta}_{\text{Ridge}} = \underset{\beta}{\operatorname{argmin}} \left[ (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta) + \lambda \beta^T \beta \right],$$




we have the Ridge solution as

$$\hat{\beta}_{\text{Ridge}} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y},$$

*Analytical solution*

where  $\mathbf{I}$  is the  $p \times p$  identity matrix.

- ▶ Ridge solution is a special case of Tikohonov solution.




  



Now, one can show that lambda here and T they have a 1 to 1 correspondence between 1 to 1 correspondence between these two parameters. But I generally there is no explicit mathematical expression now we have never found it, but one can argue that they will they will have a one-to-one function. Now following the minimization problem if you solve this minimization problem the ridge solution is X transpose.

So, you have an analytical solution X transpose X plus lambda I inverse X transpose y. So, ridge solution is a analytical solution it is a analytical solution this is a very good news for us,

though I am assume if we know the value of lambda then we can compute the ridge solution for beta. So, ridge solution is a special case of Tikohonov solution.

(Refer Slide Time: 18:42)



**LASSO Regression**


- ▶ Least Absolute Shrinkage and Selection Operator (LASSO)
- ▶ The lasso is a shrinkage method like ridge, with subtle but important differences.
- ▶ The lasso estimate is defined as
$$\hat{\beta}_{lasso} = \operatorname{argmin}_{\beta} [(y - X\beta)^T (y - X\beta) + \lambda \|\beta\|_1]$$

$L_1$
- ▶ Equivalently can be expressed as
$$\hat{\beta}_{lasso} = \operatorname{argmin}_{\beta} (y - X\beta)^T (y - X\beta)$$

subject to  $\sum_{i=1}^p |\beta_i| < t$

Next is LASSO regression least absolute LASSO stands for least absolute shrinkage and selection operator or LASSO, the LASSO shrinkage method like ridge with subtle, but different important differences. What is the difference? The difference is the penalty term instead of giving a L 2 penalty here they are putting L 1 penalty and rest of the thing is exactly same.

(Refer Slide Time: 19:23)



▶ The lasso estimate is defined as  $L_1$

$$\hat{\beta}_{lasso} = \operatorname{argmin}_{\beta} [(y - X\beta)^T (y - X\beta) + \lambda \|\beta\|_1]$$


▶ Equivalently can be expressed as

$$\hat{\beta}_{lasso} = \operatorname{argmin}_{\beta} (y - X\beta)^T (y - X\beta)$$

subject to  $\sum_{j=1}^p |\beta_j| \leq t$

cmj

Remark






So, equivalently you can show that you are trying to minimize the residual sum of square subject to constraint this and as same thing you can basically you are budgeting the total amount of beta, if you push the T towards 0, then all the beta will shrink towards 0. So, LASSO does not have close form solution like ridge.



(Refer Slide Time: 19:40)

Remark

- ▶ LASSO do not have closed form solution like Ridge.
- ▶ Computing the lasso solution is a quadratic programming problem.
- ▶ Efficient algorithms are available for computing the entire path of solutions as  $\lambda$  is varied, with the same computational cost as for ridge regression.




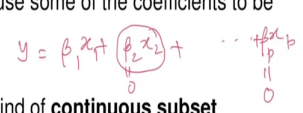


LASSO you have to solve it as a optimization problem you have to solve it as a optimization problem you cannot unfortunately you do not have analytical solution. Computer LASSO solution is essentially a quadratic programming problem turns out to be and efficient algorithms are available for computing the entire path of the solution for different values of lambda. So, with same computational cost as ridge regression. So, in that way it does not matter much.

(Refer Slide Time: 20:20)

Remark

- ▶ Because of the nature of the constraint, making  $t$  sufficiently small will cause some of the coefficients to be exactly zero.
- ▶ Thus the lasso does a kind of continuous subset selection.
- ▶ Ridge takes care of multicollinearity kind of issues.
- ▶ compromise between ridge and lasso was give Zou and Hastie (2005), known as Elastic Net penalty

$$P_{EN}(\beta) = \lambda \sum_{j=1}^p (\alpha \beta_j^2 + (1 - \alpha) |\beta_j|)$$


So, because of the nature of the constraint making  $t$  sufficiently small will cause some the coefficient to be exactly 0, ok. Thus, what does it mean some of the beta? So, if I have  $x_1 \times x_2 \dots \times x_p$  and now suppose the  $\beta_2$  turns out to be exactly 0. That means, in my model  $x_2$  does not have a effect. So, effectively  $\beta_2$  equal to 0 means that you are dropping that LASSO solution is dropping that predictor from the model.

So, this is that is how it would be, so that is why LASSO solution it is like you know for stepwise variable selection you are fitting each model. But here I am fitting only one model remember that I am fitting  $\beta_1 \times \beta_2 \times \dots \times \beta_p$  I am fitting one model with LASSO penalty and if turns out suppose  $\beta_2$  equal to 0  $\beta_p$  equal to 0. That means, effectively I am dropping  $x_2$  and  $x_p$  from the model and I am obtaining the best fitted model by that.

So, in one shot I am getting the best model while dropping the predictors or the features which has no effect in the y which has no effect in the y. So, this thing this phenomena is called continuous subset selection. Now, interestingly ridge takes care of the multicollinearity kinds of issue and LASSO take cares of the do the variable selection. So, compromise between the ridge and LASSO was given by Zou and Hastie 2005 known as Elastic Net Penalty.

(Refer Slide Time: 22:37)

Hastie (2005), known as Elastic Net penalty

$$P_{EN}(\beta) = \lambda \sum_{j=1}^p (\alpha \beta_j^2 + (1 - \alpha) |\beta_j|)$$

$0 < \alpha < 1$

cmj

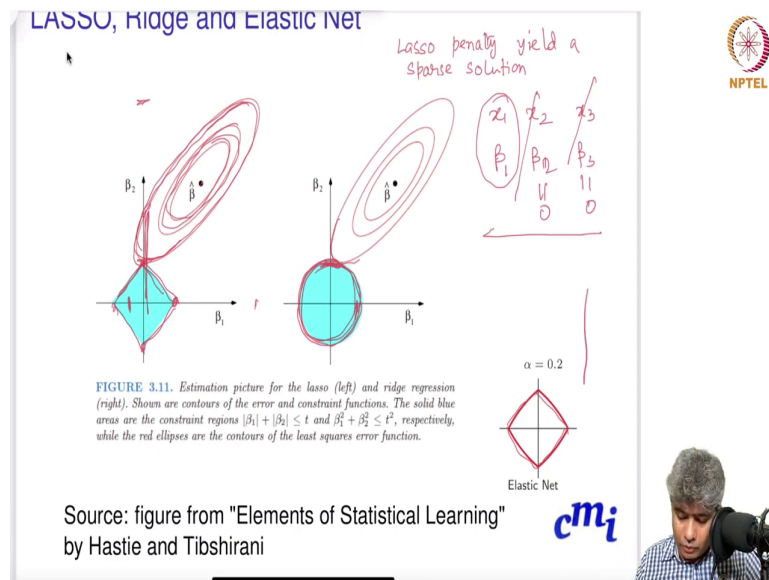
NPTEL

LASSO, Ridge and Elastic Net

The slide contains two plots. The left plot shows a diamond-shaped contour (LASSO) and a circular contour (Ridge) centered at the same point  $\hat{\beta}$  in the  $\beta_1$ - $\beta_2$  plane. The right plot shows an elliptical contour (Elastic Net) centered at the same point  $\hat{\beta}$ , which is a convex combination of the diamond and circle contours.

They invented this elastic net penalty which is essentially a convex combination of LASSO and ridge and LASSO, you can see alpha has to be between 0 and 1 and they build this new penalty.

(Refer Slide Time: 22:58)



So, here is an extract from Tibshirani's book. So, this is an example of LASSO, if LASSO how the LASSO penalty would look like and you see the LASSO penalty is very sharp has very sharp edges and because of that it is the optimization technique. So, this is typically the Bayesian interpretation of this is the posterior distribution. In Bayesian interpretation what happens in classical statistics we do all inference based using sampling distribution.



In Bayesian inference we do all the inference using posterior distribution, here they have put a posterior distribution and what you can see that for the LASSO this posterior distribution is pulled so much that for beta one it includes 0. So, the beta 1 does not have any effect; whereas, because beta 1 can be somewhere in this region, but beta 2 is typically in this region.

So,  $\beta_2$  still have effect but  $\beta_2$  does not; whereas, in this is the ridge penalty sort of a ball it creates a ball effect it does not create a and because it is much more smooth it does not pull the coefficient that much. So, still it may touch, but it is not actually containing the  $\beta_2$ . So, ridge penalty does not yield a yield a sparse solution, but LASSO penalty yields a sparse solution.

So, what I mean by sparse solution what I mean by sparse solution? If I have  $x_1$   $x_2$   $x_3$  and that means corresponding coefficient  $\beta_1$   $\beta_2$  and  $\beta_3$  if any of them becomes 0 the solution itself will force tell me that yes, it is exactly 0. That means, automatically I am dropping that predictor from the model, if  $\beta_2$  and  $\beta_3$  turns out to be exactly 0 then I can automatically the model itself is dropping and the final solution itself is dropping  $x_2$  and  $x_3$  and keeping only  $x_1$   $\beta_1$  in the model and as a result and LASSO can do that.

So, that is why LASSO penalty have a capability to yield sparse solution, but ridge does not. Whereas elastic net has an interesting property which is a convex combination of both ridge and LASSO. So, it do take care of both multicollinearity and variable selection. Whereas, the ridge do take care of the multicollinearity, but cannot handle the variable selection. On the other hand LASSO do handle variable selection, but cannot handle the multicollinearity ok.

(Refer Slide Time: 26:44)



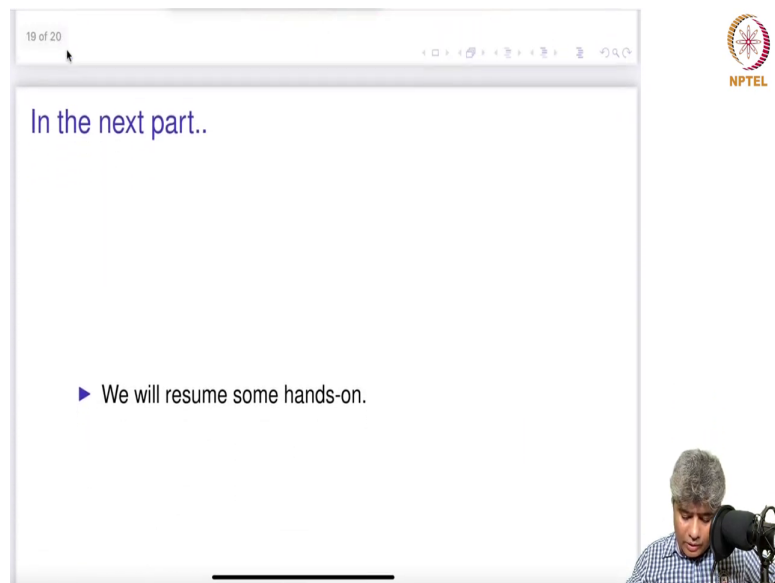
Tikhonov Regularization for multicollinearity and feature selection

- ▶ Ridge Regression takes care of multicollinearity (Hoerl and Kennard (1970))
- ▶ LASSO Regression takes care of feature selection (Tibshirani, 1996)
- ▶ ElasticNet Regression takes care of feature selection (Zou and Hastie, 2006) & multicollinearity

cmj

So, with this I have kind of summarize here Ridge Regression takes care of the multicollinearity the paper it was developed by Hoerl and Kennard in 1970. Then 1996 LASSO Regression which takes care of the feature selection and Elastic Net Regression takes care of both feature selection and multicollinearity ok nice alright.

(Refer Slide Time: 27:18)



19 of 20

In the next part..

- ▶ We will resume some hands-on.

The image shows a presentation slide with a white background and a blue border. At the top left, it says "19 of 20". The main text on the slide is "In the next part.." in blue, followed by a bullet point "▶ We will resume some hands-on." in black. In the bottom right corner, there is a small video inset showing a man with grey hair wearing a blue and white checkered shirt, looking down. To the right of the slide, there is a circular logo with a red and white design and the text "NPTEL" below it.

In the we will now resume some hands on and let us go for go to do some hands on and see you then take care bye.