What exactly happens in Multicollinearity 9

behen multicollinearity is present in a model, it can lead to several issues, including:

- 1. Difficulty in identifyly the most important predictors:

 Due to night condition between independent

 Variables, it becomes abellengly to determine

 variables, it becomes abellengly to determine

 which variable has the most significant impact

 on the dependent variable.
- 2. Inflated Standard error: Multicollinearity can lead to large, standard error for the regression coefficient, which decrease the statistical pone and can make it unalleryly to determine the and true relationship between the Independent and dependent variables.
- 3. Unitable and remeliable estimate: The regression coefficient becomes sensitive to small charges in the data, making it difficult to interpret the results accurately.

Perfect Multicollinearity

Perfect multicollineority, occurs when one independent van'able in a multiple regression model is an exact linear combination of one or more other Independent variables. In other news, there is an exact linear relationship between the independent vaniable, making it impossible to uniquely estimate the individual effects of each variable on the | percent = lox gpa + error dependent variable.

corr linear 2

 $X_1 = \alpha_1 X_2 + \alpha_0 + evens$

add decipline money example Cgpa | percent | Lpa 8.5 83

C

Perfect collinearity 7

 $x_1 = a_1 x_2 + a_0$

Cgpa | percent | Lpa 912

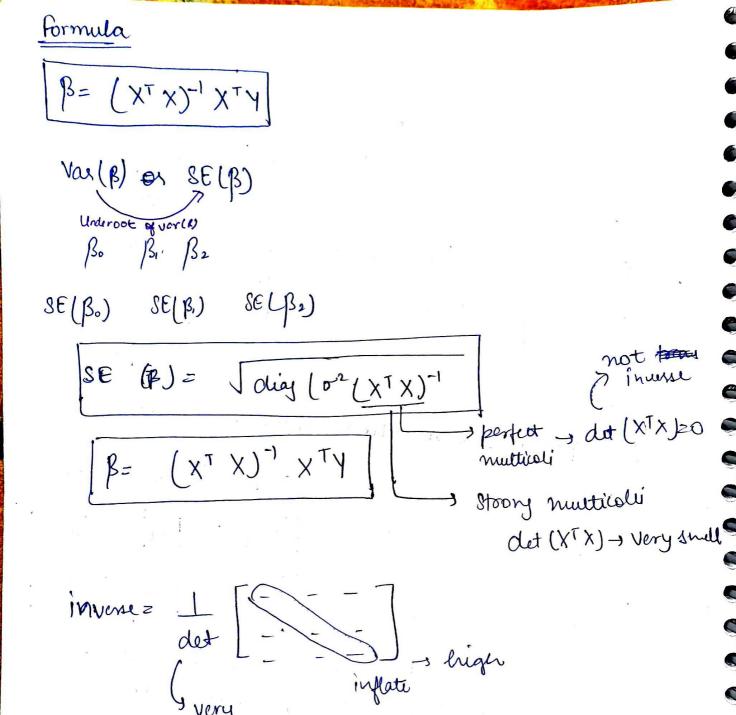
Example

-		
Cepa	Percent !	lpa
Q	80	3
6	60	4
6		_

percut = lox gpa @

lpa = Bo + BiGpa + B2 perut + euor

OLS 7 $\beta = (x^T x)^{-1} X$ > dusign Matrix input data 3×2 14 840000 8400 XTX finding Tuverse Check Beterminte (det) 140 14 8400 84 Det= 2(0) -14(0) + 140 (0) = 9 det is zero them it is singular matric and we Cant find inverse and also [B] not find. Proofe our data has mutticollinearity → runstable coefficienz froblem



Small

How to remove multicollinearity

- 1. Collect <u>more data</u>: In some cases, multicollinearly neight be a result of a limited sample size. Collecting more data, if possible, can been reduce multicollinearity and improve the stability of the model.
- O. Remove One of the brighty coullated Narriable:

 To two or more are brighty correlated, consider

 Removing one of them from the model. This

 step can been eliminate redundancy in the model

 and reduce multicollinersity. Choose the variable

 to remove based on domain knowledge, variable

 Importance, on the one neith the brighest VIF.
- 2. <u>Combine Corelated Variable</u>: If correlated Independent variable represent similar information, consider combinity them into a single variable. This combination can be done by average, summing as revery other mathematical operations, depending on the Content and the nature of the variables—

4. <u>Un partial least square sugression (PLS)</u>: PLS
is a technique that combines features of both
principal component analysis and multiple
regression. It identifies linear combinations of
the fredictor variables (Colled latent variable) that
have the highest covariance with the suspense
variable, reducing multicollinearity while retaining
must of the predictive power.

* First Find PCA > from > 2R