1. Milestone tracker for
2. Show the placement history

Polynomial Regression

Preprocessing $\rightarrow$ <u>Polynomial Features</u>

Linear Regression

Number of output features $\longrightarrow$ Input Feature ✓
              +
              Degree ✓

Linear Regression $\rightarrow$ Assumption

1. Linearity $X \longleftrightarrow Y$
2. Error should be Independent   } Interview
3. Homoscadasity
4. No Multicollinearity $\leftrightarrow$ No/ Little

> Correlation between the Independent data

Total Sqf = F1 Sqft + F2 Sqft + Underg Sqft

keep Total Sqft   or   keep 3 column

Multi Collineanty

Polynomial regression creates many new feature

$$x_1 \rightarrow x_1^2, x_1^3$$

If duplicate or redundent, will py remove them

NO, Not remove anything

Polynomial — Multicollinea

↓

Problems of Multi Collineanty ←

1. unstable Coefficients

2. large swing in the Coefficient values

3. inflated variance

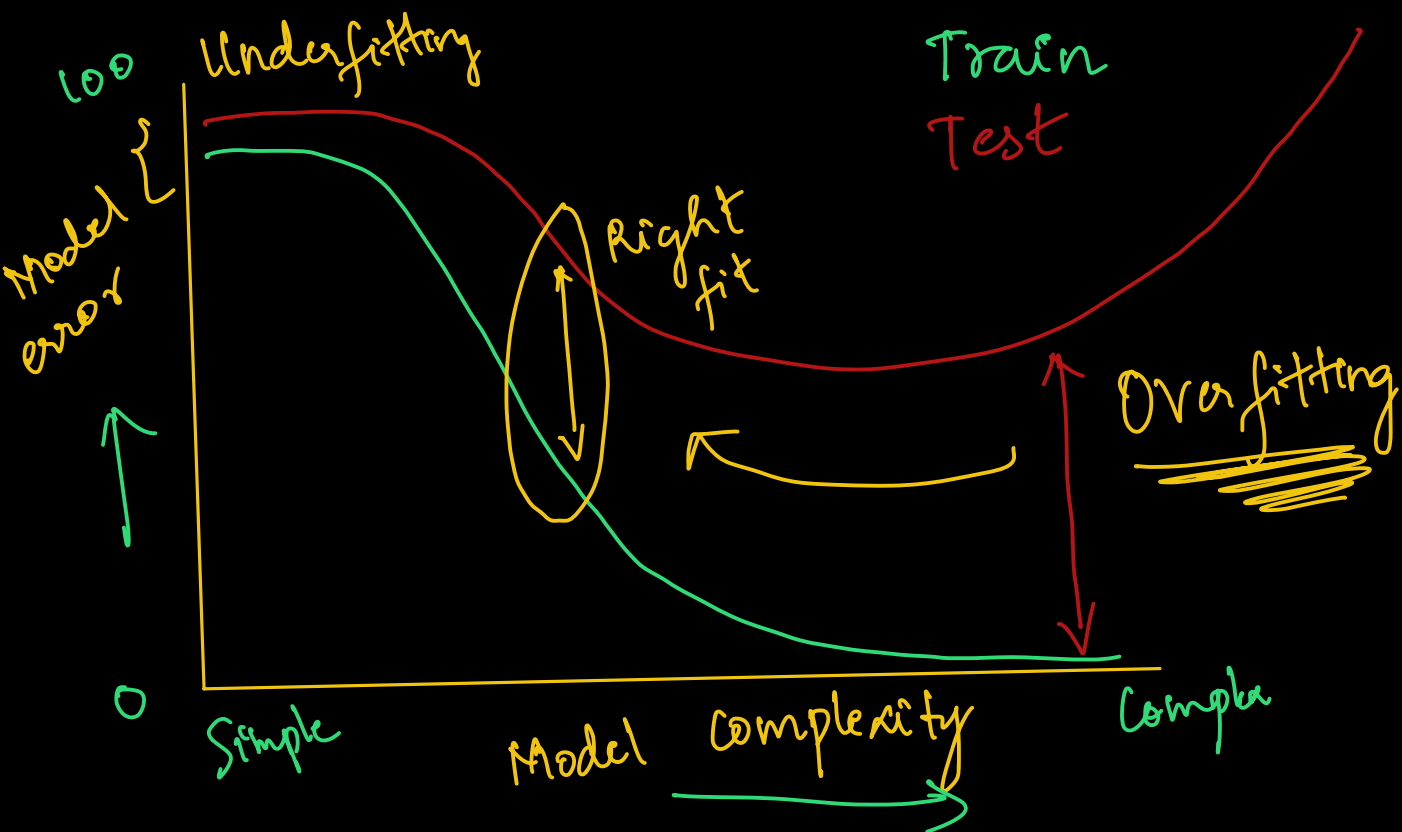# 4. Poor model generalization

Linear Regression + Polynomial features

↓

Overfit

↓ avoid

Regularization :

Underfitting   Train
                 Test

100

Model {
error

Right
fit

Overfitting

O

Simple   Model complexity →   Complex

Overfitting

Co-efficients $\rightarrow$ high ~

Gradient descent

$\begin{cases} \text{find the} & \text{Co-efficient + intercept} \\ \text{Constraint} \rightarrow & \underline{\text{Cost function}} \downarrow \text{minimal} \\ & \text{MSE} \end{cases}$ | linear regression

Regularization

$\begin{cases} \text{find the} & \text{Co-efficient + intercept} \\ \text{constraint} \rightarrow & \text{Cost function} \downarrow \text{minimal} \\ & \boxed{\text{MSE}} \quad + \\ & \boxed{\text{Co-efficient} \downarrow \text{minimal}} \end{cases}$

Penalty factor

Shopping

~ 1km   Home ✓

walk

$ 500

Shop $\rightarrow$ monthly

Multiple Products
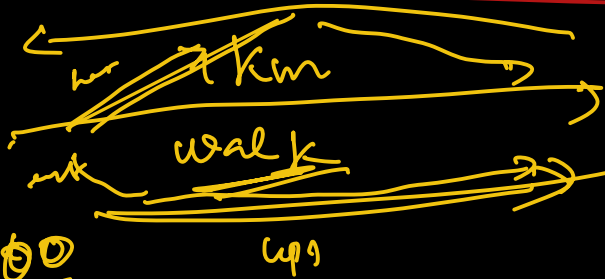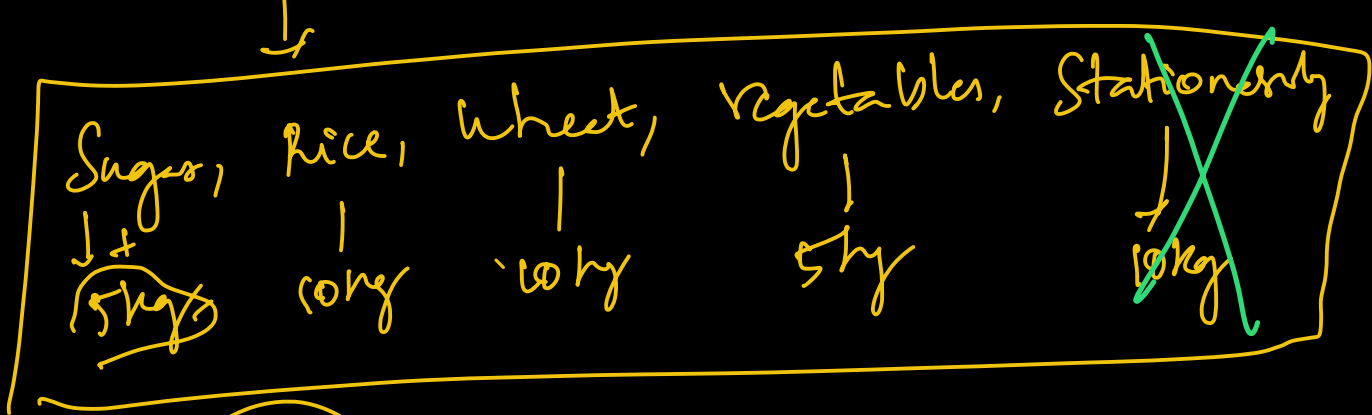
Sugar, Rice, Wheat, vegetables, Stationery

Sugar
↓ +
5kg

Rice
|
10kg

Wheat
|
10kg

vegetables
|
5kg

Stationery
|
10kg ✗

(5.0 kg)

carry and

Penalize ← Walk   Penality

If

$$\frac{Sugar}{1kg} + \frac{Rice}{2kg} + \frac{Wheat}{1kg} + \frac{Vegetable}{1kg} + \frac{State}{0kg}$$

$$1 + 2 + 1 + 1 + 2$$

7 kg

Car → driving

⇓

{ 150 N/m

f → 200 hr

h → 250 hr/m

200 h } face to   broken/-

Penality

$$so c \Rightarrow 120 \text{km/h} \longrightarrow \oint 1000 \, V$$

Cost function $= \sum \dfrac{(y - (\beta_0 + \beta_1 x))^2}{n}$ | LR

Regularization

Cost function $= \sum \dfrac{(y - (\beta_0 + \beta_1 x))^2}{n} + \lambda \sum_{0}^{n} (\beta)^2$

$\lambda = 0.1$

$\lambda = 1$

Penality factor

1. Ridge Regression
2. Lasso Regression
3. Elastic Net Regression

1. Ridge Regression L2

Cost function $= \sum \dfrac{(y - (\beta_0 + \beta_1 x))^2}{n} + \lambda \sum_{i=0}^{n} (\beta_i^2)$

Shrink all the coeff.
doest make anythin

2. Lasso

Regression L1 MSE

$$\text{Cost function} = \frac{\sum (y - (\beta_0 + \beta_1 x))^2}{n} + \lambda \sum_{i=0}^{n} |\beta_i|^1$$

Feature Selection

Zero

↓↓
Shrink all the Coeff
+
makes unimportant Coeff Zero

3. Elastic Net

$$\text{Cost func} = \frac{\sum (y - (\beta_0 + \beta_1 x))^2}{n} + \lambda \left[(\alpha) \sum |\beta| + (1-\alpha) \sum (\beta)^2 \right]$$

Gradient Descent

Cost function = MSE

find m and b

MSE ↓ minimal

Derivative

$$y = 2x^2 + 2x + 5 \quad \checkmark$$
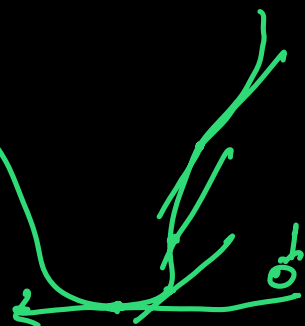
What is value of x
where y is minimum

└→ Derivative

$$y = 2x^2 + 2x + 5 + [1x + 3]$$
$$= 2x^2 + 3x + 8 \quad \longrightarrow \quad \omega$$

$$MSE = \dfrac{\sum (y-\hat{y})^2}{n} \quad \checkmark$$

$$\hat{y} = \beta_0 + \beta_1 x$$

Partial
Derivate

$$MSE = \underbrace{\dfrac{\sum (y-\hat{y})^2}{n}} + \underbrace{\sum (\beta_0^2 + \beta_1^2)} \quad \checkmark$$

$$\beta_0 = \dfrac{100}{50} \checkmark$$

$$\beta_1 = 50 \checkmark$$

$$+ \; 10,000 \; + \; 2,500$$

$$\boxed{12,500}$$

$$\left(12,800 + 150\right)$$

MSE