

## GD Steps

Pick R.V.'s  
for the  
Parameters.

- (1) Calculate SSR
- (2) Find the derivative of  
SSR w.r.t. Intercept  
to find the slope  
  
= Derivative of loss fns
- (3) Pick a R.V. for intercept  
eg. 0
- (4) Calculate the derivative  
when intercept = 0
- (5) o/p of step-4 is plugged  
in for step-size calculation  
= Slope  $\times$  learning rate  
( $\alpha$ )

⑥ 
$$\text{New-intercept} = \frac{\text{old-intercept}}{\text{Step-Size}}$$

⑦ Now plug in new intercept value into  $\frac{d(SSR)}{d\text{intercept}}$

⑧ Repeat until step-size is close to 0.

We will use SSR as our loss func.

We want to find those values for slope & intercept that gives min. SSR.

→ Take derivative of loss function

→ w.r.t. Intercept &

→ w.r.t. Slope

Gradient: when you have 2 or more derivatives of some function, they are called Gradients.

We use Gradient to descend to the lowest point in the loss func, which is the 'SSR'.

This algo is called  $\tilde{GD}$

# Learning Rate:

Start with a large value, and gradually decrease it to reach to the min. value.

## Stopping Criterion:

Repeat the process until all step-sizes are very small or we reach the maximum number of steps.