DINESH KUMAR M
AM23M022

DATA SCIENCE : THEORY AND PRACTICE

OPTIMIZATION TECHNIQUES:

Optimization criteria:-

at point $\bar{w}$, first derivative is zero,
1st non-zero higher derivative is denoted by n,

1) $n \rightarrow$ odd ; $\bar{w} \rightarrow$ inflection point

2) $n \rightarrow$ even ;

a) derivative $\rightarrow$ +ve, $\bar{w} \rightarrow$ local minimum

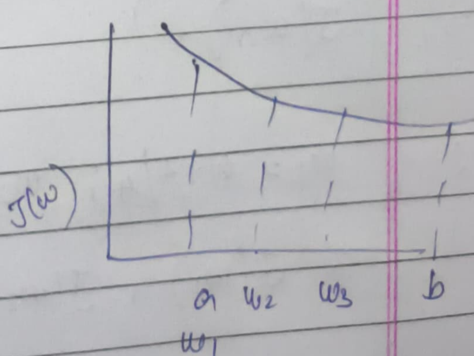
b) derivative $\rightarrow$ -ve, $\bar{w} \rightarrow$ local maximum

Aim for optimization:

Finding w for which $J(w)$ is minimum

~~Method 1:~~ Bracketing method

- Exhaustive search;



1: $\Delta w = (b-a)/n$

$w_1 = a$;

$w_2 = w_1 + \Delta w$

$w_3 = w_3 = w_2 + \Delta w$

2: $J(w_1) \geq J(w_2) \leq J(w_3)$

min lies b/w (w_1, w_3)

(or)

$w_1 = w_2$; $w_2 = w_3$; $w_3 = w_2 + \Delta w$

$\downarrow$

3. $w_1 = w_2$, $w_2 = w_3$; $w_3 = w_2 + \Delta w$

If $w_3 \leq b$;

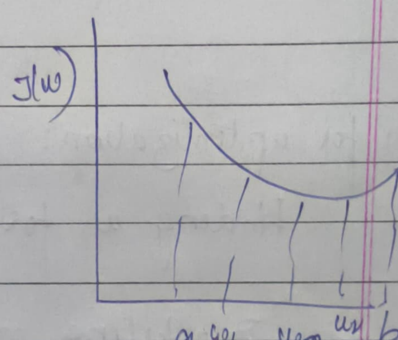
Step 2.

(or)

no min between (a, b)

Method 2: Region Elimination method

- Interval halving



1: a, b

$w_m = a + b/2$

$L = b - a$.

$J(w_m) = ?$

2: $w_1 = a + L/4$; $w_2 = b - L/4$

$J(w_1)$, $J(w_2) = ?$

3: $J(w_1) < J(w_m)$

$\Rightarrow b = w_m$

$\Rightarrow w_m = w_1$

$L = b - a$

if $|L| < \epsilon$,
stop

(or)

*) $J(w_2) < J(w_m)$

$a = w_m$; $w_m = w_2$

↓ $L = b - a$

if $|L| < \epsilon$
stop

(or)

*) $a = w_1$, $b = w_2$

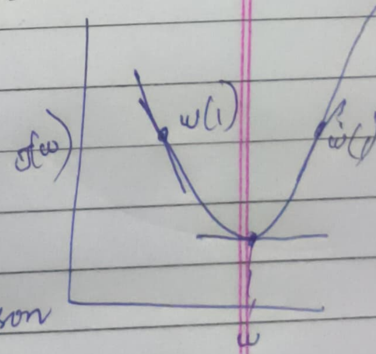
$L = b - a$

If $|L| < \epsilon$
stop

Method 3: Gradient-based approaches

Aim: Min → point where $J'(w) \simeq 0$

$$w^{k+1} = w^k - \frac{J'(w)}{J''(w)}$$



Qn: Interval halving vs Newton-raphson

Newton-raphson tends to be faster as the convergence rate is quadratic