•	· · ·	•	•	•	•	•	(D		Ν,	• • •	E7	51	י ר ר	J	•	•	•	•	C)F		•	•	•		R	D	s s	•	E	ĴТ	r R C	5 -	· · · ·	· · ·	L	.TS	s S	•	•	· · ·	
	• •		•	0	•	•	• •		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	0	•	• •	•	0	•	•	•	•	• •	• •	•	•	•	0		0 0	• •	
	• •	•	•	•	•	•	• •		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	•	•	•	•	•	•	• •		•	•	•	•	•	•	• •	
٠	• •	•		٠	٠	*		• •	٠	٠	٠			٠		٠	٠	•	٠			•		٠	٠	•	• •	•	•	٠	•	٠	٠	• •			٠	٠	٠	•	•	• •	
	• •	٠	•	٠	٠	٠		•	۰	٠	۰	٠	•	٠	•	٠	٠	٠	٠		•	٠		٠	٠	•	• •	•	٠	٠	٠	٠	٠	• •		٠	۰	٠	۰	•			
•	• •	•	•	•	•	•			•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•			•	•	•	•		• •		•	•	•	•	•	•		
		•						, ,		٠			٠													•		•						•				٠		•	•		
٠	• •	•	٠	٠	٠	٠	• •	• •	٠	*	0			٠		٠		•	0	•		*	•	٠	٠	٠	• •	•	٠	*	٠		٠	• •		0		٠	0	٠	٠	• •	
•	• •			٠	•	٠		•	٠	٠	۰			٠		٠	٠	٠	٠			٠		٠	٠		• •	•		•	٠	٠	٠	• •		٠	٠	٠	۰	•	•		
•			•	•	•	•			•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•			•	•	•	•	•			•	•	•	•	•	•		
		•	•																							*					*							•					
	• •			٠			• •			٠				٠		•		•	•					•	٠			•			•	•		• •		•			٠		٠	• •	
٠	• •	•	٠	٠		٠	• •	•	٠	٠	٠	٠	٠	٠		•	٠	•	٠	٠	•	٠	٠	٠	٠	•	• •	•	٠		٠	٠	•	• •		٠	٠	٠	٠	٠	٠	• •	
		٠		•	٠					•	0	٠	٠		٠	0		٠	0	٠	٠	٠	٠			٠		•	٠	٠	٠	0		• •	• •	0	0	٠	0	٠	•		
		•	•	•		•			•	•	•	•	•	•	•	•	•	•	•		•	•	•	•	•	•			•	•	•	•		• •		•	•	•	•	•	•		
۰		٠	٠	٠	٠						0	٠				0			0		•	٠		٠		•		•	٠	٠				• •		0		٠				• •	
				•		•	• •	• •			•			•		•	•		•			٠		•	•	•	• •	•	٠			•	•	• •	, .	•	•	٠	•	•			
																																										· ·	
	• •			٠		•			٠	•	۰			۰		٠		•	٠					٠	٠		• •	•			•	•	•	• •		٠	٠	•	٠		٠	• •	
٠	• •	•	•	٠	٠		• •	•	٠	٠	0	•	٠	٠	•	٠	٠	٠	0	•	•	٠	٠	٠		٠	• •	•	٠	*	٠	•	٠	• •		0	0	٠		٠	٠	• •	
																																										• •	
•	• •		•	•	•	•	• •	•	•	•		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	•	•	•	•	•	•	•	• •	•	•		•	•	•	• •	

$J(\theta) = - \sum_{i=1}^{N} (y_i \log \hat{y_i} + (1-y_i) \log \hat{y_i})$	$(1-\hat{y_i})$
$\frac{\partial J(\theta)}{\partial \theta_{j}} = \sum_{i=1}^{N} [\hat{y_{i}} - \hat{y_{i}}] x_{i}^{j}$	· · · · · · · · · · · · · · · · · · ·
Where $j \notin j^{th}$ element $z_{i} \notin m$ data point	
i < in data point	
$\begin{array}{c} \cdot \cdot$	
· ·	· · · · · · · · · · · · · · · · · · ·
. .	· · · · · · · · · · · · · · · · · · ·

 $\frac{\partial J(b)}{\partial j} = \sum_{i=1}^{N} [\hat{y_i} - \hat{y_i}] x_i$ 901901 991901 $\frac{\partial}{\partial \theta_2} \frac{\partial J(\theta)}{\partial \theta_1}$ 3 27(0) 907 9 0 T

 $\frac{\partial \overline{J(b)}}{\partial \overline{J(b)}} = \sum_{i=1}^{N} [\hat{y_i} - \hat{y_i}] x_i$ 90, 90, <u>9 91(0)</u> $\frac{\partial}{\partial \theta_2} \frac{\partial}{\partial \theta_1}$ 5 97(P) 202 3 0 T as Hjk entry let us compute 901 90K

$\frac{\partial \overline{J(B)}}{\partial \overline{J(B)}} = \sum_{i=1}^{N} [\hat{y_i} - \hat{y_i}] x_i^i$.
let us compute a <u>2</u> <u>3</u> <u>3</u> <u>3</u> <u>0</u> 30 <u>j</u> 30 <u>k</u>	as Hjr entry
$\frac{\partial J(\theta)}{\partial \theta_R} = \sum_{i=1}^{N} (\hat{y_i} - \hat{y_i}) \chi_i^R$	
$\frac{\partial}{\partial j} \frac{\partial J(0)}{\partial 0 k} = \frac{\partial}{\partial 0 j} \frac{\xi}{i=1} (\hat{y_i} - y_i)$	R K
. .	

$\frac{\partial J(\theta)}{\partial j} = \sum_{i=1}^{N} [\hat{y_i} - \hat{y_i}] \times_i^j$	
let us compute a <u>2</u> <u></u>	entry
$\frac{\partial J(\theta)}{\partial \theta_R} = \sum_{i=1}^{N} (\hat{y_i} - \hat{y_i}) \hat{z_i}$. .
$\frac{\partial}{\partial 0j} \frac{\partial J(0)}{\partial 0k} = \frac{\partial}{\partial 0j} \frac{\sum_{i=1}^{N} (\hat{y_i} - \hat{y_i}) x_i^k}{\sum_{i=1}^{N} \partial 0j}$	
$H_{jk} = \frac{\partial}{\partial \theta_j} \hat{\xi}_j^k \hat{\chi}_j^k$	

90j 9 <u>1(p)</u> =	$\sum_{i=1}^{N} \left[y_{i}^{2} - y_{i}^{2} \right] \times_{i}^{j}$
Hjk	$= \frac{\partial}{\partial \rho_{j}} \hat{y}_{i}^{k} \hat{x}_{i}^{k}$
	$= \sum_{i=1}^{N} \hat{y}_{i} \left(1 - \hat{y}_{i}\right) \times_{i}^{k} + \frac{\partial}{\partial \theta_{j}} \left(x_{i}^{i} \theta_{l} + x_{i}^{2} \theta_{2} + \cdots\right)$
	$= \sum_{i=1}^{N} \hat{y_{i}} \left(l - \hat{y_{i}} \right) \frac{k}{x_{i}} \hat{z_{i}}$
· · · · · · · · · · · ·	
	· · · · · · · · · · · · · · · · · · ·
	· · · · · · · · · · · · · · · · · · ·
	· · · · · · · · · · · · · · · · · · ·

· · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · · ·	$\sim \sim \sim k$	· · · · · · · · · · · · · · · ·
H jr		$(1-\hat{y};) \times_{i}^{k} \times_{i}^{j}$. .
· · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · ·		\hat{y}_{1}
· · · · · · · · · · · · · · · · ·			
· · · · · · · · · · · · · · · · · · ·		· · · · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·
· · · · · · · · · · · · · · · · · ·	· · · · · · · · · ·		
· · · · · · · · · · · · · · · · · · ·		<mark></mark> .	· · · · · · · · · · · · · · · · · · ·
· · · · · · · · · · · · · · · · · ·		· · · · · · · · · · · · ·	· · · · · · · · · · · · · · ·
· · · · · · · · · · · · · · · · · · ·			
		<mark></mark> .	· · · · · · · · · · · · · · · · · · ·
			· · · · · · · · · · · · ·
· · · · · · · · · · · · · · · · · · ·			

· · · · · ·	· · · · · · · · · · ·	· · · ·	jk	· · · · · · · · · · · · · · · · · · ·	· · ·	· · ·			Ą;	· · · ·	- c)))))		2 7 1		· · · · · · · · · · · · · · · · · · ·		· · · · · · · · · · · · · · · · · · ·	· · ·	· · · · · · · · · · · · · · · · · · ·		· · ·	· · · · · · · · · · · · · · · · · · ·
	× ×	j _X k	· · ·	· ·	· ·	· ·	· · ·	•	· ·	· ·	y i	· · ·		· ·	•	· ·	Ŷ		· ·			Ŷ		· ·
		· · ·		 	· ·	• •		÷		• •			•	• •	•	 	· ·	•	• •		· ·	•	•	• •
· · · · ·	· ·	•	· · ·	· ·	• •	• •	· ·	•	• •	• •		• •	•	• •	•	• •	· ·	•	• •	•	· ·	•	•	· ·
	· ·			• •	• •	• •		•	• •	• •	•••		•	• •	•	• •		•	• •	•		•	•	• •
· · · · ·	· · ·		· · · ·	· ·	• •	• •		•	• •	• •		• •	•	• •	•	• •	•••	•	• •	•	• •	•	•	• •
		· · ·		 	• •	••••		•		• •		• •		• •	•	• •	•••		• •	•		•	•	• •
		· · ·		• •		•••		•	• •	• •	• •	• •	•	• •	•	•••	•••		••••			•	•	•••
					J	• •		•	• •	• •	• •		•	• •	•	• •	• •		• •	•		•	•	• •
		· · ·	· · ·	· · ·	• •	• •		•		• •				• •	•	• •	• •	•	• •	•	• •			• •
	ay	J^{-2}	-; k=	5	• •	• •	· ·	•	• •	• •	••••	• •	•	• •	•	• •	• •	•	• •	•	• •	• •	•	• •
				• •		• •	• •	•	• •	• •			•	• •	•	• •	• •	•			• •		•	• •

 . .<	· · · ·	· · ·	jĸ			y	`, (-	- J;) <i>k</i> <i>z</i> ;			· · · ·	· · · · ·	· · · ·
	× ¹	x x x		· · · ·	· · · ·			<u>}</u>		· ·	Ŷ			· · · ·
· · · · ·				· · · ·	· · · ·		· · · · ·	· · · · ·		· · ·				· · · ·
· · · · ·				· · · ·	· · · ·		· · · · ·	· · · · ·		· · ·		· · · ·	· · · ·	
· · · · ·				· · · ·	· · · ·		· · · · ·			· · ·			· · · ·	· · · ·
· · · · · ·	Hjk	· · ·		ý.(+	-A')	0.0 n(týn)		NXI		г Х і	diag	lý O	lfy))X vDe m	· · · ·

H j k	$= \sum_{i=1}^{N} \hat{y}_{i} \left(l - \hat{y}_{i} \right) \chi_{i}^{k} \chi_{i}^{j}$. .
Hjk =	$\begin{array}{c} \mathbf{x}^{T} \\ \mathbf{x}^{J} \\ 1 \times \mathbf{N} \\ \begin{array}{c} \mathbf{y}_{n}(\mathbf{i} - \mathbf{y}_{1}) & 0 \\ \mathbf{y}_{n}(\mathbf{i} - \mathbf{y}_{n}) \\ \mathbf{y}_{n}(\mathbf{i} - \mathbf{y}_{n}) \\ \mathbf{y}_{n}(\mathbf{i} - \mathbf{y}_{n}) \\ \mathbf{y}_{n}(\mathbf{i} - \mathbf{y}_{n}) \end{array}$. .
HjR	= xj ^T diag (ŷ O(rŷ))x ^k	. .
	. .	
		· · · · · · · · · · · · · · · · · ·

$H_{jk} = \sum_{i=1}^{n} \hat{y}_{i} \left(l - \hat{y}_{i} \right) \chi_{i}^{k} \chi_{i}^{j}$
$H'_{jk} = \chi_{j}^{T} \begin{bmatrix} \hat{y}_{i}(r_{\hat{y}_{i}}) & 0.0 \end{bmatrix} \chi_{NX1}^{k} \\ 1 \times N \begin{bmatrix} y_{i}(r_{\hat{y}_{i}}) & 0.0 \end{bmatrix} \chi_{NX1}^{k} \\ y_{i}(r_{\hat{y}_{i}}) \end{bmatrix}_{NXN}^{k}$
Hjk = xj ^T diag (ý O(rý)) ^k = Xj ^T D x ^k
What con nee say about D?

$H_{jk} = \sum_{i=1}^{N} \hat{y}_{i} \left(I - \hat{y}_{i} \right) \chi_{i}^{k} \chi_{i}^{j}$
$H_{jk} = \chi_{j}^{T} \begin{bmatrix} \hat{y}_{1}(l-\hat{y}_{1}) & 0.0 \end{bmatrix} \chi_{N\times 1}^{k}$ $I_{XN} \begin{bmatrix} \hat{y}_{1}(l-\hat{y}_{1}) & 0.0 \end{bmatrix} \chi_{N\times 1}^{k}$ $\tilde{y}_{N}(l+\hat{y}_{2}) \end{bmatrix}_{N\times n}$
$H_{jk} = x_{j}^{T} diag \left(\hat{y}_{O}(r_{\hat{y}}) \right) x^{k} = X_{j}^{T} D x^{k}$
What can use say about D^2 Each $\hat{y} \in [0_1^2]$; $1-\hat{y} \in [0_1^2]$ Man $\hat{y} [1-\hat{y}] = 0.5 \times 0.5 = 0.25$

$H_{jk} = x_{j}^{T} \operatorname{diag} \left(\hat{y}_{O(r\hat{y})} \right) x^{k} = x_{j}^{T} D x^{k}$	•
What con use say about D^2 Fach $\hat{y} \in [0_1^2]$; $I-\hat{y} \in [0_1^2]$ Man $\hat{y} [I-\hat{y}] = 0.5 \times 0.5 = 0.25$	
Thus D is of type	•
$D = \int 0 \leq D(1 \leq 0.25) 0 0 1 \leq 0.25$	· · ·
$0 \leq 0 \leq 1$	•

Hjk =	xj diag lý Olry) x =	x ^T D	× * * * * * * *
What con w Each $\hat{y} \in [c$	$[k Say about D^{?}$ $[b_{1}i]; [-\tilde{y} \in [b_{1}i]]$
Man ý [1-ý]	= 0.5×0.5 = 0.25 diagonal matrix	Intere	diagonal
entries are	51w 0 and 0.25	· · · · · · · ·	· · · · · · · ·
. .		· · · · · · · ·	

· · · · · · · · · · · · · · · · · · ·	· · ·		· · ·	ŀ	1	j k	e e e	•	•	•		•	X	ד ג	ð	Lio	2		(¥		<u>(</u>	~ 3)))))	k X	• • •	•	- - -	· · · · · ·	7	Ĵ	T -)		,	2	•	•	•	•	· · ·	· · ·	• • •
· · · · · ·	· · ·	h	r r r		\mathbf{t}	•	•	ð	Ø	23	•	•		Н	•		l	0	k		l	٠£	2	?	· · ·	•	•	•	•	· · ·	•	•	•	•	•	•	•	•	•	•	· ·	· · ·	•
• •	• •	•	• •	•	•		•	•	•	•	•	•	•	•	•		•	•	•	•	•	•	•	•	•	•	*	•	•	• •	•	٠	•	•	•	•	•	•	0	•	• •		
• •	• •	•	• •	•	•	÷	•		•	•		•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	0	• •	•	•	•	0	•	•	•	•		•	• •		•
• •	• •		• •	۰		٠	٠		٠	•		٠	٠	٠	٠		•	۰	٠		٠			۰	٠	٠	٠	٠	٠		•	٠	٠	۰	٠	٠		•	٠	٠	• •		•
• •	• •	٠	• •	0	٠		٠	•	•	•		•	•	•	•		•	•	•			•		•	•		•	•	•		٠	٠	•	•		•	•	•	•	•			•
	• •	•		•		•	•	•	•			•	•	•	•		•	•	•	•	•	•	•	•	•	•	•	•	•		•	•	•	•	•	•		•	•	•			
				•		*	•	٠	*	•		*	٠		٠			٠	٠			•		٠		٠	*	•	٠		*	•	٠	٠		٠	•	•		*			•
• •	• •		• •	0	٠	٠	٠		٠			٠	٠	•			٠		٠		٠				•	٠	٠		0		٠	٠	٠		٠	٠		•			• •		•
• •	• •	•	• •	٠		•	•	۰	٠			٠	٠	•	٠		•	٠	٠		٠			٠	•	٠	•	•	•				٠	٠	•	٠		•		•	• •		•
• •	• •	•	• •	٠	٠	•	٠	•	•	٠		•	٠	•	•		•	٠	•		•	•		•	•	•	•	•	٠		٠	٠	•	٠		•	٠	•	•	•			•
0 0	• •	٠	• •	۰		٠	•	•	•	•	•	•	•	•	•			•	•	٠	•	٠	•	•	•	•		•	•	• •	•	٠	•	•	•	•	•	•	•	•	• •		•
• •	• •	•		•	•	•	•		•	•		•			•		•	•	•	•		•	•	•		•		•			•	•	•	•		•	•	•					
																					٠					•			0														
	• •	٠	• •	0		٠		٠	•	•		•	٠	•	•		•	٠						٠	•	•			٠			٠	٠	٠		•	٠	•		•			•
	• •		• •	٠	٠	*	٠	٠			•	٠				٠		0	٠	٠		٠		0		٠			•		٠		٠	٠			•		•	•			
			• •	•			•		•			•	٠	·	٠		•	٠						٠	·		•	·	•				•		•	•		·		•	• •		•
	• •	•	• •	٠		•	٠	٠	٠		•	٠	٠	٠	٠	٠	٠	٠	٠	•	٠	٠		٠	٠	٠	٠	٠	٠		*	•	٠	٠	٠	٠		•	•	•	• •		•
• •																																											
• •																																											
• •																																											
• •																																											

Hjk	= xj ^T diag lý OL	$(\hat{y}) \times^{k} = \chi^{j^{T}}$	$\sum_{k=1}^{n} \sum_{k=1}^{n} \sum_{k$
what doe	s H look l		. .
$H_{21} = X$. .	· ·
· · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · · · ·
· · · · · · · · · · · · · · · · · · ·			
· ·	· · · · · · · · · · · ·	· · · · · · · · · · · ·	· · · · · · · · · · · · · ·

Hjk = xj ^T diag lýol	$(r_{y})_{x}^{k} = x^{i^{T}} D x^{k}$
what does 4 look &	Ste?
$H_{II} = X^{T} D X^{I}$	
$H_{21} = \chi^2 D \chi$. .
 	· · · · · · · · · · · · · · · · · · ·

Hjk = xj ^T diag (ŷo(rŷ)) ^k	$= x^{jT} D x^{k}$
what does H look like?	. .
$H_{11} = x^{T} D x^{T}$ $H_{21} = x^{2} D x$	
= X D X	

•	· · ·	•	•	•	•	•	ŀ	-	•			•	X	Т					•	•	•	•		•	•	•		•	•	•	•	•	•	• •	· · ·	•	•	•	•	•	•	•	•	•	•	•
•	· · ·	Q		•	•		46						ł	1	• •	•	j <u>(</u>	5	•	•	P	. \$		n D N			N	rt	ri		L	•	•	• • • •	· · ·		•	•	•	•	•	•	•	•	•	•
•	· ·	•	•	•	•	•	•	•	•	•	•	•	•		0 0 0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• • • •	· · ·	•	•	•	•	•	•	•	•	•	•	•
•	· ·	•		•	•	•	•	•	•	•		•			•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	· · ·		•	•	•	•	•	•	•	•	•	•
•	· · ·	•	•	•	•	•	•	•		•		•			• • •	•	•	•	•	•	•	•		•	•	•	•	•	•	•	•	•	•	• •	· · ·		•	•	•	•	•	•	•	•	•	•
٠	· ·	٠	٠	٠	٠	٠	٠	٠		٠	٠	۰				0	٠	٠	٠	٠	0	٠	٠	٠	٠	٠	0	٠	0		٠	٠		• •		٠	÷	÷	0	٠	٠	÷	٠	÷	0	٠
•	· ·	•	•	•	•	•	•	•	•	•	•	•	•		0 0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	• •	• •	•	•	•	•	•	•	•	•	•	•	•
	· ·																																													

 		·. His	P.J.D. ⇒	J(d) is conver
· · · · · ·		dii Zi ²	where	04 dii 40.25
 			. .	· · · · · · · · · · · · · · ·
 	let Xv=Z	
 			· · · · · · · · · · · · · · ·
 	THE	ντχτρχ	\mathbf{v}	· · · · · · · · · · · · · · ·
\mathbb{Q} :	Prove H	is. P.s.D n	ntrix	· · · · · · · · · · · · · ·
· · · · · ·	· · · · · · · · · · · ·	· · · · · · · · · · · · ·	· · · · · · · · · · ·	· · · · · · · · · · · · · ·
· · · · · ·	$H = x^{T}$	$\Sigma_{\mathbf{X}}$	· · · · · · · · · ·	· · · · · · · · · · · · · · ·

Iter ativel	Remeighted least squares
I) First	uder update Rule
	$- \propto \Delta J(\vartheta)$
Typi cally	g: gradient 7J(7)
. .	H: Hessian
· · · · · · · · · · · · ·	. .

Iter atively	p Remeighte	d least	Squares	· · · · · · · · · · · · · ·
II) Second (Inder Update	Rule	. .	· · · · · · · · · · · · · ·
θ =	θĻ –	нg
· · · · · · · · · · · · · ·	· · · · · · · · · · · · ·	· · · · · · · ·	 	· · · · · · · · · · · · ·
· · · · · · · · · · · · · · ·	· · · · · · · · · · · · · ·	· · · · · · · · ·	· · · · · · · · · · ·	· · · · · · · · · · · · · ·
· · · · · · · · · · · · · · ·				
· · · · · · · · · · · · · · ·		· · · · · · · ·	· · · · · · · · · ·	· · · · · · · · · · · · ·

Ţ	teratively Remeighted least squares	•
II)	Second Order update Rule	•
· · · · · ·	$\theta_{th} = \theta_t - H^2 g$	•
Four	logistic regression: $\theta_{t+1} = \Theta_t - (x^T O x)^T x^T (\hat{y} - y)$	•
· · · · · · ·		•

Iteratively Reweighted least squares	· · ·
I) Second Order update Rule	· ·
$\theta_{t+1} = \theta_t - H^{-1}g$	· · ·
Fou logistic negression:	· ·
$\theta_{H} = \Theta_{L} - (X' O X) X (y - y)$	· ·
$= (x^{T} o x \overline{y})^{T} (x - y \theta x \overline{y} - x)^{T}$	· ·
$\theta_{t+1} = (x D x)^T x^T D [x \theta_t - \overline{0}^T (y^2 - y^2)]$	· · ·

Iteratively Remeighted le	ost squares
I) Second Order update Rul $\theta_{t+1} = (x^T \Omega x)^T x^T \Omega [x \theta]$	$e = \bar{o}(y - y)$
$= (x^T D x)^T X^T D z_t$; Zt = X+ - 0 (ŷ-y)
(ontrast w/ meighted Rinear	regression
$\hat{\theta} = (\hat{x}^T O x \hat{j} x^T D y)$. .

Iteratively Reweighted le	ost squares
I) Second Onder update Rul $\theta_{t+1} = (x^T D x)^T x^T D [x \theta]$ $= (x^T D x)^T x^T D z_t$	$e = \bar{0}(\hat{y} - y) \int_{zt}^{z} (\hat{y} - y) \int$
(ontrast w/ meighted linear	· · · · · · · · · · · · · · · · · · ·
$\hat{\theta} = (\hat{x} D x \hat{j} x D \hat{x})$. .

Iter atively	Remeighted	least squ	ares	· · · · · · · · · · · ·
II) Second Oro		ule	· · · · · · · ·	
	DxJxTDzt	;2f	= X9f-	- ō (ŷ -y)
(onthast wl	meighted Rinea	r regressit	1	· ·
	y j'x Dy	· · · · · · · · ·	· · · · · · · · ·	. .
	· · · · · · · · · · · · · ·			. .

Iteratively Remeighted best squares II) Second Onder update Rule ō' (ŷ-y) $\theta_{++1} = (x^T D x)^T X^T D [X \theta_{+}]$ = (xTDxJxTDzt ; Zt = X++ - 0 (ý-y) w meighter linear negression (onthast $D = \left[\widehat{y}_{1}(+\widehat{y}) \right]$ $\hat{\theta} = (\hat{x} \hat{D} \hat{x} \hat{J} \hat{x} \hat{D} \hat{y})$

Iteratively Remeighted bast squares II) Second Onder update Rule ō' (ŷ-y) $\theta_{++1} = (x^T D x)^T X^T D [X \theta_{+}]$ = (xTDxJxTDzt ; Zt = X++ - 0 (ý-y) w meighter linear regression $D = \left[\widehat{y_1} \left(l + \widehat{y_1} \right) \right]$ (onthast $\hat{\theta} = (\hat{x} \hat{v} \hat{v} \hat{v}) = \hat{\theta}$