Logistic Regression

Saturday, July 8, 2023 2:17 PM

	Linear Regression	Logistic Regression
Model Type	Regression	Classification
Expression	$y = b_0 + b_1x_1 + bx_2 + \cdots + b_K x_K$	$y = \frac{1}{1 + e^{-(b_0 + b_1x_1 + \cdots + b_Kx_K)}}$
Method of estimation of b's	Least Squares	Maximum Likelihood
Error function minimized	$Z = \sum (y_i - \hat{y}_i)^2$	log loss

Binary : y : 0 or 1

Multiclass : y : 0 or 1 or 2

OVR : One Vs Rest of all

Training

y	X_1 X_2 ... X_K
n_1 { 0	
0	
n_2 { 1	
1	
n_3 { 2	
2	
2	
n_4 { 3	
3	

$n_1 + n_1 + n_3 + n_4 = n$ Total training obs.

0^s Vs Rest of all except 0^s

y	X_1 .... X_K	
0		
0		
0	n_1	
1		n obs
1		
1	$n - n_1$	
1		

lr_0

1^{st} Vs Rest of all except 1^s

y	X_1 X_2 ... X_K	
1		
1	n_2	
1		
0		
0		
0		
0	$n - n_2$	
0		
0		

lr_1

lr_2 lr_3

Testing

X_1 X_2 ... X_K	$\hat{y}_{lr_0}$ (1)	$\hat{y}_{lr_1}$ (1)	$\hat{y}_{lr_2}$ (1)	$\hat{y}_{lr_3}$ (1)	$\hat{y}$-pred
	0.8	0.9	0.6	0.7	1
					0

	0.8	0.9	0.6	0.7	1
	0.4	0.2	0.3	0.1	0

y	$X_1 \cdots X_K$
0	
0	
⋮	
0	
1	
1	
⋮	
1	
2	
⋮	
z	

$$\text{Softmax} = \frac{e^{s_1}}{e^{s_1} + e^{s_2} + e^{s_3}}$$