

$X \in \mathbb{R}^{n \times p}$
 FEATURES:

- SOME ARE DATES

- MANY NAs

* 1, 2, 1 DEWON,

* " "

* - | ← (CHARACTER)

* "[]"

- SOME VALS ARE QUAL



$$\nabla \ell(\beta)$$

$$\hat{\beta} \leftarrow \hat{\beta} - \alpha \nabla \ell(\beta)$$

(SLIDE 40 IN REGULARIZATION)

$$\ell(\beta) = \sum_{i=1}^n \left(y_i \beta^T X_i - \log(1 + e^{\beta^T X_i}) \right)$$

INTRO SLIDES TO (141),

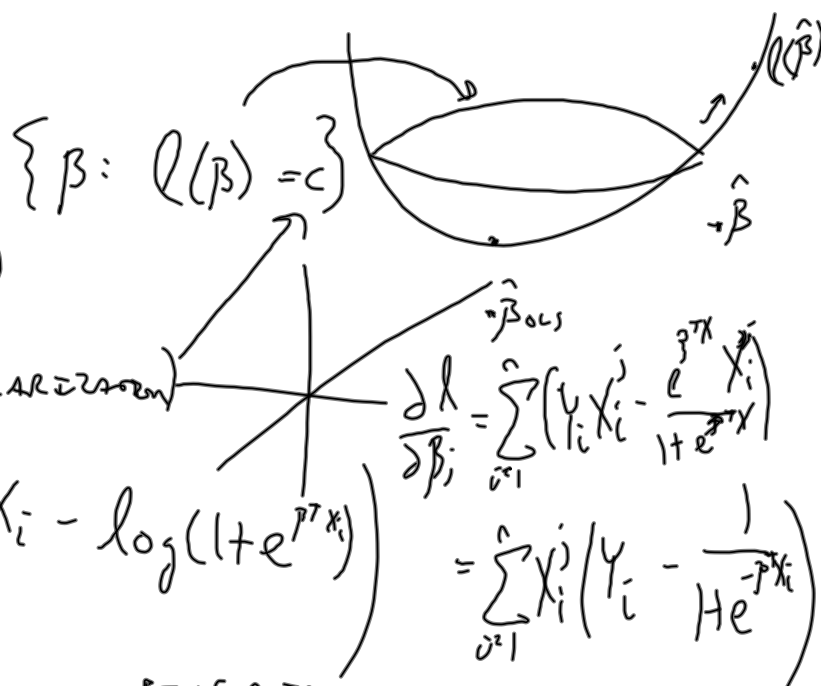
INITIALIZE $\hat{\beta} \leftarrow$ SMALL SUBSET OF DATA

1) SAVE $L = \ell(\hat{\beta})$

FOR $j = 1, \dots, P$

$$2) \hat{\beta}_j = \hat{\beta}_j + \alpha \frac{\partial \ell}{\partial \beta}$$

3) $|L - \ell(\hat{\beta}_j)| > \text{TOL}$ $\|\hat{\beta}_j - \hat{\beta}_{j-1}\| > \text{TOL}$



(BATCH) $\sum_{i=1}^n$ L.O.
 LIKELIHOOD
 DERIVATIVE

STOCHASTIC L.O.

$N = \{1, \dots, n\}$

DRAW $N \subseteq N$ S.T. N IS SMALL

$\sum_{i \in N}$



FORWARD SELECTION:

$$AIC = -\log \text{likelihood} + df$$

$$j = 1, \dots, p :$$

$$AIC(j) \quad (\text{at } j^*)$$

$$j^* j^* :$$

$$AIC(j)$$