Math 342W/642/742W

Recitation – Day #12 (3.25.25)

I. Types of Modeling

Fill in the missing cells of the table below regarding the various modeling types:

Response	**Type of Modeling**	**g returns...**	**Example $\mathcal{A}$**
$\mathcal{Y} \subset \mathbb{R}$	Regression	$\hat{y} \in \mathbb{R}$	OLS
$\mathcal{Y} \in \{C_1, \ldots, C_k\}$	Classification	$\hat{y} \in \{C_1, \ldots, C_k\}$	KNN
$\mathcal{Y} = \{0, 1\}$	Binary Classification	$\hat{y} \in \{0, 1\}$	KNN, SVM
$\mathcal{Y} = \{0, 1\}$	Probability Estimation	$\hat{p} \in (0, 1)$	Logistic Regression
$\mathcal{Y} = (0, \infty)$	Survival/Churn	$\hat{y} \in \mathbb{R}$	Weibull Regression
$\mathcal{Y} = \{0, 1, 2, \ldots\}$	Count	$\hat{y} \in \{0, 1, \ldots\}$	Poisson Regression
$\mathcal{Y} = \{C_1, C_2, \ldots, C_k\}$	Probability Estimation	$\hat{p}_i = \mathbb{P}(Y = C_i \mid \boldsymbol{x})$	Multilogistic Regression
$\mathcal{Y} = \{C_1 < C_2 < \ldots < C_k\}$	Probability Estimation	same column vector above	Proportional Odds
$\mathcal{Y} \in (0, 1)$	Regression	$\hat{y} \in (0, 1)$	Beta Regression

II. Logistic Regression

(i) What is the response space for ***logistic regression***?

$$\mathcal{Y} = \{0, 1\}$$

(ii) What type of modeling is ***logistic regression***?

Probability Estimation

(iii) What is the function f_{pr}? And what are we trying to find the "best guess" of?

$$f_{pr}(\boldsymbol{x}) : \mathbb{R}^{p+1} \to (0,1) \text{ and is the best guess of } \mathbb{P}\left(Y = 1 \mid \boldsymbol{x}\right)$$

(iv) What is the ***link function***, Φ that we will use for logistic regression?

$$\Phi : \mathbb{R} \to (0,1) \;\text{ where }\; \Phi(u) = \frac{e^u}{1+e^u} = \frac{1}{1+e^{-u}}$$

(v) What is the candidate set of functions $\mathcal{H}_{pr}$ for approximating f_{pr}?

$$\mathcal{H}_{pr} = \left\{ \frac{1}{1+e^{-\boldsymbol{x}\cdot\boldsymbol{w}}} \;\middle|\; \boldsymbol{w} \in \mathbb{R}^{p+1} \right\}$$

(vi) What is the name that we give to this model based upon the candidate set $\mathcal{H}_{pr}$?

Generalized Linear Model (GLM)

(vii) What is the algorithm for which we compute $g_{pr} = \mathcal{A}(\mathbb{D}, \mathcal{H}_{pr})$?

The algorithm is called the "*maximum likelihood*" where

$$\mathcal{A}: \quad \boldsymbol{b} = \underset{\boldsymbol{w}\in\mathbb{R}^{p+1}}{\arg\max}\{\mathbb{P}(\mathbb{D})\}$$

$$\begin{aligned}
\mathbb{P}(\mathbb{D}) &= \mathbb{P}(Y_1 = y, Y_2 = y_2, \ldots, Y_n = y_n \mid \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \\
&= \mathbb{P}(Y_1 = y_1 \mid \boldsymbol{x}_1) \cdot \mathbb{P}(Y_2 = y_2 \mid \boldsymbol{x}_2) \cdots \mathbb{P}(Y_n = y_n \mid \boldsymbol{x}_n) \\
&= \prod_{i=1}^{n} \mathbb{P}(Y_i = y_i \mid \boldsymbol{x}_i) \\
&= \prod_{i=1}^{n} (f_{pr}(\boldsymbol{x}_i))^{y_i} \cdot (1 - f_{pr}(\boldsymbol{x}_i))^{1-y_i} \\
&= \prod_{i=1}^{n} \left(\frac{1}{1+e^{-\boldsymbol{x}\cdot\boldsymbol{w}}}\right)^{y_i} \cdot \left(\frac{1}{1+e^{\boldsymbol{x}\cdot\boldsymbol{w}}}\right)^{1-y_i}
\end{aligned}$$