Lecture 10

Multinomial Logistic Regression

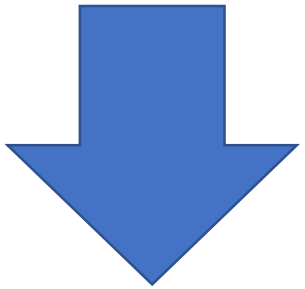
Suman Samui

What we discussed so far…

- Logistic regression model → Probabilistic model
- Where from the name ‘logit’ comes from?
- What are the important aspects to train a LR model?

Logistic Regression

- Binary classification model → probabilistic
- It can be extended to multiclass classification problem in different ways.



Softmax Regression

(Multinomial Logistic Regression)

Multinomial Logistic Regression

- Labeled dataset with N classes ($N > 2$)

- The Softmax Regression model first computes a score $s_k(x)$ for each class k

- Then it estimates the probability of each class by applying the *softmax function* (also called the *normalized exponential*) to the scores.

Softmax score for class k

$$s_k(\mathbf{x}) = \mathbf{x}^T \boldsymbol{\theta}^{(k)}$$

The scores are generally called logits or log-odds.

Note that each class has its own dedicated parameter vector $\theta^{(k)}$

$\theta^{(k)}$ → Column vector with (n+1) elements

x^T → Row vector with (n+1) elements

Softmax function

$$\hat{p}_k = \sigma(\mathbf{s}(\mathbf{x}))_k = \frac{\exp\left(s_k(\mathbf{x})\right)}{\sum_{j=1}^{K} \exp\left(s_j(\mathbf{x})\right)}$$

the estimated probability

Classes → K

sklearn.linear_model.LogisticRegression

class sklearn.linear_model.**LogisticRegression**(*penalty='l2', *, dual=False, tol=0.0001, C=1.0, fit_intercept=True, intercept_scaling=1, class_weight=None, random_state=None, solver='lbfgs', max_iter=100, multi_class='auto', verbose=0, warm_start=False, n_jobs=None, l1_ratio=None*)