
Lecture Notes for Machine Learning in Python

Professor Eric Larson
Logistic Regression

Class Logistics and Agenda

- Welcome back to lecture!
- Logistics
 - A2 Deadline
 - Next week: A3
- Agenda
 - Logistic Regression
 - Solving
 - Programming
 - Finally some real python!

Solving Logistic Regression



Setting Up Binary Logistic Regression

- From flipped lecture:

$$\hat{y} = w^T \hat{x}$$

$$p(y^{(i)} | x^{(i)}, w) = \frac{1}{1 + \exp(-w^T x^{(i)})}$$

$$p(y^{(i)} = 0 | x^{(i)}, w) = 1 - \frac{1}{1 + \exp(-w^T x^{(i)})}$$

$$L(w) = \prod_{y^{(i)}=1} p(y^{(i)}=1 | x^{(i)}, w) \prod_{y^{(i)}=0} p(y^{(i)}=0 | x^{(i)}, w)$$

$$\text{MAX } L(w)$$

$$w^* = \underset{w}{\text{ARGMAX}} L(w)$$

maximize!

Binary Solution for Update Equation

- Video Supplement (also on canvas):
 - <https://www.youtube.com/watch?v=FGnoHdjFrJ8>
- General Procedure:
 - Simplify $L(w)$ with logarithm, $l(w)$

$$l(w) = \sum_i y^{(i)} \ln(g(w^T x^{(i)})) + (1 - y^{(i)}) \ln(1 - g(w^T x^{(i)}))$$

- Take Gradient

$$= - \sum_i (y^{(i)} - g(w^T x^{(i)})) x_j^{(i)}$$

- Use gradient inside update equation for \mathbf{w}

Binary Solution for Update Equation

- Use gradient inside update equation for \mathbf{w}

$$= - \sum_i (y^{(i)} - g(\mathbf{w}^T \mathbf{x}^{(i)})) \mathbf{x}_j^{(i)}$$
$$\mathbf{w} := \mathbf{w} + \alpha \begin{bmatrix} \sum_i (y^{(i)} - g(\cdot)) \mathbf{x}_1^{(i)} \\ (\quad) \quad \quad \quad \mathbf{x}_2^{(i)} \\ \vdots \\ \sum_i (y^{(i)} - g(\cdot)) \mathbf{x}_N^{(i)} \end{bmatrix}$$

Reinvent sklearn **Logistic Regression**

Programming
Vectorization
Regularization
Multi-class extension



Other Tutorials:

<http://blog.yhat.com/posts/logistic-regression-python-rodeo.html>

http://scikit-learn.org/stable/auto_examples/linear_model/plot_iris_logistic.html

For Next Lecture

- **Next time:** More gradient based optimization techniques for logistic regression
- **Next Next time:** SVMs in-class assignment

Lecture Notes for Machine Learning in Python

Replaced with Jupyter

Professor Eric Larson
Optimization Techniques for Logistic Regression

Scratch Paper

Scratch Paper

Scratch Paper