Lecture 14 (Logistic Regression and Generalised Models)

1 Hypothesis Function

$$h_\theta(x) = \frac{1}{y_i - e^{-\theta^T x}}$$

We use *gradient ascent* instead of *gradient descent* since

$$LL(\theta)$$

is concave.

2 Generalised Linear Models

We compute $P(y_i; \eta)$ where η is a function of θ, x_i and $y_i\eta$ belongs to exponential family distribution.

$$P(y_i; \eta) = b(y) \exp\left(\eta y - a(\eta_i)\right)$$

(HW: Prove Bernoulli and Normal are special cases of the above equation) The log-liklihood is given by:

$$LL(\eta) = \sum_{i=1}^{m} \left(\log(b(y_i)) + \eta y_i - a(\eta_i)\right)$$

$$\implies \nabla_\theta(LL(\eta)) = \sum_{i=1}^{m} \nabla_\theta(\eta_i)\left(y_i - a'(\eta_i)\right)$$

Following assumptions are made:

1. η_i is a linear function of x_i
2. $h_\theta(x) = E[y|x;\theta] = g(\eta)$
3. $g^{-1}(\phi)$ is linearly dependent on x