CS 229 PS7 Q2

Chwan-Hao TUNG
861052182
12/1/2016
CS229 PS7 Q2

part a.

As a member of the exponential family:

$$\begin{aligned}
p(y) &= \binom{y+r-1}{y} \cdot q^y \cdot (1-q)^r \\
&= \binom{y+r-1}{y} \cdot e^{y \log q + r \log (1-q)} \\
&= h(y) \cdot e^{\theta y - A(\theta)}
\end{aligned}$$

where

$$\begin{aligned}
h(y) &= \binom{y+r-1}{y} \\
\phi(y) &= y \\
\theta &= \log (q) \\
A(\theta) &= -r \log (1 - e^{\theta})
\end{aligned}$$

The relation between θ and q is described as $\theta = \log (q)$

part b.

$$\begin{aligned}
E[y|x] &= \Delta_{\theta} A(\theta) \\
&= -r \frac{d}{d\theta} \log (1 - e^{\theta}) \\
&= \frac{re^{\theta}}{1 - e^{\theta}} = \psi^{-1}(\theta) = \mu
\end{aligned}$$

$$\begin{aligned}
\mu &= \frac{re^{\theta}}{1 - e^{\theta}} \\
\theta &= \log (\frac{\mu}{r + \mu})
\end{aligned}$$

The canonical link function is therefore,

$$g(\mu) = \psi(u) = \theta = \log\left(\frac{\mu}{r+\mu}\right)$$

Given learned parameters w and x, to predict q we have,

$$\begin{aligned}
\theta(x) &= x^T w \\
\theta &= \log(q)
\end{aligned}$$

Therefore,

$$\begin{aligned}
q &= e^{\theta} \\
q &= e^{x^T w}
\end{aligned}$$

part c.

$$L(w) = \sum_i (x_i{}^T w y_i - A(x_i{}^T w))$$

The Newton-Paphson update for updating w, the learned parameters is as follows,

$$\begin{aligned}
w &\leftarrow w - \frac{\Delta_w L(w)}{\Delta\Delta_w L(w)} \\
w &\leftarrow w - \sum_i \frac{x_i y_i - \frac{r x_i e^{x_i w}}{1-e^{x_i w}}}{\frac{-r x_i{}^2 e^{x_i w}}{(1-e^{x_i w})^2}}
\end{aligned}$$