PASS-GLM for Bernoulli and Poisson

Yoel Sanchez Araujo

1 Polynomial approximation to e^z

First we need to arrive at an approximation of the exponential function: e^{Xw}. We approximate it with a set of chebyshev polynomials of the form:

$$\begin{aligned} e^{Xw} \approx & c_0 + c_1\Big(\sum_k \mathbf{x}_k^T\mathbf{w}\Big) + c_2\Big(\sum_k \mathbf{x}_k^T\mathbf{w}\Big)^2 \\ & = c_0 + c_1X\mathbf{w} + \mathbf{w}^T\Big(c_2\sum_k \mathbf{x}_k\mathbf{x}_k^T\Big)\mathbf{w} \\ & = c_0 + c_1X\mathbf{w} + \mathbf{w}^T M\mathbf{w} \end{aligned} \tag{1}$$

2 Bernoulli approximation

Starting with the familiar log-likelihood for the Bernoulli GLM:

$$\mathcal{L} = \mathbf{y^T}log\Big[\sigma(x^Tw)\Big] + (\mathbf{1} - \mathbf{y}^T)log\Big[1 - \sigma(x^Tw)\Big] \tag{2}$$

we set $M = c_2\sum_k \mathbf{x}_k\mathbf{x}_k^T$, and set:

$$\begin{aligned} \sigma(Xw) &= \frac{e^{x^Tw}}{1+e^{x^Tw}} \\ 1 - \sigma(Xw) &= \frac{1}{1+e^{x^Tw}} \end{aligned} \tag{3}$$

if we plug this back into the expression for the log-likelihood and simplify we get:

$$\begin{aligned} \mathcal{L} &= \mathbf{y^T}log\Big[\sigma(x^Tw)\Big] + (\mathbf{1} - \mathbf{y}^T)log\Big[1 - \sigma(x^Tw)\Big] \\ &= \mathbf{y^T}log\Big[\frac{e^{x^Tw}}{1+e^{x^Tw}}\Big] + (\mathbf{1} - \mathbf{y}^T)log\Big[\frac{1}{1+e^{x^Tw}}\Big] \\ &= \mathbf{y^T}X\mathbf{w} - \mathbf{y^T}log\Big[1+e^{x^Tw}\Big] - log\Big[1+e^{x^Tw}\Big] + \mathbf{y^T}log\Big[1+e^{x^Tw}\Big] \end{aligned} \tag{4}$$

the following approximation is made: $log\Big[1 + e^{X\mathbf{w}}\Big] \approx c_0 + c_1 X\mathbf{w} + \mathbf{w}^T M\mathbf{w}$. This gives the approximation of the log-likelihood as:

$$\mathcal{L}_{approx} = \mathbf{y}X\mathbf{w} - c_0 - c_1 X\mathbf{w} - \mathbf{w}^T M\mathbf{w} \tag{5}$$

manipulating the log terms and multiplying the terms involving $\mathbf{y}$ out we get: taking the derivative, setting it to zero and solving for $\mathbf{w}$ we get:

$$\begin{aligned} &\frac{\partial}{\partial \mathbf{w}}\Bigg(\mathbf{y}X\mathbf{w} - c_0 - c_1 X\mathbf{w} - \mathbf{w}^T M\mathbf{w}\Bigg) \\ &0 = \mathbf{y}X - c_1 X - 2M\mathbf{w} \\ &\hat{\mathbf{w}}_{approx} = \frac{1}{2}M^{-1}X\Big(\mathbf{y} - c_1\Big) \end{aligned} \tag{6}$$

3 Poisson approximation

As before we start with the log-likelihood of the relevant family of models:

$$\mathcal{L} = \sum_k y_k \times log\big[e^{\mathbf{x}_k^T\mathbf{w}}\big] - e^{\mathbf{x}_k^T\mathbf{w}} \tag{7}$$

simplifying this and adding on our approximation of e^z as shown above we have get the approximation to the log-likelihood:

$$\mathcal{L}_{approx} = \mathbf{y}X\mathbf{w} - c_0 - c_1 X\mathbf{w} - \mathbf{w}^T M\mathbf{w} \tag{8}$$

taking the derivative, setting it to zero and solving for $\mathbf{w}$ we get:

$$\begin{aligned} &\frac{\partial}{\partial \mathbf{w}}\Bigg(\mathbf{y}X\mathbf{w} - c_0 - c_1 X\mathbf{w} - \mathbf{w}^T M\mathbf{w}\Bigg) \\ &0 = \mathbf{y}X - c_1 X - 2M\mathbf{w} \\ &\hat{\mathbf{w}}_{approx} = \frac{1}{2}M^{-1}X\Big(\mathbf{y} - c_1\Big) \end{aligned} \tag{9}$$

4 Finding the coefficients

Solve the following least-squares problem, an m'th order polynomial over a grid of points defined by:

$$\begin{aligned} &T_0 = \mathbf{1}_N \\ &T_1 = \mathbf{x} \\ &T_{n+1} = 2 \times \mathbf{x} \times T_n - T_{n-1} \end{aligned} \tag{10}$$

solve for $\mathbf{c}$:

$$e^{\mathbf{x}} = \mathbf{Tc} \tag{11}$$

there is some more thing to take into account, which is that the polynomials in $\mathbf{T}$ are weighted by the function $w(\mathbf{x}) = \frac{1}{1-\sqrt{\mathbf{x}^2}}$