7) $M(x, \vec{\theta}) = \dfrac{\theta_0}{\theta_1 + e^{-\theta_2 x}}$, $x^2(\vec{\theta}) = \sum_{i=1}^{n} \left(y_i - M(x_i, \vec{\theta})\right)^2$

c) Las derivadas parciales de la métrica estan dadas por:

$$\frac{\partial(x^2(\vec{\theta}))}{\partial\theta_i} = \sum_{i=1}^{n} 2\left(y_i - M(x_i, \vec{\theta})\right)\left(-\frac{\partial(M(x_i,\vec{\theta}))}{\partial\theta_i}\right)$$

En el paso anterior aplicamos la regla de la cadena $z = f(g(x)) \Rightarrow z' = f'(g(x))\, g'(x)$.

Reescribiendo tenemos:

$$\frac{\partial(x^2(\vec{\theta}))}{\partial\theta_j} = -2\sum_{i=1}^{n} \left(y_i - M(x_i, \vec{\theta})\right)\frac{\partial(M(x_i,\vec{\theta}))}{\partial\theta_j}$$

d) Vectorialmente, el descenso del gradiente esta dado por:

Primero, vease que es posible generalizar las derivadas parciales mostradas anteriormente de la siguiente forma:

$$\vec{\nabla} x^2(\vec{\theta}) = \begin{pmatrix} \dfrac{\partial(x^2(\vec{\theta}))}{\partial\theta_0} \\ \dfrac{\partial(x^2(\vec{\theta}))}{\partial\theta_1} \\ \dfrac{\partial(x^2(\vec{\theta}))}{\partial\theta_2} \end{pmatrix} = \begin{pmatrix} -2\sum_{i=1}^{n}(y_i - M(x_i,\vec{\theta}))\dfrac{d(M(x_i,\vec{\theta}))}{d\theta_0} \\ -2\sum_{i=1}^{n}(y_i - M(x_i,\vec{\theta}))\dfrac{d(M(x_i,\vec{\theta}))}{d\theta_1} \\ -2\sum_{i=1}^{n}(y_i - M(x_i,\vec{\theta}))\dfrac{d(M(x_i,\vec{\theta}))}{d\theta_2} \end{pmatrix}$$

$$\vec{\nabla} x^2(\vec{\theta}) = -2\sum_{i=1}^{n}\left(y_i - M(x_i,\vec{\theta})\right)\nabla_{\theta} M(x_i,\vec{\theta})$$

Ahora consideramos la definición del descenso del gradiente y sustituimos lo mostrado anteriormente

$$\vec{\theta}^{(K+1)} = \vec{\theta}^{K} - \gamma \vec{\nabla} \chi^2(\vec{\theta})$$

Entonces:

$$\vec{\theta}^{(K+1)} = \vec{\theta}^{K} - \gamma\left(-2\sum_{i=1}^{n}\left(y_i - M(x_i,\vec{\theta})\right)\vec{\nabla}_{\theta} M(x_i,\vec{\theta})\right)$$

donde, de lo mostrado anteriormente, $\vec{\nabla}_{\theta} M(x_i,\vec{\theta})$ es:

$$\nabla_{\theta} M(x_i,\vec{\theta}) = \left[\frac{\partial M(x_i,\vec{\theta})}{\partial \theta_0}, \frac{\partial M(x_i,\vec{\theta})}{\partial \theta_1}, \frac{\partial M(x_i,\vec{\theta})}{\partial \theta_2}\right]$$

Además, nótese que $\nabla_{\theta} M(x_i,\vec{\theta})$ es el transpuesto del Jacobiano para que vectorialmente todas las operaciones tengan sentido dimensional.

Esto último concuerda con la otra definición del descenso del gradiente:

$$\vec{x}^{n+1} = \vec{x}^{n} - \gamma \mathbb{J}^{T}(\vec{x}^{n}) \cdot \vec{G}(\vec{x}^{n})$$