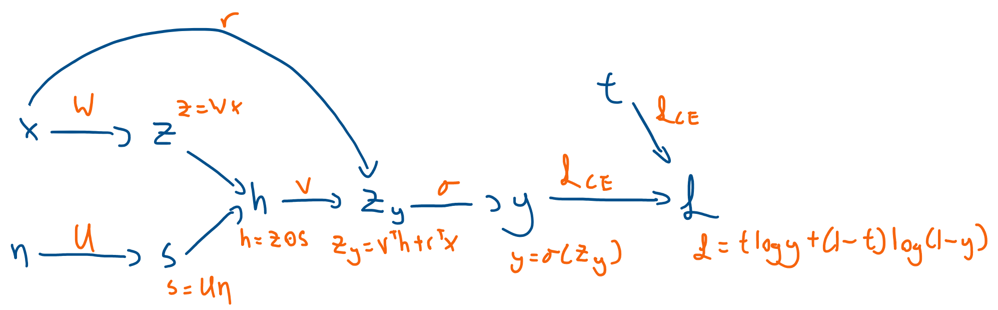
2.3

Sunday, March 3, 2024 1:42 PM

1. Since $z \in \mathbb{R}^c$, then $s \in \mathbb{R}^c$
   Since $s \in \mathbb{R}^c$, then $s = U\eta = R^{c \times b} R^b = R^c$
   Since $z \in \mathbb{R}^{c'}$, $z = Wx = \mathbb{R}^{c \times a} R^a$
   Since $h \in \mathbb{R}^c$, $v \in \mathbb{R}^c$
   Since $x \in \mathbb{R}^a$, $r \in \mathbb{R}^a$

   $W \in \mathbb{R}^{c \times a}$, $U \in \mathbb{R}^{c \times b}$, $v \in \mathbb{R}^c$, $r \in \mathbb{R}^a$

2. 

3. $\bar{\mathcal{L}} = 1$
   $\bar{y} = \bar{\mathcal{L}} \frac{\partial \mathcal{L}}{\partial y} = \frac{t-y}{y(1-y)}$

   $\bar{z}_y = \bar{y} \frac{\partial y}{\partial z_y} = \left(\frac{t-y}{y(1-y)}\right) \sigma'(z_y) = \left(\frac{t-y}{y(1-y)}\right) (\sigma(z_y)(1-\sigma(z_y)))$

   $\bar{h} = \bar{z}_y \frac{\partial z_y}{\partial h} = \left(\frac{t-y}{y(1-y)}\right) (\sigma'(z_y)) (v)$ * $\sigma'(z_y) = (\sigma(z_y)(1-\sigma(z_y)))$

   $\bar{v} = \bar{z}_y \frac{\partial z_y}{\partial v} = \left(\frac{t-y}{y(1-y)}\right) \sigma'(z_y) h$

   $\bar{r} = \bar{z}_y \frac{\partial z_y}{\partial r} = \left(\frac{t-y}{y(1-y)}\right) \sigma'(z_y) x$

   $\bar{z} = \bar{h} \frac{\partial h}{\partial z} = \left(\frac{t-y}{y(1-y)}\right) (\sigma'(z_y)) (v) \odot \mathrm{diag}(s)$ * $z \odot s = \mathrm{diag}(z) s = \mathrm{diag}(s) z$

   $\bar{s} = \bar{h} \frac{\partial h}{\partial z} = \left(\frac{t-y}{y(1-y)}\right) (\sigma'(z_y)) (v) \odot \mathrm{diag}(z)$

   $\bar{\eta} = \bar{s} \frac{\partial s}{\partial \eta} = (U)^T \left(\frac{t-y}{y(1-y)}\right) (\sigma'(z_y)) (v) \odot \mathrm{diag}(z)$

   $\bar{U} = \bar{s} \frac{\partial s}{\partial U} = \left(\frac{t-y}{y(1-y)}\right) (\sigma'(z_y)) (v) \odot \mathrm{diag}(z) (\eta)^T$

   $\bar{W} = \bar{z} \frac{\partial z}{\partial W} = \left(\frac{t-y}{y(1-y)}\right) (\sigma'(z_y)) (v) \odot \mathrm{diag}(s) x^T$

   $\bar{x} = \bar{z} \frac{\partial z}{\partial x} + \bar{z}_n \frac{\partial z_y}{\partial x} = W^T \left(\frac{t-y}{y(1-y)}\right) (\sigma'(z_y)) (v) \odot \mathrm{diag}(s) + \left(\frac{t-y}{y(1-y)}\right) \sigma'(z_y) r$