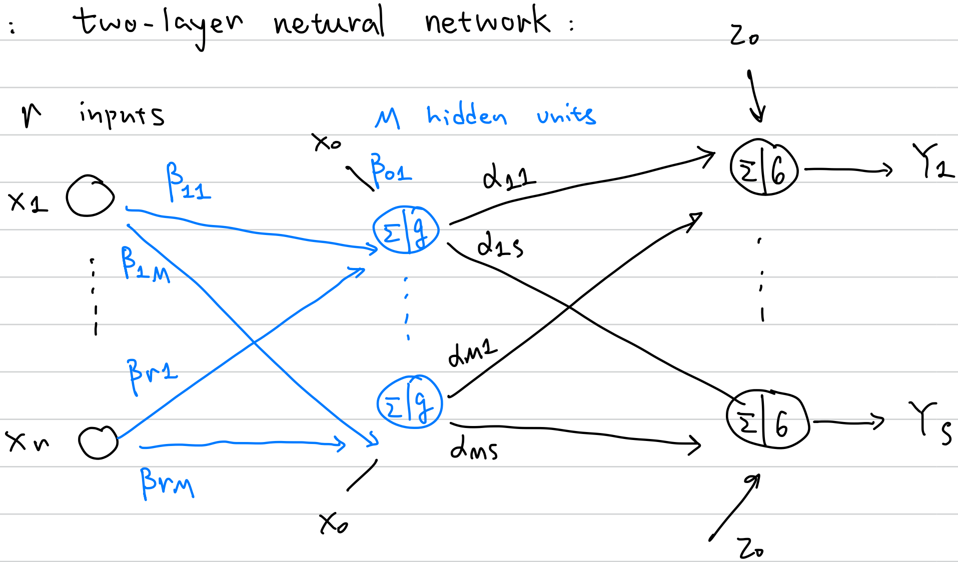
Question 1

$$f(u) = \frac{e^{u} - e^{-u}}{e^{u} + e^{-u}} = \frac{-(e^{u} + e^{-u}) + 2e^{u}}{e^{u} + e^{-u}} = -1 + \frac{2e^{u}}{e^{u} + e^{-u}}$$

$$= -1 + \frac{2}{1 + e^{-2u}} = -1 + 2g(2u).$$

2): two-layer netural network :



σ - linear fn.

The output of node k for the network with hidden layer activation function f has the form:

$$y_k^f = \alpha_{0k}^f + \sum_{j=1}^{M} \alpha_{jk}^f f\Big(\underbrace{\beta_{0j}^f + \sum_{i=1}^{r} \beta_{ij}^f x_i}_{h_j^f}\Big)$$

$$= \alpha_{0k}^f + \sum_{j=1}^{M} \alpha_{jk}^f f(h_j^f)$$

$$\Rightarrow y_k^f = \alpha_{0k}^f + \sum_{j=1}^{M} \alpha_{jk}^f \left[2g(2h_j^f) - 1 \right]$$

$$= \sum_{j=1}^{M} 2\alpha_{jk}^f g(2h_j^f) + \left[\alpha_{0k}^f - \sum_{j=1}^{M} \alpha_{jk}^f \right]$$

The output node k for the network with hidden layer activation function g has the form:

$$y_k^g = \alpha_{0k}^g + \sum_{j=1}^{M} \alpha_{jk}^g \, g(h_j^g)$$

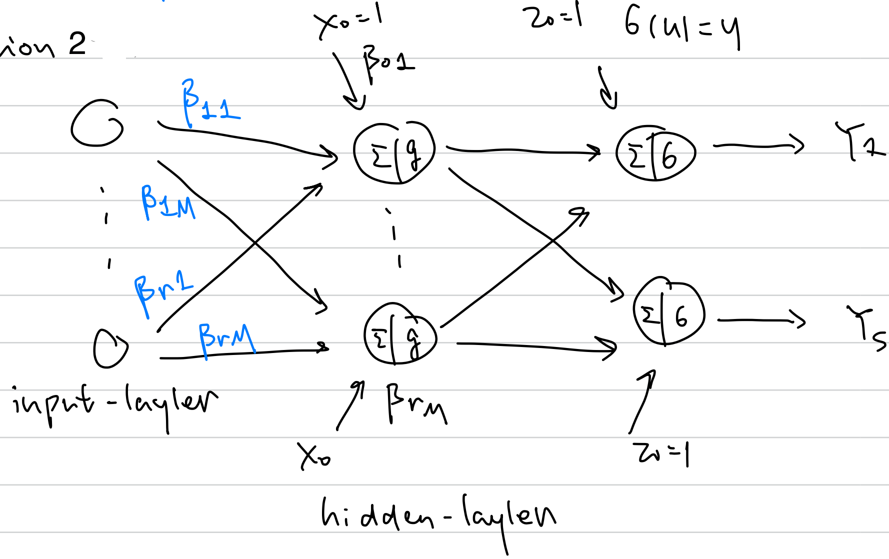
$$\beta_{0j}^g + \sum_{i=1}^{m} \beta_{ij}^g x_j$$

For output is equivalent $\Leftrightarrow$ $y_k^f = y_k^g$

$$\Rightarrow \begin{cases} h_j^g = 2h_j^f \\ \alpha_{jk}^g = 2\alpha_{jk}^f \\ \alpha_{k0}^g = \alpha_{k0}^f - \sum_{j=1}^{M} \alpha_{jk}^f \end{cases} \Longrightarrow \beta_{0j}^g = 2\beta_{0j}^f, \quad \beta_{ij}^g = 2\beta_{ij}^f$$

$\beta \approx 0.$

Question 2



For k^{th} hidden unit,

$$g\left(\beta_{0k} + \sum_{i=1}^{r} \beta_{ik} X_k\right) = g(\beta_{0k} + X^T \beta_k)$$

$$= \frac{1}{1 + \exp\{-\beta_{0k} - x^T\beta_k\}} = \frac{\exp\{\beta_{0k}\}}{\exp\{\beta_{0k}\} + \exp\{-x^T\beta_k\}}.$$

where $\beta_k = (\beta_{1k}, \ldots, \beta_{rk})$.

$$= \frac{\dfrac{\exp\{\beta_{0k}\}}{1+\exp\{\beta_{0k}\}}}{\dfrac{\exp\{\beta_{0k}\} + \exp\{-x^T\beta_k\} + 1 - 1}{1+\exp\{\beta_{0k}\}}} = \frac{\exp\{\beta_{0k}\}}{1+\exp\{\beta_{0k}\}} \left(\frac{1}{1 + \dfrac{\exp\{-x^T\beta_k\} - 1}{1+\exp\{\beta_{0k}\}}} \right)$$

if $-\beta_k^T x \to 0 \Rightarrow \exp\{-x^T\beta_k\} - 1 \to -x^T\beta_k$

$$\Rightarrow \left(\frac{1}{\dfrac{\exp\{-x^T\beta_k\} - 1}{1+\exp\{\beta_{0k}\}} + 1} \right) \to \left(\frac{1}{\dfrac{-x^T\beta_k}{1+\exp\{\beta_{0k}\}} + 1} \right) = \left(\frac{1}{1 - \dfrac{x^T\beta_k}{1+\exp\{\beta_{0k}\}}} \right)$$

By using geometric expansion $\frac{1}{1-x} \approx 1 + x$, if $|x| \leq 1$.

if $x^T\beta_k \to 0 \Rightarrow \left| \dfrac{x^T\beta_k}{1+\exp\{\beta_{0k}\}} \right| < 1$

$$\Rightarrow \left(\frac{1}{1 - \dfrac{x^T\beta_k}{\exp\{\beta_{0k}\}}} \right) \approx 1 + \frac{x^T\beta_k}{\exp\{\beta_{0k}\}}$$

$$\Rightarrow g(\beta_{0k} + x^T\beta_k) \approx \underbrace{\frac{\exp\{\beta_{0k}\}}{1+\exp\{\beta_{0k}\}} \left(1 + \frac{x^T\beta_k}{\exp\{\beta_{0k}\}} \right)}_{\text{linear f'n of } x.}$$