Ex 02

$$\tilde{z}_l = z_{l-1} \cdot B_l + b_l$$

$$\tilde{z}_l = (z_{l-2} \cdot B_{l-1} + b_{l-1}) \cdot B_l + b_l$$

$$\tilde{z}_l = z_{l-2} \cdot B_{l-1} \cdot B_l + b_{l-1} \cdot B_l + b_l$$

$$\tilde{z}_l = (z_{l-3} \cdot B_{l-2} + b_{l-2}) \cdot B_{L-1} \cdot B_L + b_{l-1} \cdot B_L + b_L$$

$$\tilde{z}_l = z_{l-3} \cdot B_{l-2} \cdot B_{l-1} \cdot B_L + b_{L-2} \cdot B_{L-1} \cdot B_L + b_{L-1} \cdot B_L + b_L$$

$$\tilde{z}_l = z_0 \times B_1 \cdot B_2 \cdots B_L + (b_1 \cdot B_2 + b_2 + \cdots b_{L-2} \cdot B_{L-1} \cdot b_{L-1}) B_L + b_L$$

$$\tilde{z}_l = x \cdot (B_1 \cdots B_L) + (b_1 B_2 + b_2 + \cdots b_{L-2} \cdot B_{L-1} \cdot b_{L-1}) \cdot B_L + b_L$$

$$\tilde{z}_l = x \cdot \boxed{W} + \boxed{b}\, B_L + b_L$$

$\because$ ϕ_L is identity Function

z_L can be written as:

$$z_L = X \cdot \boxed{W} + \boxed{b} B_L + b_L$$

$\therefore$ if ϕ_L is identity Function, any network (depth > 1) is equivalent to a 1-layer neural network.