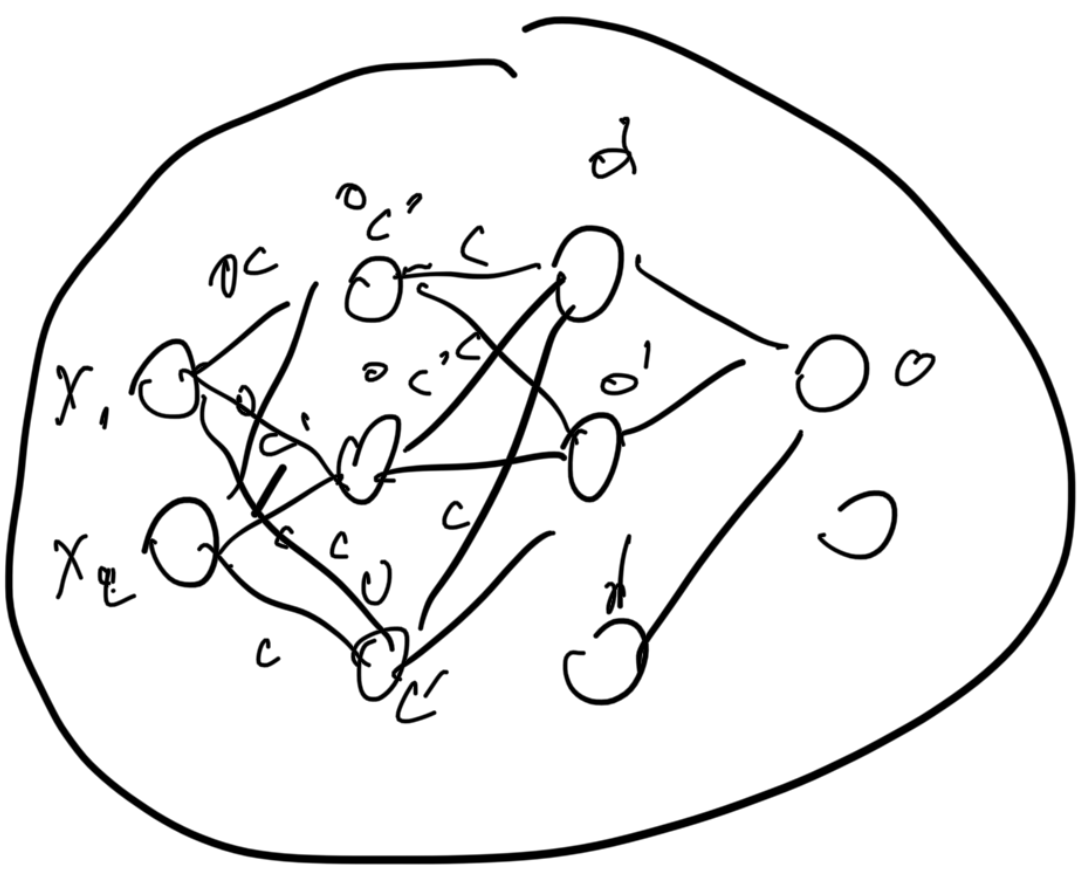
$$y = Xw$$

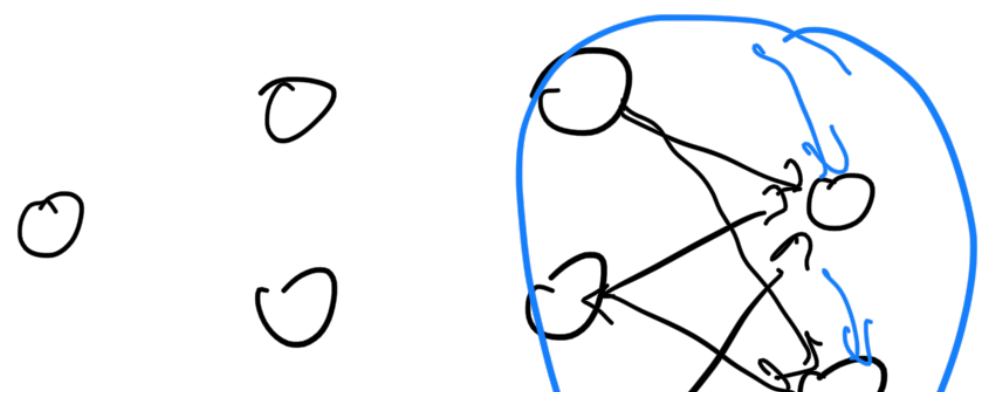
$$X^{-1}y = \underbrace{X^{-1}X}_{I}w = w$$

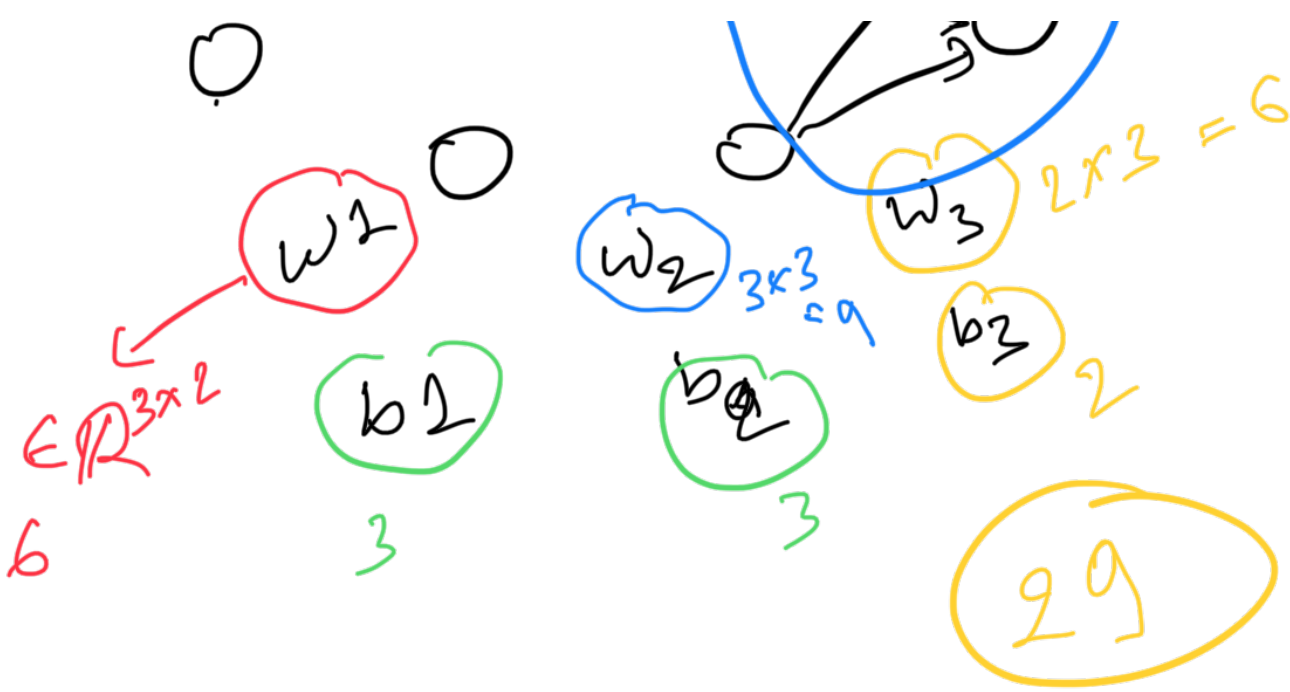
$$X^T y = X^T X w$$

$$w^* = (X^T X)^{-1} X^T y$$

Normal Equation



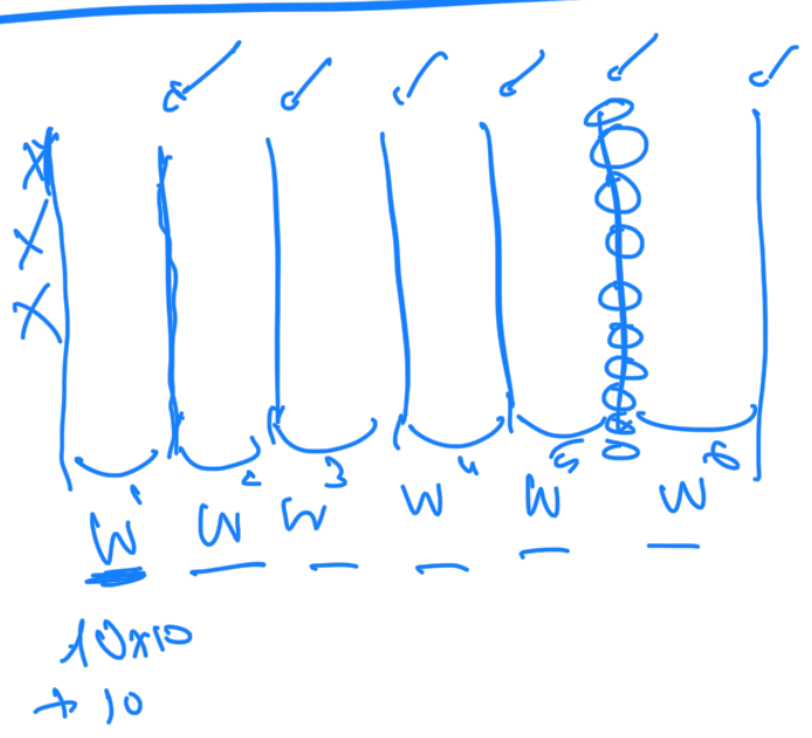




Input layer with $n_i = n_x = 10$

5 hidden layers $n_h = 10$

output layer with $n_o = 10$



$(10 \times 10 + 10) \times 6 = 660$

} bach $n_y = 10$

$$[5 \times 10 + 10] +$$
$$[10 \times 10 + 10] \times 2 +$$
$$(5 \times 10 + 5)$$

layer1 = nn.Linear (input_dim, hidden_dim)

layer2 = " " (hidden, output_dim)