hw5

Part 1

$$\frac{\partial J}{\partial w} = \frac{\partial J}{\partial M_1}\frac{\partial M_1}{\partial w}$$
$$= -(y-\hat{y}^T)\frac{\partial O}{\partial M_1}\frac{\partial M_1}{\partial w}$$
$$= -(y-\hat{y}^T)f^{'}(M_1U+b_2)diag(U)f^{'}(Wx_{j:j+h-1}+b_1)X^T_{j:j+h-1}$$