Problem 1.

In the lecture, we see that the derivative of $y = f(z)$ is needed for backpropagation. Derive $\frac{\partial y}{\partial z}$ for the sigmoid function, i.e. $y = \frac{1}{1+e^{-z}}$.

Problem 2.

Give the matrix notation of backpropagation that processes a batch of samples at a time. In other words, consider the forward pass

$$Z = XW + repmat(b^{\top}, M, 1)$$
$$Y = f(Z)$$

where $Y,\ Z \in \mathbb{R}^{M\times N_y}$, $X \in \mathbb{R}^{M\times N_x}$, $W \in \mathbb{R}^{N_x\times N_y}$, $b \in \mathbb{R}^{N_y}$. Assume $\frac{\partial J}{\partial Y} \in \mathbb{R}^{M\times N_y}$ is known. How can we compute $\frac{\partial J}{\partial X}$, $\frac{\partial J}{\partial W}$, and $\frac{\partial J}{\partial b}$?