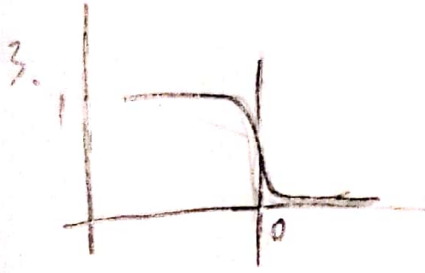Section 7.

1. Large z means the predicted value and y has the similar direction, small z means the opposite direction.

2. 1) $L = \frac{1}{2} Tr\{(X\theta - Y)^T (X\theta - Y)\}$.

3.

4.