

1) a) learning function: $h_{\theta}(X) = X\theta$, Lost/Cost: $J(\theta) = 1/2m(X\theta - y)^T(X\theta - y)$, Learning goal: $\min_{\theta} J(\theta)$

b) $\theta = (X^T X)^{-1} X^T y$

2) a) $\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} 3 \end{bmatrix}$
 $\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix}$

b) $\begin{bmatrix} 6 & -3 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \end{bmatrix}$
 $\begin{bmatrix} 1 & 7 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \end{bmatrix}$
 $\begin{bmatrix} -3 & 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \end{bmatrix}$
 $\begin{bmatrix} 0 & 7 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 5 & 0 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 25 \end{bmatrix}$
 $\begin{bmatrix} 0 & 15 & 20 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 5 \end{bmatrix}$
 $\begin{bmatrix} 1 & -50 & 160 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 22 \end{bmatrix}$
 $\begin{bmatrix} 1 & 0 & -120 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 52 \end{bmatrix}$