2. (1) $\nabla f = \begin{bmatrix} 2x_1(x_1^2 - x_2) + x_1 - 1 \\ x_2 - x_1^2 \end{bmatrix} = 0 \Rightarrow$ critical point : $[1,1]^T$

$Hf = \begin{bmatrix} 6x_1^2 - 2x_2 + 1 & -2x_1 \\ -2x_1 & 1 \end{bmatrix}$, At $[1,1]$, $Hf = \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}$.

$(5-\lambda)(1-\lambda) - 4 = 0 \Rightarrow \lambda = 3 \pm 2\sqrt{2}$, $[1,1]^T$ is local minimum,

Since $f(\vec{x}) \geq 0$ and $f([1,1]^T) = 0$, $[0,0]^T$ is global minimum.

(2) $\vec{x}_0 = [2,2]^T$

$\nabla f(\vec{x}_0) = \begin{bmatrix} 9 \\ -2 \end{bmatrix}$ $Hf(\vec{x}_0) = \begin{bmatrix} 21 & -4 \\ -4 & 1 \end{bmatrix}$

$Hf(\vec{x}_0)\vec{s}_0 = -\nabla f(\vec{x}_0)$

$\vec{s}_0 = [-\frac{1}{5}, \frac{6}{5}]^T$

$\vec{x}_1 = \vec{x}_0 + \vec{s}_0 = [1.8, 3.2]^T$

(3) $f(\vec{x}_0) = 2.5$, $f(\vec{x}_1) = 0.3208$. Compared with $f(\vec{x}_0)$, $f(\vec{x}_1)$ is much closer to minimum (0). It's a good step.

(4) If compared with $f(\vec{x}_{n-1})$, $f(\vec{x}_n)$ is far away from the minimum, it's a bad step.