④ Let $f(x,y) = x^2 + 2y^2$. Starting with $x_1 = (1,1)$

a) What is the minimal function value that can be achieved with one step of the gradient descent, i.e. find the minimum of $f(x_2)$.

Gradient descent: $\boxed{x_{k+1} = x_k - \gamma \nabla f(x_k)}$

$$\boxed{\nabla f(x) = \begin{bmatrix} 2x \\ 4y \end{bmatrix}}$$

$$x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \gamma \cdot \begin{bmatrix} 2x_1 \\ 4y_1 \end{bmatrix} = \underline{\underline{\begin{bmatrix} 1-2\gamma \\ 1-4\gamma \end{bmatrix}}}$$

$f(x_2) = (1-2\gamma)^2 + 2(1-4\gamma)^2$  To find the minima, take derivative and set to

$$\frac{\partial f(x_2)}{\partial \gamma} = 2(1-2\gamma)\cdot(-2) + 4(1-4\gamma)\cdot(-4) = 0$$

$$-4 + 8\gamma - 16 + 64\gamma = 0$$

$$20 = 72\gamma$$

$$\underline{\underline{\gamma = \frac{5}{18}}}$$

$$f(x_2) = \left(1 - \frac{10}{18}\right)^2 + 2\left(1 - \frac{20}{18}\right)^2 = \frac{16}{81} + \frac{2}{81} = \underline{\underline{\frac{2}{9}}}$$

b) How close to the actual minimum $x^*$ of function $f$ can we get with one step of the gradient descent, i.e., find minimum of the distance from $x^*$ to $x_2$.

$\boxed{x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}}$

We need to find $\min \|x_2 - x^*\| = \min \|x_2\| \overset{\text{monotone increasing}}{=} \underline{\underline{\min \|x_2\|^2}}$

$$\boxed{\|x_2\|^2 = (1-2\gamma)^2 + (1-4\gamma)^2}$$

$$\frac{\partial \|x_2\|^2}{\partial \gamma} = 2(1-2\gamma)\cdot(-2) + 2(1-4\gamma)\cdot(-4) = 0$$

$$-4 + 8\gamma \cancel{-} \cancel{4}\ 8 + 32\gamma = 0$$

$$12 = 40\gamma$$

$$\underline{\underline{\gamma = \frac{3}{10}}}$$

$$\|x_2 - x^*\| = \|x_2\| = \sqrt{\left(1-\frac{6}{10}\right)^2 + \left(1-\frac{12}{10}\right)^2}$$

$$= \sqrt{\frac{16}{100} + \frac{4}{100}} = \sqrt{\frac{20}{100}} = \underline{\underline{\sqrt{\frac{1}{5}}}} = \underline{\underline{\frac{1}{\sqrt{5}}}}$$