(2) Determine the optimal learning rates γ, μ for the Polyak GD for function

$$f(x,y,z) = x^2 + 2y^2 - 2yz + 4z^2 + 3x - 4y + 5z$$

Let's check the eigenvalues of the Hessian to get optimal α and β parameters.

$$\nabla f = \begin{bmatrix} 2x+3 \\ 4y-2z-4 \\ -2y+8z+5 \end{bmatrix} \qquad \nabla^2 f = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & -2 \\ 0 & -2 & 8 \end{bmatrix}$$

Find eigenvalues: $\det(\nabla^2 f - \lambda I) = \begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & 4-\lambda & -2 \\ 0 & -2 & 8-\lambda \end{vmatrix} = (2-\lambda)\left((4-\lambda)\cdot(8-\lambda)-(+4)\right) = 0$

$$(2-\lambda)(32-8\lambda-4\lambda+\lambda^2-4)=0$$

$$\lambda_1 = 2 \qquad (\lambda^2 - 12\lambda + 28) = 0$$

$$\lambda_{2,3} = \frac{12 \pm \sqrt{144-112}}{2} = \frac{12 \pm \sqrt{32}}{2} = \frac{12 \pm 4\sqrt{2}}{2} = 6 \pm 2\sqrt{2}$$

$\alpha = 2$, $\beta = 6 + 2\cdot\sqrt{2}$. We found out that f is 2-strongly convex and $6+2\cdot\sqrt{2}$ - smooth.

USED NESTEROV by MISTAKE

Since f is strongly convex we use THEOREM 5.3 from lecture notes:

FOR:

$$\gamma = \frac{1}{\beta}, \quad \mu = \frac{\sqrt{\varkappa}-1}{\sqrt{\varkappa}+1}, \quad \varkappa = \frac{\beta}{\alpha}: \qquad \boxed{f(x_{k+1}) - f(x^*) \le \frac{\alpha+\beta}{2}\left(\frac{\sqrt{\varkappa}-1}{\sqrt{\varkappa}}\right)^k \|x_1 - x^*\|^2}$$

$$\gamma = \frac{1}{6+2\cdot\sqrt{2}} \approx 2.97, \quad \mu = \frac{\sqrt{\frac{6+2\cdot\sqrt{2}}{2}}-1}{\sqrt{\frac{6+2\cdot\sqrt{2}}{2}}+1} = \frac{\sqrt{3+\sqrt{2}}-1}{\sqrt{3+\sqrt{2}}+1} \approx 0.36$$

We use THEOREM 5.2 from lecture notes:

$$\gamma = \frac{4}{(\sqrt{\alpha}+\sqrt{\beta})^2}, \quad \mu = \left(\frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\beta}+\sqrt{\alpha}}\right)^2 \quad \text{satisfies:} \quad \boxed{\|x_{k+1} - x^*\| \le \left(\frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\beta}+\sqrt{\alpha}} + \varepsilon_k\right)^k \cdot \|x_1 - x^*\|}$$

$$\gamma = \frac{4}{(\sqrt{2}+\sqrt{6+2\cdot\sqrt{2}})^2} \approx 0.21 \qquad \mu = \left(\frac{\sqrt{6+2\cdot\sqrt{2}}-\sqrt{2}}{\sqrt{6+2\cdot\sqrt{2}}+\sqrt{2}}\right)^2 = 0.13$$