Minimize the function

$$f(\boldsymbol{x}) = f(x, y) = (x-1)^2 + (y+1)^2 + xy \tag{1}$$

using naive n-dimensional optimization by hand for four steps, cycling through the direction vectors

$$\boldsymbol{v_1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{v_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

starting with the guess

$$\boldsymbol{x_0} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Eq. (1) is simple enough that you should not need to use an optimization program to carry out each of the 1-dimensional optimizations; instead, you should be able to do these by hand using single-variable calculus.

At each step, calculate the 'true relative error' in $\boldsymbol{x}$ by comparing against the correct value of the minimum, and the 'approximate relative error' by comparing against the previously-computed value (i.e., compare step 1 against step 0, step 2 against step 1, etc.). These errors can be calculated using the following definitions from lecture 2:

$$\text{True relative error} = \frac{||\boldsymbol{x} - \boldsymbol{x}_{\text{true}}||}{||\boldsymbol{x}_{\text{true}}||}, \text{ and} \tag{2}$$

$$\text{Approximate relative error} = \frac{||\boldsymbol{x}_{\text{current}} - \boldsymbol{x}_{\text{previous}}||}{||\boldsymbol{x}_{\text{previous}}||}, \tag{3}$$

where the norm of a 2-dimensional vector can be calculated using

$$||\boldsymbol{x}|| = ||[x, y]|| = \sqrt{x^2 + y^2}. \tag{4}$$

Show all your work, and create a table of the following kind:

	Step 0	Step 1	Step 2	Step 3	Step 4
x	-1				
y	2				
$f(x, y)$	11				
True relative error	$5/(2\sqrt{2})$				
Approx. relative error	-				

Ans. In the first step, we minimize $f(\underline{x})$ along the vector $\underline{v}_1 = [1, 1]$ starting from the point $[-1, 2]$, i.e., minimize the function

$$f(\underline{x}_0 + \alpha\underline{v}_1) = f([-1, 2] + \alpha \cdot [1, 1]) = 3\alpha^2 + 3\alpha + 11$$

over α, which comes out to

$$6\alpha + 3 = 0 \implies \alpha = -1/2.$$

Hence, the new value of $\underline{x}$ is

$$\underline{x}_1 = \underline{x}_0 + \alpha\underline{v}_1 = [-1, 2] - \frac{1}{2}[1, 1] = [-3/2, 3/2].$$

In the second step, we minimize $f(\underline{x})$ along the vector $\underline{v}_2 = [0, 1]$ starting from the point $[-3/2, 3/2]$, i.e., minimize the function

$$f(\underline{x}_1 + \alpha\underline{v}_2) = f([-3/2, 3/2] + \alpha \cdot [0, 1]) = \alpha^2 + 3.5\alpha + 10.25$$

over α, which comes out to

$$2\alpha + 3.5 = 0 \implies \alpha = -1.75.$$

Hence, the new value of $\underline{x}$ is

$$\underline{x}_2 = \underline{x}_1 + \alpha\underline{v}_2 = [-3/, 3/2] - 1.75 \cdot [0, 1] = [-1.5, -0.25].$$

In the third step, we minimize $f(\underline{x})$ along the vector $\underline{v}_1 = [1, 1]$ starting from the point $[-1.5, -0.25]$, i.e., minimize the function

$$\begin{aligned} f(\underline{x}_2 + \alpha\underline{v}_1) = f([-1.5, -0.25] + \alpha \cdot [1, 1]) &= f([\alpha - 1.5, \alpha - 0.25]) \\ &= (\alpha - 2.5)^2 + (\alpha + 0.75)^2 + (\alpha - 1.5)(\alpha - 0.25) \end{aligned}$$

over α, which comes out to

$$\alpha = 0.875.$$

Hence, the new value of $\underline{x}$ is

$$\underline{x}_3 = \underline{x}_2 + \alpha\underline{v}_1 = [-1.5, -0.25] + 0.875 \cdot [1, 1] = [-0.625, 0.625].$$

In the fourth step, we minimize $f(\underline{x})$ along the vector $\underline{v}_2 = [0, 1]$ starting from the point $[-0.625, 0.625]$, i.e., minimize the function

$$\begin{aligned} f(\underline{x}_3 + \alpha\underline{v}_2) = f([-0.625, 0.625] + \alpha \cdot [0, 1]) &= f([-0.625, \alpha + 0.625]) \\ &= (-1.625)^2 + (\alpha + 1.625)^2 + (-0.625)(\alpha + 0.625) \end{aligned}$$

over α, which comes out to

$$\alpha = -1.3125.$$

Hence, the new value of $\underline{x}$ is

$$\underline{x}_4 = \underline{x}_3 + \alpha\underline{v}_2 = [-0.625, 0.625] - 1.3125 \cdot [0, 1] = [-0.625, -0.6875].$$

We can summarize these results in the following table,

	Step 0	Step 1	Step 2	Step 3	Step 4
x	-1	-1.5	-1.5	-0.625	-0.625
y	2	$+1.5$	-0.25	$+0.625$	-0.6875
$f(x,y)$	11	10.25	7.1875	4.8906	3.1679
True relative error	1.7677	1.7500	1.3835	1.3125	1.0376
Approx. relative error	-	0.3162	0.8249	0.8137	1.4849

- (15 points) Illustrate the steps from problem 1 graphically on the following plot, which shows contours of $f(\boldsymbol{x})$. Label the points $\boldsymbol{x}_0$ through $\boldsymbol{x}_4$, as well as the direction vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ drawn at the appropriate locations. Altogether, you will have five points connected by four lines: two lines will have the direction of $\boldsymbol{v}_1$ and two will have the direction of $\boldsymbol{v}_2$. You may either use repeated labels or label the successive directions $[\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4]$.

Ans. The completed figure is below.

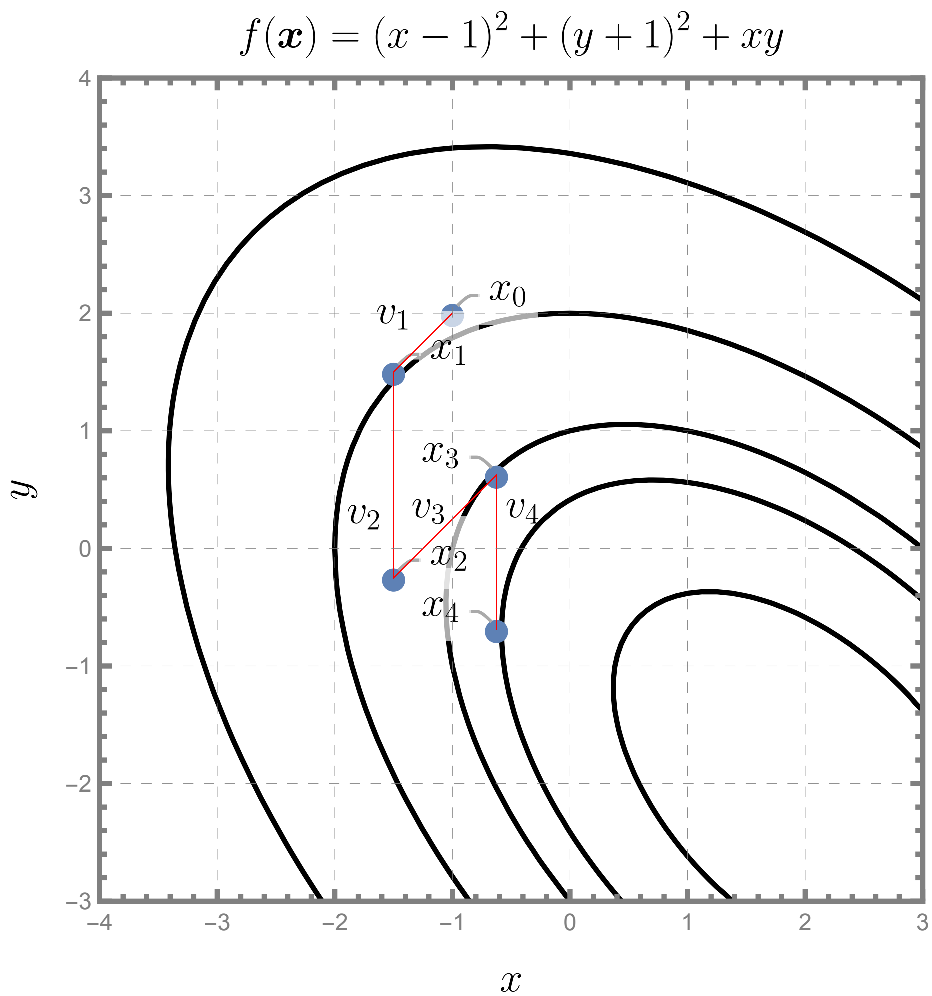

Figure 1: Contours of $f(\boldsymbol{x})$ from eq. (1).