1.a The Rosenbrock function is $f(\tilde{x}) = 100(\mathrm{x}_2 - {x_1}^2)^2 + (1 - x_1)^2$

We will compute **gradient** and **hessian** of this function and also we will show $\tilde{x}_* = (1,1)$ is a local minimizer of this function.

$f(\tilde{x}) = 100(\mathrm{x}_2 - {x_1}^2)^2 + (1 - x_1)^2$

For gradient computation,

$\nabla f(\tilde{x}) = <$ -100.2.2.$\mathrm{x}_1(x_2 - {x_1}^2) - 2(1 - x_1)$, $2.100(x_2 - {x_1}^2) >$

$\nabla f(\tilde{x}) = < 400\mathrm{x}_1^3 - 400x_1x_2 + 2x_1 - 2$, $200x_2 - 200{x_1}^2 >$

$\nabla f(1,1) = < 400.1^3 - 400.1.1 + 2.1 - 2$, $200.1 - 200.1^2 >$

$\nabla f(1,1) = < 0 , 0 >$ ==> Gradient is zero.

For hessian computation,

We will compute two times gradient of this function to find hessian.

$\dfrac{\partial f}{\partial x_1} = 400.1^3 - 400.1.1 + 2.1 - 2$, $\dfrac{\partial f}{\partial x_2} = 200\mathrm{x}_2 - 200{x_1}^2$ ==> First gradient.

$\dfrac{\partial^2 f}{\partial {x_1}^2} = 1200\mathrm{x}_1^2 - 400x_2 + 2$, $\dfrac{\partial^2 f}{\partial x_1 \partial x_2} = \text{-}400\mathrm{x}_1$

$\dfrac{\partial^2 f}{\partial x_1 \partial x_2} = \text{-}400\mathrm{x}_1$, $\dfrac{\partial^2 f}{\partial {x_2}^2} = 200$ ==> Second gradient.

$$\mathrm{H}_f(\tilde{x}) = \nabla^2 f(\tilde{x}) = \begin{bmatrix} 1200{x_1}^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

$$\mathrm{H}_f(1,1) = \nabla^2 f(1,1) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$$

If determinant of Hessian matrix is positive, then we show that Hessian is positive definite.

$$\det(\nabla^2 f(1,1)) = \det \begin{pmatrix} 802 & -400 \\ -400 & 200 \end{pmatrix} = 802.200 - 400.400 = 400 \quad ==> \text{ Hessian is positive definite.}$$

1.b We have three variables that are vector $\tilde{a} \in \mathbb{R}^n$, $\tilde{x} \in \mathbb{R}^n$ and symmetric matrix $\mathrm{A} \in \mathbb{R}^{n \times n}$. For symmetric matrix, its transpose is equals to itself. $(\mathrm{a}_{1,0} = a_{0,1}, a_{2,0} = a_{0,2}..)$

We can symbolize these variables in terms of their types like this:

$$\tilde{a} = \begin{bmatrix} \vec{v}_0 \\ \vec{v}_1 \\ \vec{v}_2 \\ \vdots \\ \vec{v}_{n-1} \end{bmatrix}, \quad \tilde{x} = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{bmatrix}, \quad A = \begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,n-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,n-1} \\ a_{2,0} & a_{2,1} & \cdots & a_{2,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,0} & a_{n-1,1} & \cdots & a_{n-1,n-1} \end{bmatrix}$$

Now, we will compute gradient and Hessian matrices of $f_1(\tilde{x}) = \tilde{a}^T \tilde{x}$ and $f_2(\tilde{x}) = \tilde{x}^T A \tilde{x}$

For $f_1(\tilde{x}) = \tilde{a}^T \tilde{x}$ **,**

$$f_1(\tilde{x}) = \tilde{a}^T \tilde{x} = \begin{bmatrix} \vec{v}_0 & \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_{n-1} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{bmatrix}$$

$$\mathrm{f}_1(\tilde{x}) = \vec{v}_0 x_0 + \vec{v}_1 x_1 + \vec{v}_2 x_2 + \ldots + \vec{v}_{n-1} x_{n-1}$$

$$(\textbf{Gradient} \text{ of } \mathrm{f}_1) \quad \nabla f_1(\tilde{x}) = f_1(\tilde{x}) = \tilde{a}^T \tilde{x} = \begin{bmatrix} \vec{v}_0 \\ \vec{v}_1 \\ \vec{v}_2 \\ \vdots \\ \vec{v}_{n-1} \end{bmatrix}$$

We should take 2 times gradient of f_1 in order to find Hessian matrix.

If we take gradient of $\nabla f_1(\tilde{x})$ in terms of $\tilde{x}$, we can see easily that all elements are **zero**.

$$(\textbf{Hessian} \text{ of } \mathrm{f}_1) \quad \nabla^2 f_1(\tilde{x}) = f_1(\tilde{x}) = \tilde{a}^T \tilde{x} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}_{n \times n}$$

For $f_2(\tilde{x}) = \tilde{x}^T A \tilde{x}$**,**

$$f_1(\tilde{x}) = \tilde{x}^T A \tilde{x} = \begin{bmatrix} x_0 & x_1 & x_2 & \dots & x_{n-1} \end{bmatrix} \begin{bmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,n-1} \\ a_{1,0} & a_{1,1} & \dots & a_{1,n-1} \\ a_{2,0} & a_{2,1} & \dots & a_{2,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,0} & a_{n-1,1} & \dots & a_{n-1,n-1} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{bmatrix}$$

$$(\textbf{Gradient of } f_2) \quad \nabla f_2(\tilde{x}) = f_2(\tilde{x}) = \tilde{x}^T A \tilde{x} = \begin{bmatrix} 2x_0a_{0,0} + x_1a_{0,1} + x_1a_{1,0} + \ldots + x_{n-1}a_{0,n-1} + x_{n-1}a_{n-1,0} \\ x_0a_{1,0} + x_0a_{0,1} + 2x_1a_{1,1} + \ldots + x_{n-1}a_{1,n-1} + x_{n-1}a_{n-1,1} \\ x_0a_{2,0} + x_0a_{0,2} + \ldots + x_2a_{2,2} + \ldots + x_{n-1}a_{n-1,2} + x_{n-1}a_{2,n-1} \\ \vdots \\ x_0a_{n-1,0} + x_0a_{0,n-1} + x_1a_{n-1,1} + x_1a_{1,n-1} + \ldots + 2x_{n-1}a_{n-1,n-1} \end{bmatrix}$$

Since A is symmetric matrix , A is equal to A^T. Therefore, we can use that $a_{0,1} = a_{1,0}$, $a_{0,2} = a_{2,0}$, $a_{0,3} = a_{3,0}$, . . . , $a_{0,n-1} = a_{n-1,0}$ and we can simplify our matrix using this property.

$$(\textbf{Gradient of } f_2) \quad \nabla f_2(\tilde{x}) = f_2(\tilde{x}) = \tilde{x}^T A \tilde{x} = \begin{bmatrix} 2x_0a_{0,0} + 2x_1a_{0,1} + \ldots + 2x_{n-1}a_{0,n-1} \\ 2x_0a_{1,0} + 2x_1a_{1,1} + \ldots + 2x_{n-1}a_{1,n-1} \\ 2x_0a_{2,0} + \ldots + 2x_2a_{2,2} + \ldots + 2x_{n-1}a_{2,n-1} \\ \vdots \\ 2x_0a_{n-1,0} + 2x_1a_{n-1,1} + \ldots + 2x_{n-1}a_{n-1,n-1} \end{bmatrix}$$

$$(\textbf{Hessian of } f_2) \quad \nabla^2 f_2(\tilde{x}) = f_2(\tilde{x}) = \tilde{x}^T A \tilde{x} = \begin{bmatrix} 2a_{0,0} & 2a_{0,1} & \ldots & 2a_{0,n-1} \\ 2a_{1,0} & 2a_{1,1} & \ldots & 2a_{1,n-1} \\ 2a_{2,0} & 2a_{2,1} & \ldots & 2a_{2,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 2a_{n-1,0} & 2a_{n-1,1} & \ldots & 2a_{n-1,n-1} \end{bmatrix}$$

1.c Our function is $f(x_1, x_2) = (x_1 + x_2{}^2)^2$ and we will substitute $x_1 = 1 + \alpha(-1)$, $x_2 = 0 + \alpha(1)$ in order to find critical points and lastly, we will prove that these are minimizers.

Firstly, we will substitute $x_1 = 1 + \alpha(-1)$, $x_2 = 0 + \alpha(1)$.

$f(1 - \alpha \ , \ 0 + \alpha) = (1 - \alpha + \alpha^2)^2$

Now, we will take derivative of f to find critical points.

$$\begin{aligned} f'(1 - \alpha \ , \ 0 + \alpha) &= 2(2\alpha - 1)(\alpha^2 - \alpha + 1) \\ &= 2(2\alpha^3 - 2\alpha^2 + 2\alpha - \alpha^2 + \alpha - 1) \\ &= 2(2\alpha^3 - 3\alpha^2 + 3\alpha - 1) \\ &= 2(\alpha^3 + \alpha^3 - 3\alpha^2 + 3\alpha - 1) \quad ==> (\alpha^3 - 1) = (\alpha - 1)(\alpha^2 + \alpha + 1) \\ &= 2(2\alpha - 1)(\alpha^2 - \alpha + 1) ==> \textbf{That is final result} \end{aligned}$$

Critical points are $=> \alpha = \frac{1}{2}$, $\alpha = \frac{1 + \sqrt{3}i}{2}$, $\alpha = \frac{1 - \sqrt{3}i}{2}$

After finding critical points, we will compute gradient of f for critical points. If the result is zero, then these are minimizers.

$f'(1-\alpha\ ,\ 0+\alpha) = 2(2\alpha-1)(\alpha^2-\alpha+1)$

For $\alpha = \frac{1}{2}$,

$f'(1/2\ ,\ 1/2) = 2(2.0,5--1)((0.5)^2-0.5+1) = 0$

Our result is 0, therefore there is one minimizer when α **is** $\frac{1}{2}$**. Rest of these do not provide our result.**