Assignment 1

1. a. $f(x) = (4x_1^2 - x_2)^2$

$$\nabla f(x) = \begin{bmatrix} 16x_1(4x_1^2 - x_2) \\ -2(4x_1^2 - x_2) \end{bmatrix} \quad Hf(x) = \begin{bmatrix} 128x_1^2 + 16(4x_1^2 - x_2) & \\ -16x_1 & 2 \end{bmatrix}$$

$\nabla f(x) = 0 = 16x_1(4x_1^2 - x_2)$

$\boxed{x_2 = 4x_1^2}$ → all stationary points follow this relation.

When sub. x_2 in the Hessian:

$$Hf(x) = \begin{bmatrix} 2(8x_1)^2 & \\ -2x_1 & 2 \end{bmatrix}$$ Plugging in $x_1 = 0$

$$Hf(0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

All points are global minimizers since $f(x)$ doesn't not go below 0.

b. $g(x) = x_1^2 + 4x_1x_2 + x_2^2 + x_1 - x_2$

$$\nabla g(x) = \begin{bmatrix} 2x_1 + 4x_2 + 1 \\ 4x_1 + 2x_2 - 1 \end{bmatrix} \quad Hf(x) = \begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix}$$

$\nabla g(x) = 0 = 2x_1 + 4x_2 + 1 = 0$

$x_1 = -2x_2 - 1/2$

$4(-2x_2 - 1/2) + 2x_2 - 1 = 0$

$x_2 = -1/2$

$x_1 = 1/2$

Solve for eigen values of Hessian:

$$\det\begin{bmatrix} 2-\lambda & 4 \\ 4 & 2-\lambda \end{bmatrix} = 0$$

$(2-\lambda)^2 - 16 = 0$

$\lambda^2 - 4\lambda - 12 = 0$

$(\lambda - 6)(\lambda + 2) = 0$

$\lambda = 6, -2$ → So the Hessian is indefinite.

Since the Hessian is indefinite, the point $(1/2, -1/2)$ is a saddle point.

C. - f(x) is not coercive because along the path where the function f(x) has $x_2 = 4x_1^2$ the function will not increase to ∞.

- g(x) is not coercive because there is only 1 saddle point at $(1/2, -1/2)$, and no other stationary points that there are directions which go to $-\infty$. An example of this is the direction $\hat{i} - \hat{j}$.

D. MatLab code and graphs attached.

Note the points plotted in the f(x) graph represent the parabola $x_2 = 4x_1^2$.

2. $A \in \mathbb{R}^{n \times n}$ and is Symmetric (So $A = A^T$).

$$R_A(x) = \frac{x^T A x}{\|x\|^2}, \quad \forall x \neq 0.$$

$$\frac{x^T A x}{\|x\|^2} = \frac{x^T A x}{x^T x} = \frac{(Uy)^T A (Uy)}{(Uy)^T (Uy)} = \frac{y^T U^T A U y}{U^T y^T U y} = \frac{y^T \Lambda y}{y^T y}$$

$$= \frac{\sum_{i=1}^{n} \lambda_i y_i^2}{\|y\|^2}.$$

Considering $n=2$, $\frac{\sum_{i=1}^{2} \lambda_i y_i^2}{\|y\|^2} = \frac{\lambda_1 y_1^2 + \lambda_2 y_2^2}{\|y\|^2}$

From This, The $\|y\|^2$ gives a weight to each λ between 0 - 1. So assuming λ_1 is The minimum and λ_2 is The maximum, it can be seen That The smallest the $R_A(x)$ can be is λ_1 and The Largest as λ_2.

. So

$$\lambda_{min} \leq R_A(x) \leq \lambda_{max} \quad \forall x \neq 0.$$

3. a. $A \in \mathbb{R}^{n \times n}$ is a Symmetric matrix and The Diagonals are non-negative elements.

$$|A_{ii}| \geq \sum_{j \neq i} |A_{ij}| \quad \forall i \in \{1 \ldots n\} \rightarrow A \succeq 0$$

Assume $\exists \lambda < 0$ Such That $Ax = \lambda x$. Then

$$|A_{ii} - \lambda||x_i| = \left|\sum_{j \neq i} A_{ij} x_j\right|$$

$$\left|\sum_{j \neq i} A_{ij} x_j\right| \leq \sum_{j \neq i} |A_{ij}||x_j| \leq |A_{ii}||x_j|$$

So

$$|A_{ii} - \lambda| |x_i| \leq |A_{ii}||x_j|$$

Assuming x_i is the biggest value in $\vec{x}$. Then we have

$$|A_{ii} - \lambda| |x_i| \leq |A_{ii}||x_j| \leq |A_{ii}||x_i|$$

So if λ is negative This would violate The inequality above. So since $\lambda \geq 0 \rightarrow A \succeq 0$.

3. b. Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix with positive diagonal elements.

$$|A_{ii}| > \sum |A_{ij}| \quad \forall i \in \{1, \dots n\} \rightarrow A \succ 0$$

Assume $\exists \lambda < 0$ S.T. $Ax = \lambda x$ Then

$$|A_{ii} - \lambda||x_i| = \left|\sum_{j \neq i} A_{ij} x_j\right| \leq \sum_{j \neq i} |A_{ij}||x_j| < |A_{ii}||x_j|$$

So $|A_{ii} - \lambda||x_i| < |A_{ii}||x_j|$

Assume x_i is the biggest value in $\vec{x}$. Then we have

$$|A_{ii} - \lambda||x_i| < |A_{ii}||x_j| < |A_{ii}||x_i|$$

If $\lambda \leq 0$ The inequality would be violated.

So $\lambda > 0 \rightarrow A \succ 0$.

3. c. i $\begin{bmatrix} -3 & 1 & 1 \\ 1 & -4 & 1 \\ 1 & 1 & -5 \end{bmatrix}$ $Ax = \lambda x$

$$\det \begin{bmatrix} -3-\lambda & 1 & 1 \\ 1 & -4-\lambda & 1 \\ 1 & 1 & -5-\lambda \end{bmatrix} = 0$$

$$(-3-\lambda) \begin{vmatrix} -4-\lambda & 1 \\ 1 & -5-\lambda \end{vmatrix} - \begin{vmatrix} 1 & 1 \\ 1 & -5-\lambda \end{vmatrix} + \begin{vmatrix} 1 & -4-\lambda \\ 1 & 1 \end{vmatrix} = 0$$

$$(-3-\lambda)\left[(-4-\lambda)(-5-\lambda)-1\right] - \left[-5-\lambda-1\right] + \left[1-(-4-\lambda)\right] = 0$$

$$(-3-\lambda)(\lambda^2+9\lambda+19) + 5 + \lambda + 1 + 1 + 4 + \lambda = 0$$

$$-\lambda^3 - 9\lambda^2 - 19\lambda - 3\lambda^2 - 27\lambda - 57 + 5 + \lambda + 6 + \lambda = 0$$

$$-\lambda^3 - 12\lambda^2 - 44\lambda - 46 = 0$$

$$\lambda^3 + 12\lambda^2 + 44\lambda + 46 = 0$$

$\lambda = -1.786, -4.539, -5.675$ (Found by plotting on Desmos).

Since all λ is negative this matrix is <u>negative Definite</u>.

3 c. ii $\begin{bmatrix} 2 & -4 & 0 \\ -4 & 8 & 0 \\ 0 & 0 & -3 \end{bmatrix}$ $Ax = \lambda x$

$$\det \begin{bmatrix} 2-\lambda & -4 & 0 \\ -4 & 8-\lambda & 0 \\ 0 & 0 & -3-\lambda \end{bmatrix} = 0$$

$$(-3-\lambda) \begin{vmatrix} 2-\lambda & -4 \\ -4 & 8-\lambda \end{vmatrix} = 0$$

$$(-3-\lambda)\left[(2-\lambda)(8-\lambda) - 16\right] = 0$$

$$(-3-\lambda)\left[\lambda^2 - 10\lambda\right] = 0$$

$$-\lambda^3 + 10\lambda^2 - 3\lambda^2 + 30\lambda = 0$$

$$\lambda^3 - 7\lambda^2 - 30\lambda = 0$$

$$\lambda(\lambda^2 - 7\lambda - 30) = 0$$

$\lambda = 0$ or $\lambda^2 - 7\lambda - 30 = 0$

$$(\lambda - 10)(\lambda + 3) = 0$$

$$\lambda = 10, -3$$

Since there is at least 1 positive and negative eigen value, the matrix is indefinite.

3. c. iii

$$\begin{bmatrix} 2 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix} \rightarrow \det \begin{bmatrix} 2-\lambda & 2 & 0 & 0 \\ 2 & 2-\lambda & 0 & 0 \\ 0 & 0 & 3-\lambda & 1 \\ 0 & 0 & 1 & 3-\lambda \end{bmatrix} = 0$$

$$(2-\lambda)\begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & 3-\lambda & 1 \\ 0 & 1 & 3-\lambda \end{vmatrix} - 2\begin{vmatrix} 2 & 0 & 0 \\ 0 & 3-\lambda & 1 \\ 0 & 1 & 3-\lambda \end{vmatrix} = 0$$

$$(2-\lambda)\left((2-\lambda)\begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix}\right) - 2\left(2\begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix}\right) = 0$$

$$(2-\lambda)(2-\lambda)\left[(3-\lambda)^2 - 1\right] - 4\left[(3-\lambda)^2 - 1\right] = 0$$

$$(\lambda^2 - 4\lambda + 4)(\lambda^2 - 6\lambda + 8) - 4(\lambda^2 - 6\lambda + 8) = 0$$

$$\lambda^4 - 6\lambda^3 + 8\lambda^2 - 4\lambda^3 + 24\lambda^2 - 32\lambda + \cancel{4\lambda^2} \cancel{- 24\lambda} \cancel{+ 32} \ldots$$

$$\cancel{-4\lambda^2} \cancel{+ 24\lambda} \cancel{- 32} = 0$$

$$\lambda^4 - 10\lambda^3 + 32\lambda^2 - 32\lambda = 0$$

$$\lambda(\lambda^3 - 10\lambda^2 + 32\lambda - 32) = 0$$

$$\lambda(\lambda + 2)(\lambda + 4)(\lambda + 4) = 0$$

$$\lambda = 0, 2, 4$$

↳ Double root.

Since all eigen values are ≥ 0 This matrix is <u>positive semidefinite.</u>

4. $f(x) = \frac{1}{2}x^TQx - c^Tx + d$; $Q \in R^{n\times n}$ is symmetric
$c \in R^n$; $d \in R$

if $Q \succ 0 \to f$ is coercive

Starting with $c^Tx \leq \|c\|\|x\|$
So $-c^Tx \geq -\|c\|\|x\|$
So as $\|x\| \to \infty$ the right side would go to $-\infty$
which mean that $\lim_{\|x\|\to\infty} -c^Tx = -\infty$.

However we know that $\lambda_{min} \leq R_a(x) \leq \lambda_{max}$
also since $Q \succ 0$ we have that
$x^TQx > 0$.
Now multiplying both sides by $\frac{\|x\|^2}{\|x\|^2}$ we have

$$\frac{1}{2}\frac{x^TQx}{\|x\|^2}\|x\|^2 = \frac{1}{2}R_a(x)\|x\|^2 \geq \frac{1}{2}\lambda_{min}\|x\|^2.$$

Since we know $\lambda_{min} > 0$ then as $\|x\|^2 \to \infty$
the right side goes to ∞. and since $\frac{1}{2}R_a(x)\|x\|^2$
is lower bounded we know it will also go to ∞.

So as $\|x\|^2 \to \infty$ $\frac{1}{2}x^TQx \to \infty$.

Since $\frac{1}{2}x^TQx$ goes to ∞ faster than $-c^Tx$
goes to $-\infty$, $\lim_{\|x\|\to} f(x) = \infty \to f$ is coercive.

5. $f(x) = 4x_1^2 + 2x_1x_2 + 2x_2^2$

a. The Function, $f(x)$, is coercive because as

$\lim_{\|x\| \to \infty} f(x) = \infty$.

b. $\nabla f(x) = \begin{bmatrix} 8x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{bmatrix}$ $\quad Hf(x) = \begin{bmatrix} 8 & 2 \\ 2 & 4 \end{bmatrix}$

$8x_1 + 2x_2 = 0 \qquad 2x_1 + 4x_2 = 0$

$x_2 = -4x_1 \qquad x_2 = -\frac{1}{2}x_1$

Only way for This To be true is if $x_1 = x_2 = 0$. Since $Hf(x) > 0$ for all x. The point $(0,0)$ is a global minimizer

c. Plot / code attached To File.

d. Plot attached to File.

e. The Largest Step size Found was $\gamma \doteq 0.2265$. This was Found Through guess and Check. The Convergence is better For $\gamma = 0.025$, using The Max Step Size was very Slow Versus a Step Size OF 0.025.