098311 Optimization 1 Spring 2018
HW 2

Chen Tessler 305052680
Orr Krupnik 302629027

April 17, 2018

Problem 1. Find the global minimum and maximum points of the linear function $f(x, y) = 7x + 12y$ over the set $S = \{(x, y) : 2x^2 + 6xy + 9y^2 - 2x - 6y \leq 24\}$.

Solution The set S is an ellipse, and is therefore a bounded set. The set S also contains its boundary, which makes it a closed set. Since S is both closed and bounded, it is a compact set, and by Weierstrass Theorem, we know global maxima and minima exist over the set.
The linear function $f(x, y) = 7x + 12y$ has no stationary points (since $\nabla f = (7, 12)^T \neq 0 \ \forall x, y$). A local optimum of the function would require $\nabla f(x, y) = 0$. Therefore, there does not exist a local optimum in $int(S)$. Since we know optimal points do exist over the set (see above), we can limit our search to $bd(S) = \{(x, y) : 2x^2 + 6xy + 9y^2 - 2x - 6y - 24 = 0\}$.
Note that for some constant C, the line $7x+12y = C$ defines a constant-valued line of $f(x, y)$. An infinite number of such lines exists; however, we are interested in the ones intersecting the set S, and specifically, the lines with maximal and minimal values of C.
The lines described above can either intersect the ellipse S or be tangent to it. The intersecting lines will not provide us with the required solution, since they have points in $int(S)$. Therefore, we can look for the tangent lines with the maximal and minimal values of C. We plug in the line equation to the ellipse equation in order to find the tangent lines:

$$\begin{aligned} 2x^2 + 6x\Big(-\frac{7}{12}x + \frac{C}{12}\Big) + 9\Big(-\frac{7}{12}x + \frac{C}{12}\Big)^2 - 2x - 6\Big(-\frac{7}{12}x + \frac{C}{12}\Big) - 24 = 0 \\ 2x^2 - \frac{7}{2}x + \frac{C}{2} + \frac{49}{16}x^2 - \frac{7C}{8}x + \frac{C}{16} - 2x + \frac{7}{2}x - \frac{C}{2} - 24 = 0 \\ \frac{81}{16}x^2 - \Big(\frac{7C}{8} + 2\Big)x + \frac{C}{16} - 24 = 0 \end{aligned}$$

To find lines which are tangents, we are looking for values of C for which the above quadratic equation has only one real root, which means only one point of intersection with the ellipse. The roots of this equation are:

$$x_{1,2} = \frac{8}{81}\Big(\frac{7C}{8} + 2\Big) \pm \frac{8}{81}\sqrt{\Big(\frac{7C}{8} + 2\Big)^2 - \frac{81C}{64} - 486}$$

For a single root, we require:

$$\frac{49}{64}C^2 + \frac{7}{2}C + 4 - \frac{81}{64}C - 486 = 049C^2 + 143C - 30848 = 0$$
$$C_{1,2} = -\frac{143}{98} \pm \sqrt{143^2 + 4 \cdot 49 \cdot 30848}$$
$$C_1 = 23.674, C_2 = -26.592$$

The resulting values of C are the values of the global minimum and maximum of f over the set S.

Problem 2. Let $f(x) = x^T Ax + 2b^T x + c$, where $A \in \mathbb{R}^{n\times n}$ is symmetric, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Suppose that $A \succeq 0$. Show that f is bounded below over $\mathbb{R}^n$ if and only if $b \in \text{Range}(A)$.

Solution

- f is bounded from below $\Rightarrow b \in \text{Range}(A)$:
  Let us assume $b \notin \text{Range}(A)$. Therefore, we can write b as $b = \sum_{i=1}^n \alpha_i A_i + q \triangleq Ay + q$, where $\{\alpha\}_{i=1}^n$ are some coefficients, $\{A\}_{i=1}^n$ are the columns of A and $q \in \mathbb{R}^n$ is some vector such that $q \in \text{Null}\{A\}$ (assuming $\text{Null}\{A\} \neq \{\mathbf{0}\}$; see below for the opposite case).

  $$f(\beta q) = \beta^2 q^T Aq + 2\beta y^T Aq + 2\beta q^T q + c$$

  note that $Aq = \mathbf{0}$ since $q \in \text{Null}(A)$. Therefore:

  $$f(\beta q) = 2\beta q^T q + c = 2\beta \sum_{i=1}^n q_i^2 + c \xrightarrow{\beta\to-\infty} -\infty$$

  This gives us a path along which $f(x) \to -\infty$, in contradiction to the fact that f is bounded from below. It remains to address the case in which $\text{Null}\{A\} = \{\mathbf{0}\}$. In this case, $x^T Ax > 0$ for any $x \neq \mathbf{0} \in \mathbb{R}^n$, and therefore A is PD. This, in turn, means that A has n non-zero eigenvalues, and is of full rank. Therefore, $\text{Range}(A) = \mathbb{R}^n$, since it is spanned by n linearly independent vectors. This means $b \in range(A)$, as required.

- $b \in \texttt{Range}(A) \Rightarrow f$ is bounded from below:
  Notice that since $b \in \text{Range}(A)$, b is a linear combination of the vectors of $A \Rightarrow b = Ay$.

  $$f(x) = x^T Ax + 2b^T x + c = x^T Ax + 2y^T Ax + c$$
  $$\nabla_x f(x) = 2Ax + 2Ay = 2A(x + y)$$
  $$\nabla_x^2 f(x) = 2A \succeq 0$$

  f is twice continuously differentiable (as it is a linear combination of twice continuously differentiable functions) defined over $\mathbb{R}^n$. We know that $\nabla_x^2 f(x) \succeq 0$ for any $x \in \mathbb{R}^n$. Using the theorem from lecture 2, we can now conclude that any stationary point of f is also a global minimum of f. Additionally, observing $\nabla_x f(x)$ we see that $x \equiv -y$ is a stationary point of $f \Rightarrow f(-y) = y^T Ay + 2y^T Ay + c > -\infty$ is a global minimum $\Rightarrow f(x)$ is bounded from below over $\mathbb{R}^n$.

Problem 3. 1. For each of the following matrices determine, without computing the eigenvalues, whether they are positive or negative definite or semidefinite or indefinite.

a)

$$A = \begin{pmatrix} 2 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

b)

$$B = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 3 & 3 \\ 2 & 3 & 3 \end{pmatrix}$$

c)

$$C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

d)

$$D = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 0 \end{pmatrix}$$

2. For each of the following matrices determine, without computing the eigenvalues, the interval of α for which they are positive or negative definite or semidefinite or indefinite.

a)

$$D_\alpha = \begin{pmatrix} -1 & \alpha & -1 \\ \alpha & -4 & \alpha \\ -1 & \alpha & -1 \end{pmatrix}$$

b)

$$E_\alpha = \begin{pmatrix} 1 & \alpha & 0 & 0 \\ \alpha & 2 & \alpha & 0 \\ 0 & \alpha & 2 & \alpha \\ 0 & 0 & \alpha & 1 \end{pmatrix}$$

Solution

1. a) By definition:

$$\begin{aligned} xAx^T &= 2x_1(x_1+x_2) + 2x_2(x_1+x_2) + x_3(3x_3+x_4) + x_4(x_3+3x_4) \\ &= 2(x_1+x_2)^2 + (x_3+x_4)^2 + 2(x_3^2+x_4^2) \geq 0 \end{aligned}$$

for an example vector $x = (x_1, -x_1, 0, 0)$ notice that the above $xAx^T = 0$ hence matrix A is positive semi-definite.

b) Since B has two identical lines, we have $|B| = 0$ and therefore at least one zero eigenvalue. Since $tr(B) > 0$, there is also at least one positive eigenvalue. Therefore, B can either be PSD or indefinite. If we find at least one vector x such that $x^T Bx < 0$ we can conclude B is indefinite. We have:

$$\begin{aligned} x^T Bx &= x_1(2x_1+2x_2+2x_3) + x_2(2x_1+3x_2+3x_3) + x_3(2x_1+3x_2+3x_3) \\ &= 2x_1^2 + 3x_2^2 + 3x_3^2 + 4x_1x_2 + 6x_2x_3 + 5x_1x_3 \end{aligned}$$

For instance, for $x = \left(\frac{1}{2}, \frac{1}{2}, -1\right)^T$ we get $x^T Bx = -\frac{1}{2} < 0$, and therefore B is indefinite.

c) C is a symmetric matrix with $|C_{i,i}| \geq \sum_{j \neq i} |C_{i,j}| \forall i = 1, 2, 3$, and only non-negative diagonal elements. Therefore, C is diagonally dominant, and PSD.

d) Since D has two linearly-dependant lines, we have $|D| = 0$ and therefore at least one zero eigenvalue. Since $tr(D) > 0$, there is also at least one positive eigenvalue. Therefore, D can either be PSD or indefinite. If we find at least one vector x such that $x^T Dx < 0$ we can conclude D is indefinite. We have:

$$x^T Dx = x_1^2 + 4x_1x_2 + 6x_1x_3 + 4x_2^2 + 12x_2x_3$$

For the $x = (1, 1, -1)^T$ we have $x^T Dx = 1+4-6+4-12 = -9 < 0$ and therefore D is indefinite.

2. a) Let us look at $-D_\alpha$:

$$-D_\alpha = \begin{pmatrix} 1 & -\alpha & 1 \\ -\alpha & 4 & -\alpha \\ 1 & -\alpha & 1 \end{pmatrix}$$

Its principal minors are 1, $4 - 2\alpha$ and 0. Therefore, for any $|\alpha| < 2$, we have all principal minors positive and the last one zero, and therefore $-D_\alpha$ is PSD, and D_α is NSD.

For $\alpha \leq -2$ and $\alpha \geq 2$, let us look at $x^T D_\alpha x$ for some $x = (x_1, x_2, x_3)^T$:

$$x^T Dx = -4x_2^2 - (x_1+x_3)^2 + 2\alpha x_2(x_1+x_3)$$

The vector $x = (1, 1, 1)^T$ gives $x^T D_\alpha x > 0$ for any $\alpha > 2$, and accordingly $x = (1, -1, 1)^T$ gives $x^T D_\alpha x > 0$ for any $\alpha > 2$. Therefore, for these values D_α is indeterminate.

For the edge case of $\alpha = \pm 2$, we have two zero eigenvalues (since all three rows are linearly dependent) and one negative eigenvalue (since $tr(D_\alpha) < 0$, which has D_α as NSD.

b) Firstly, E_α is strictly diagonally dominant for $|\alpha| < 1$ and diagonally dominant for $|\alpha| = 1$, with positive values on the diagonal. Therefore, E_α is PD for $|\alpha| < 1$ and PSD for $|\alpha| = 1$.
For all other α, note that:

$$x^T E_\alpha x = \underbrace{x_1^2 + 2\alpha x_1 x_2 + x_2^2}_{(1)} + \underbrace{x_2^2 + 2x_2 x_3 + x_3^2}_{(2)} + \underbrace{x_3^2 + 2\alpha x_3 x_4 + x_4^2}_{(3)}$$

For any $\alpha > 1$, parts (1), (2) and (3) are greater than $(x_1 + x_2)^2$, $(x_2 + x_3)^2$ and $(x_3 + x_4)^2$ accordingly, and therefore greater than 0 for any $x = (x_1, x_2, x_3, x_4)^T \neq \overrightarrow{0}$ and E_α is PD. The same logic holds with a replacement of variables: $\tilde{x}_i = -x_i$ for $i = 1, 2, 3, 4$, and then for any $\alpha < -1$ we have that E_α is also PD.

Problem 4. Let A be an $n \times n$ positive semidefinite matrix.

(a) Show that for any $i \neq j$

$$A_{i,i} A_{j,j} \geq A_{i,j}^2$$

(b) Show that if for some $i \in \{1, ..., n\}$ $A_{i,i} = 0$ then the i-th row of A consists of zeros.

Solution

(a) Since A is PSD, for any $x \in \mathbb{R}^n$ it holds that $x^T A x \geq 0$. Particularly, let us choose x such that $x_i \neq 0$ and $x_j = \neq 0$, while all other entries are 0. We now have:

$$x^T A x = A_{i,i} x_i^2 + A_{j,j} x_j^2 + 2x_i x_j A_{i,j}$$

Assuming $A_{i,j} > 0$ we can select $x_i = \sqrt{A_{j,j}}$ and $x_j = -\sqrt{A_{i,i}}$. This works because for a PSD matrix, all elements along the diagonal (including $A_{i,i}$ and $A_{j,j}$ are non-negative). We get:

$$\begin{aligned}
A_{i,i} x_i^2 + A_{j,j} x_j^2 + 2x_i x_j A_{i,j} &= (\sqrt{A_{i,i}} x_i + \sqrt{A_{j,j}} x_j)^2 - 2x_i x_j \sqrt{A_{i,i} A_{j,j}} + 2x_i x_j A_{i,j} = \\
&= (\sqrt{A_{i,i} A_{j,j}} - \sqrt{A_{j,j} A_{i,i}})^2 + 2A_{i,i} A_{j,j} - 2\sqrt{A_{i,i} A_{j,j}} A_{i,j} = \\
&= 2A_{i,i} A_{j,j} - 2\sqrt{A_{i,i} A_{j,j}} A_{i,j} \geq 0 \\
&\Rightarrow 2A_{i,i} A_{j,j} \geq 2\sqrt{A_{i,i} A_{j,j}} A_{i,j} \\
&\Rightarrow \sqrt{A_{i,i} A_{j,j}} \geq A_{i,j} \\
&\Rightarrow A_{i,i} A_{j,j} \geq A_{i,j}^2
\end{aligned}$$

On the other hand, if $A_{i,j} < 0$ we can select $x_i = \sqrt{A_{j,j}}$ and $x_j = \sqrt{A_{i,i}}$, and we get:

$$\begin{aligned}
A_{i,i} x_i^2 + A_{j,j} x_j^2 + 2x_i x_j A_{i,j} &= (\sqrt{A_{i,i}} x_i - \sqrt{A_{j,j}} x_j)^2 + 2x_i x_j \sqrt{A_{i,i} A_{j,j}} + 2x_i x_j A_{i,j} = \\
&= (\sqrt{A_{i,i} A_{j,j}} - \sqrt{A_{j,j} A_{i,i}})^2 + 2A_{i,i} A_{j,j} + 2\sqrt{A_{i,i} A_{j,j}} A_{i,j} = \\
&= 2A_{i,i} A_{j,j} + 2\sqrt{A_{i,i} A_{j,j}} A_{i,j} \geq 0 \\
&\Rightarrow 2A_{i,i} A_{j,j} \geq -2\sqrt{A_{i,i} A_{j,j}} A_{i,j} \\
&\Rightarrow \sqrt{A_{i,i} A_{j,j}} \geq -A_{i,j} \\
&\Rightarrow A_{i,i} A_{j,j} \geq A_{i,j}^2
\end{aligned}$$

Therefore, the inequality holds for any $A_{i,j} \neq 0$. For $A_{i,j} = 0$, the inequality holds trivially since all elements along the diagonal of a PSD matrix are non-negative.
One last case we have to consider is where $A_{i,i} = A_{i,j} = 0$. In this case: $0 \leq x^T Ax = 2x_i x_j A_{i,j}$, and since this holds for any x_i and x_j, we are restricted to $A_{i,j} = 0$, in which case the inequality also holds.

(b) The proof for this is follows directly from the last case we have shown above. We have that $\forall j, 0 = A_{i,i}A_{j,j} \geq A_{i,j}^2$. Therefore, since $A_{i,j}^2$ is non-negative, we have $A_{i,j} = 0$ for all entries in the same row.

Problem 5. Let $Q \in \mathbb{R}^{n \times n}$ be a positive definite matrix. Show that the so-called Q-norm, defined by $||x||_Q = \sqrt{x^T Q x}$ is indeed a norm.

Solution

- Nonnegativity: since Q is PD, we have $x^T Qx > 0$ for any $x \neq 0$, and therefore the same holds for $||x||_Q$.
- Positive Homogenity: $||\lambda x||_Q = \sqrt{\lambda x^T Q \lambda x} = \sqrt{\lambda^2 x^T Q x} = |\lambda| \sqrt{x^T Qx} = |\lambda| ||x||_Q$.
- Triangle Inequality: Using the spectral decomposition theorem, we can write $Q = U^T DU = U^T D^{1/2} D^{1/2} U = U^T D^{1/2T} D^{1/2} U$, where $D^{1/2} = D^{1/2T}$ is a diagonal matrix with each value along the diagonal being the square root of the value in the respective position in D (this exists since Q is PD, which means all of its eigenvalues are positive). Using the above decomposition, we have:

$$\begin{aligned} ||x+y||_Q &= \sqrt{(x+y)^T Q(x+y)} = \sqrt{(x+y)^T U^T D^{1/2T} \underbrace{D^{1/2}U}_{S}(x+y)} = \\ &= \sqrt{(x+y)^T S^T S(x+y)} = \sqrt{(S(x+y))^T S(x+y)} = ||S(x+y)||_2 \leq \\ &\overset{(a)}{\leq} ||Sx||_2 + ||Sy||_2 = \sqrt{x^T S^T Sx} + \sqrt{y^T S^T Sy} = \sqrt{x^T Qx} + \sqrt{y^T Qy} = ||x||_Q + ||y||_Q \end{aligned}$$

  Where (a) is the triangle inequality over the vector 2-norm.

Since the Q-norm satisfies all three requirements of an induced norm, it is indeed a norm.

Problem 6. For each of the following functions, find all the stationary points and classify according to whether they are saddle points, strict/non-strict local/global minimum/maximum points.

1. $f(x_1, x_2) = 2x_1^2 + x_2^2 - 2x_1 x_2 + 2x_1^3 + x_1^4$
2. $f(x_1, x_2) = 2x_1^3 - 3x_1^2 - 6x_1 x_2(x_1 - x_2 - 1)$
3. $f(x_1, x_2) = (x_1^2 + x_2^2)e^{-x_1^2 - x_2^2}$
4. $f(x_1, x_2) = 2x_1^3 - 6x_2^2 + 3x_1^2 x_2$

Solution

1. We begin by finding the stationary points of f:

$$0 = \nabla f(x_1, x_2) = \begin{pmatrix} 4x_1 - 2x_2 + 6x_1^2 + 4x_1^3 \\ 2x_2 - 2x_1 \end{pmatrix} \Rightarrow x_2 = x_1 \Rightarrow 2x_1 + 6x_1^2 + 4x_1^3 = 0$$

$$2x_1(2x_1^2 + 3x_1 + 1) = 0 \Rightarrow x_1 = x_2 = 0, -1, -\frac{1}{2}$$

Now, we can look at the classification of the Hessian of f at each of the stationary points, in order to classify them:

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 4 + 12x_1 + 12x_1^2 & -2 \\ -2 & 2 \end{pmatrix}$$

$$\nabla^2 f(0,0) = \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix} \Rightarrow \texttt{tr}(\nabla^2 f(0,0)) = 6 > 0\texttt{, } \texttt{det}(\nabla^2 f(0,0)) = 4 > 0$$

$$\Rightarrow \texttt{both eigenvalues are strictly positive} \Rightarrow \nabla^2 f(0,0) \succ 0$$

$$\nabla^2 f(-1,-1) = \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix} \Rightarrow \text{identical to the above.}$$

$$\nabla^2 f(-\frac{1}{2}, -\frac{1}{2}) = \begin{pmatrix} 1 & -2 \\ -2 & 2 \end{pmatrix} \Rightarrow \texttt{tr}(\nabla^2 f(-\frac{1}{2}, -\frac{1}{2})) = 3 > 0\texttt{, } \texttt{det}(\nabla^2 f(-\frac{1}{2}, -\frac{1}{2})) = -2 < 0$$

$$\Rightarrow \texttt{indefinite}$$

Since the first two points have a PD Hessian, they are strict local minima (we shall check if they are global shortly). The third point has an indefinite Hessian, which means it is a saddle point.
The value of f at both $(0,0)$ and $(-1,-1)$ is 0. In order to check if these are global minima, let us look at the last part of the function: $2x_1^3 + x_1^4 = x_1^3(2 + x_1)$. This attains a negative value when $-2 < x_1 < 0$ only. The minimal value it attains can be found by differentiating the 1-d function $g(y) = y^3(2 + y)$:

$$0 = g'(y) = 6y^2 + 4y^3 \Rightarrow y^2(6 + 4y) = 0 \Rightarrow y = 0\texttt{, } y = -\frac{3}{2}$$

Let us look at the value of f at this minimal point of $x_1 = -\frac{3}{2}$:

$$f(x_1, x_2) = 2x_1^2 + x_2^2 - 2x_1x_2 + 2x_1^3 + x_1^4 = (x_1 - x_2)^2 + x_1^2 + 2x_1^3 + x_1^4$$

$$\geq x_1^2 + 2x_1^3 + x_1^4 \overset{x_1 = -\frac{3}{2}}{=} \frac{9}{4} - \frac{27}{8} + \frac{81}{16} > 0$$

This means the function is bounded from below by 0, which is the value of the strict local minma we have found. Consequentially, both these strict local minima are also global minima.

2. We begin by finding the stationary points of f:

$$0 = \nabla f(x_1, x_2) = \begin{pmatrix} 6x_1^2 - 6x_1 - 12x_1x_2 + 6x_2^2 + 6x_2 \\ -6x_1^2 + 12x_1x_2 + 6x_1 \end{pmatrix} \Rightarrow 6x_2^2 + 6x_2 = 0 \Rightarrow x_2 = 0, -1$$

$$\Rightarrow \text{the stationary points are:} (0,0), (0,-1), (1,0), (-1,-1)$$

We shall now calculate the Hessian at each of the stationary points:

$$\nabla^2 f(x_1,x_2) = \begin{pmatrix} 12x_1 - 6 - 12x_2 & -12x_1 + 12x_2 + 6 \\ -12x_1 + 12x_2 + 6 & 12x_1 \end{pmatrix}$$

$$\nabla^2 f(0,0) = \begin{pmatrix} -6 & 6 \\ 6 & 0 \end{pmatrix} \Rightarrow \texttt{tr}(\nabla^2 f(0,0)) = -6 < 0\texttt{,}\ \texttt{det}(\nabla^2 f(0,0)) = -36 < 0$$
$$\Rightarrow \texttt{both eigenvalues are strictly negative} \Rightarrow \nabla^2 f(0,0) \prec 0$$
$$\nabla^2 f(0,-1) = \begin{pmatrix} 6 & -6 \\ -6 & 0 \end{pmatrix} \Rightarrow \texttt{tr}(\nabla^2 f(0,-1)) = 6 > 0\texttt{,}\ \texttt{det}(\nabla^2 f(0,-1)) = 36 > 0$$
$$\Rightarrow \texttt{both eigenvalues are strictly positive} \Rightarrow \nabla^2 f(0,-1) \succ 0$$
$$\nabla^2 f(1,0) = \begin{pmatrix} 6 & -6 \\ -6 & 12 \end{pmatrix} \Rightarrow \texttt{tr}(\nabla^2 f(1,0)) = 18 > 0\texttt{,}\ \texttt{det}(\nabla^2 f(1,0)) = 108 > 0$$
$$\Rightarrow \texttt{both eigenvalues are strictly positive} \Rightarrow \nabla^2 f(1,0) \succ 0$$
$$\nabla^2 f(-1,-1) = \begin{pmatrix} 18 & 6 \\ 6 & -12 \end{pmatrix} \Rightarrow \texttt{tr}(\nabla^2 f(-1,-1)) = 6 > 0\texttt{,}\ \texttt{det}(\nabla^2 f(-1,-1)) = -252 < 0$$
$$\Rightarrow \texttt{indefinite}$$

Finally, to determine if these points are global minima / maxima, we observe the behavior as $||\bar{x}|| \to \infty$. Specifically we consider the following vector $\bar{x} = (x_1, x_2)^T = (x, 0)^T$.

$$f(\bar{x}) = f(x,0) = 2x^3 - 3x^2 = x^2(2x-3)$$

Notice that for $x > \frac{3}{2}$, $f(\bar{x})$ is strictly positive and for $x \geq 2$: $f(\bar{x}) = x^2(2x-3) \geq x^2 \overset{x\to\infty}{\to} \infty$.
Additionally, notice that for $x < \frac{3}{2}$, $f(\bar{x})$ is strictly negative and for $x \leq 1$: $f(\bar{x}) = x^2(2x-3) \leq -x^2 \overset{x\to-\infty}{\to} -\infty$.
We have shown that not only the function is non-coercive, it reaches both $+\infty$ and $-\infty$. Hence the stationary points are local minima / maxima points. $(0,0)$ strict local maximum, $(0,-1)$ strict local minimum, $(1,0)$ strict local minimum, $(-1,-1)$ saddle point.

3. Note that f receives a constant value on circles centered on the origin, where $(x_1^2+x_2^2) = R^2$. Therefore, we can look at f as a 1-d function: $f(x_1,x_2) = g(R) = R^2 e^{-R^2}$. Now, we can find optima for this 1-d function:

$$0 = g'(R) = 2Re^{-R^2} - 2R^3 e^{-R^2} = 2Re^{-R^2}(1-R^2)$$
$$\Rightarrow \texttt{minimum at}\ R = 1, R = 0$$
$$g''(R) = 2e^{-R^2}(2R^4 - 5R^2 + 1)$$
$$g''(0) = 2 > 0,\ g''(1) = -4\frac{1}{e} < 0$$

Therefore, $g(R)$ has a strict local maximum at $R = 1$, which means f has local maxima along the circle $x_1^2 + y_1^2 = 1$. Note that these maxima are non-strict, since along the

circle the value is constant. Additonally, $g(R)$ has a strict local minimum at $R = 0$, which means that the point $x_1 = x_2 = 0$ is a strict local minimum of f.
Lets now observe the behavior as $||\bar{x}|| \to \infty$, $f(\bar{x}) = g(||\bar{x}||^2) = g(R) = R^2 e^{-R^2}$. Not only is $g(R)$ strictly positive, we also know that this 1-d function goes to 0^+ as $R \to \infty$. Following the above, the circle with $R = 1$ also defines the set of global maxima and the point defined by $R = 0 (x_1 = x_2 = 0)$ is a strict global minimum of f.

4. Let us look at the gradient of f:

$$0 = \nabla f(x_1, x_2) = \begin{pmatrix} 6x_1^2 + 6x_1 x_2 \\ -12x_2 + 3x_1^2 \end{pmatrix} \to x_2 = \frac{x_1^2}{4}\texttt{,}\ 6x_1^2 + \frac{3}{2}x_1^3 = 0$$
$$\Rightarrow x_1^2(4 + x_1) = 0 \to \texttt{stationary points are:}\quad (0,0), (-4,4)$$

Let us look at the Hessian at both points:

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 12x_1 + 6x_2 & 6x_1 \\ 6x_1^2 & -12 \end{pmatrix}$$
$$\nabla^2 f(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & -12 \end{pmatrix} \preceq 0 \texttt{ since the eigenvalues are } 0, -12 \Rightarrow NSD$$
$$\nabla^2 f(-4,4) = \begin{pmatrix} -24 & -24 \\ 96 & -12 \end{pmatrix} \Rightarrow \texttt{tr}(\nabla^2 f(-4,4)) = -36 < 0\texttt{, det}(\nabla^2 f(-4,4)) > 0$$
$$\Rightarrow \texttt{Hessian is indefinite}$$

Therefore, $(-4,4)$ is a saddle point, and we have to investigate further into $(0,0)$, to find if it is a local maximum or a saddle point as well.
Note that along the line $(t,0)$, the function goes up as t rises and down as t goes negative. Therefore, we can conclude $(0,0)$ is also a saddle point.

Problem 7. Let A^α be the $n \times n$ matrix $(n > 1)$ defined by:

$$A^\alpha_{i,j} = \begin{cases} \alpha\texttt{,} & i = j \\ 1\texttt{,} & i \neq j \end{cases}$$

Show that A^α is positive semi-definite if and only if $\alpha \geq 1$.

Solution We can look at A^α as:

$$A^\alpha = (\alpha - 1) I_{n \times n} + \mathbb{1}_{n \times n}$$

Where $\mathbb{1}$ is a matrix for which all elements are 1, and I is the identity matrix. Note that $\mathbb{1}_{n \times n}$ is a positive semi-definite matrix, since it is rank 1 and therefore has $n-1$ zero eigenvalues and one eigenvalue of n (since a matrix with all identical rows has one eigenvalue which is the sum of the rows). The matrix $(\alpha - 1) I_{n \times n}$ has n eigenvalues which are all $\alpha - 1$, since it is diagonal and therefore has its eigenvalues along its main diagonal. Therefore, $(\alpha - 1) I_{n \times n}$ is PD when $\alpha > 1$ and ND when $\alpha < 1$.
With the above observations, we can now start our proof:

- $\alpha \geq 1 \Rightarrow$ PSD: For $\alpha \geq 1$, we have for some $x \in \mathbb{R}^n$:

$$x^T A^\alpha x = \underbrace{x^T(\alpha - 1)I_{n\times n}x}_{\geq 0} + \underbrace{x^T \mathbb{1}_{n\times n}x}_{\geq 0} \geq 0$$

Where the underbraces follow from our observations above. Since $x^T A^\alpha x \geq 0$ for any $x \in \mathbb{R}^n$, A^α is PSD.

- PSD $\Rightarrow \alpha \geq 1$: Let us suppose $\alpha < 1$. Consider the vector $x = (1, -1, 0, ..., 0)^T \in \mathbb{R}^n$:

$$x^T A^\alpha x = (1^2 + (-1)^2)\alpha + 2 \cdot 1 \cdot (-1) = 2\alpha - 2 \overset{\alpha<1}{<} 0$$

This contradicts the fact that A^α is PSD (since $x^T A^\alpha x \geq 0$ for all $x \in \mathbb{R}^n$.

Problem 8. Let f be a twice continuously differentiable function on $U \in \mathbb{R}^n$, and let x_0 be an interior point of U, such that $\nabla^2 f(x_0) \succ 0$. Prove that there exists $R > 0$ such that the Hessian $\nabla^2 f(x)$ is positive definite for every point $x \in B(x_0, R)$.

Solution Since f is twice continuously differentiable over U, each element in $\nabla^2 f$ is continuous in x. Let us define $g(x) = y^T \nabla^2 f(x) y$ for any $y \in \mathbb{R}^n$. This is a function $\mathbb{R}^n \to \mathbb{R}$, which is continuous in x since its value is a linear combination of the elements of the Hessian, which are all continuous in x.
Assume the value of $g(x)$ at x_0 is ϵ_0. We have $\epsilon_0 = g(x_0) = y^T \nabla^2 f(x_0) y > 0$ for any y, since $\nabla^2 f$ is PD. By definition of a continuous function, for any ϵ there exists some $R > 0$, such that for any $||x - x_0|| < R$, $|g(x) - g(x_0)| < \epsilon$. Specifically, this holds for ϵ_0. This means that for any point x such that $||x - x_0|| < R$, or alternatively, $x \in B(x_0, R)$, we have $y^T \nabla^2 f(x) y = g(x) > 0$, for any y. Therefore, for any $x \in B(x_0, R)$, $\nabla^2 f(x) \succ 0$.