אופטימיזציה 1 - 098311

גיליון בית מס' 2 - חורף תשפ"א 2021

אסף גנדלר 301727715 עידו צ'רנינסקי 312544596

5 בנובמבר 2020

Problem 1:

a)

Find the maximum and minimum of $f(x) = a^T x$ over the set: $B[0, \alpha] = \{x \in \mathbb{R}^n : \|x\|_2 \leq \alpha\}$.

First, let's explain why the maximum and minimum are attained.

$f(x) = a^T x = \sum_{i=1}^{n} a_i x_i$ is a continuous function over $\mathbb{R}^n$.

$B[0, \alpha] \subset \mathbb{R}^n$ is a non empty compact set.

Therefore, using the Weierstrass Theorem, $f(x)$ attains a minimum and a maximum value over $B[0, \alpha]$.

Let's show that all the partial derivatives of $f(x)$ exist over all $\mathbb{R}^n$:

$$
\begin{aligned}
f(x; e_i) &= \lim_{t \to 0} \frac{f(x + te_i) - f(x)}{t} = \lim_{t \to 0} \frac{a^T (x + te_i) - a^T x}{t} = \\
&= \lim_{t \to 0} \frac{a^T x + a^T te_i - a^T x}{t} = \lim_{t \to 0} \frac{a^T te_i}{t} = \lim_{t \to 0} a^T e_i = a_i
\end{aligned}
$$

$$\nabla f(x) = a$$

The optimum points can be obtained on the boundary or in an interior point.

Let's first check the possibility of an interior point, because all the partial derivatives exist in all $\mathbb{R}^n$ in must be attained in a stationery point, let's find the stationery points:

$$\nabla f(x) = 0$$

$$a = 0$$

a is a non zero vector, thus, there are no stationery points, and the optimum points are not interior points.

Let's check on the boundary, we need to solve:

$$\max_{x \in R^n} / \min \left\{a^T x : \|x\|_2 = \alpha\right\}$$

using the Cauchy-Schwartz Inequality:

$$\left|a^T x\right| \leq \|a\|_2 \|x\|_2 = \alpha \|a\|_2$$

$$-\alpha \|a\|_2 \le a^T x \le \alpha \|a\|_2$$

in addition for $x = \alpha \frac{a}{\|a\|_2}$:

$$a^T x = a^T \alpha \frac{a}{\|a\|_2} = \alpha \frac{\|a\|_2^2}{\|a\|_2} = \alpha \|a\|_2$$

and for $x = -\alpha \frac{a}{\|a\|_2}$

$$a^T x = -a^T \alpha \frac{a}{\|a\|_2} = -\alpha \frac{\|a\|_2^2}{\|a\|_2} = -\alpha \|a\|_2$$

We have found an upper and lower bounds that are attained, therefore:

$$\max_{x \in R^n} / \min \left\{ a^T x : \|x\| = \alpha \right\} = \pm \alpha \|a\|$$

and:

$$\arg \max_{x \in R^n} / \min \left\{ a^T x : \|x\| = \alpha \right\} = \pm \alpha \frac{a}{\|a\|}$$

Since the maximum/minimum are attained, and since we didn't find any other optimum points those are the global maximum and minimum of $f(x)$.

b)

Find the optimum points of $f(x, y) = 3x - 5y$ over the constraint $U = \{x, y : x^2 + y^2 - 2y - 3 \le 0\}$.
Let's look at U:

$$x^2 + y^2 - 2y - 3 = x^2 + y^2 - 2y - 1 + 1 - 3 = x^2 + (y-1)^2 - 4$$

$$U = \left\{ x, y : x^2 + (y-1)^2 \le 4 \right\} = B \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}, 2 \right]$$

Define:

$$a = \begin{pmatrix} 3 \\ -5 \end{pmatrix}, z = \begin{pmatrix} x \\ y-1 \end{pmatrix}$$

$$a^T z = 3x - 5y + 5$$

Notice that:

$$\arg\max / \min \left\{ 3x - 5y + 5 : (x, y)^T \in U \right\} = \arg\max / \min \left\{ 3x - 5y : (x, y)^T \in U \right\}$$

Therefore, we can solve the left problem instead. Since $z = \begin{pmatrix} x \\ y-1 \end{pmatrix}$:

$$(x, y)^T \in U \iff \|z\|_2^2 \le 4 \iff \|z\|_2 \le 2$$

Thus, the left problem is given by:

$$\max / \min \left\{a^T z : \|z\|_2 \leq 2\right\}$$

which is the exact problem we solved at section a and the solution is:

$$z = \pm 2 \frac{a}{\|a\|_2} = \pm 2 \frac{\begin{pmatrix} 3 \\ -5 \end{pmatrix}}{\sqrt{3^2 + 5^2}} = \pm 2 \frac{\begin{pmatrix} 3 \\ -5 \end{pmatrix}}{\sqrt{34}} = \pm \begin{pmatrix} \frac{6}{\sqrt{34}} \\ -\frac{10}{\sqrt{34}} \end{pmatrix}$$

Hence, the solution to the original problem is attained at:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \pm \begin{pmatrix} \frac{6}{\sqrt{34}} \\ 1 - \frac{10}{\sqrt{34}} \end{pmatrix}$$

and the maximum\minimum values are (maximum :

$$f(x, y) = \pm 3 \frac{6}{\sqrt{34}} \mp 5 \left(1 - \frac{10}{\sqrt{34}}\right)$$

Problem 2:

a)

Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$ be two symmetric matrices. Prove that both A and B are positive semi-definite if and only if:

$$C \triangleq \begin{pmatrix} A & 0_{n\times m} \\ 0_{m\times n} & B \end{pmatrix} \succeq 0$$

First direction:

$A, B \succeq 0 \Rightarrow C \succeq 0$

Proof:

Let $x \in \mathbb{R}^{n+m}$:

$$x = \begin{pmatrix} | \\ x_1 \\ | \\ x_2 \\ | \end{pmatrix}$$

where $x_1 \in \mathbb{R}^n$ and $x_2 \in \mathbb{R}^m$

$$\begin{aligned} x^TCx &= \sum_{i=1}^{n+m}\sum_{j=i}^{n+m} x_ix_jC_{i,j} = \sum_{i=1}^{n}\sum_{j=i}^{n} x_ix_jC_{i,j} + \sum_{i=1}^{n}\sum_{j=n+1}^{n+m} x_ix_jC_{i,j} + \sum_{i=n+1}^{n+m}\sum_{j=1}^{n} x_ix_jC_{i,j} + \sum_{i=n+1}^{n+m}\sum_{j=n+1}^{n+m} x_ix_jC_{i,j} \\ &= \sum_{i=1}^{n}\sum_{j=i}^{n} x_{1,i}x_{1,j}A_{i,j} + \sum_{i=1}^{n}\sum_{j=n+1}^{n+m} x_ix_j \cdot 0 + \sum_{i=n+1}^{n+m}\sum_{j=1}^{n} x_ix_j \cdot 0 + \sum_{i=1}^{m}\sum_{j=1}^{m} x_{2,i}x_{2,j}B_{i,j} \\ &= x_1^TAx_1 + x_2^TBx_2 \geq 0 + 0 = 0 \end{aligned}$$

The last inequality holds since A and B are positive semi-definite meaning:

$$\forall x_1 \in \mathbb{R}^n, x_2 \in \mathbb{R}^m : x_1^TAx_1 \geq 0, x_2^TBx_2 \geq 0$$

Second direction:

$C \succeq 0 \Rightarrow A, B \succeq 0$

Let's assume by contradiction that at least one of the matrices A and B, is not positive semi-definite (without the loss of generality we choose A):

thus:

$$\exists x_1 \in \mathbb{R}^n : x_1^TAx < 0$$

by choosing:

$$x = \begin{pmatrix} | \\ x_1 \\ | \\ 0_m \\ | \end{pmatrix}$$

we get:

$$x^T C x = x_1^T A x_1 + 0_m^T B 0_m = x_1^T A x_1 < 0$$

and that is a contradiction to the given fact that $C \succeq 0$.

To summarize, we got:

$$\boxed{A, B \succeq 0 \iff C \succeq 0}$$

b)

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

First, it can be seen that C is a symmetric matrix

Let's try to rule out options:

C is **not** negative definite\semi-definite since it's diagonal elements are all positive (hence its trace is positive).

C is **not** positive definite since it does not have a full rank (hence its determinant equals to zero and not positive).

We could not determine easily whether the matrix is positive semi-definite or indefinite. Let's try by definition:

Let $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$.

$$\begin{aligned} x^T C x &= \sum_{i=1}^{3}\sum_{j=1}^{3} x_i \cdot x_j \cdot C_{i,j} = 2x_1^2 + 3x_2^2 + 3x_3^2 + 4x_1x_2 + 4x_1x_3 + 6x_2x_3 \\ &= 2\left(x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3\right) + x_2^2 + x_3^2 + 2x_2x_3 \\ &= 2\left(x_1 + x_2 + x_3\right)^2 + \left(x_2 + x_3\right)^2 \geq 0 \end{aligned}$$

hence:

$$\boxed{C \succeq 0}$$

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

First, it can be seen that D is a symmetric matrix

Let's try to rule out options:

D is **not** negative definite\semi-definite since it's diagonal elements are all positive (hence its trace is positive)

D is **not** positive definite since it does not have a full rank (hence its determinant equals to zero and not positive).

D is diagonally dominant (not strictly) with positive diagonal elements thus, it is **positive semi-definite**.

c)

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

First, E is a symmetric matrix $\forall \alpha \in \mathbb{R}$.

Moreover, $\forall \alpha \in \mathbb{R}$, E is not negative definite\semi-definite since it's diagonal elements are all positive (hence its trace is positive).

Thus, we have three options:

Positive definite

As shown in class, E is positive definite if and only if $D_1(E) > 0, D_2(E) > 0, D_3(E) > 0, D_4(E) > 0$

$$D_1(E) = 1 > 0,\ \forall\alpha \in \mathbb{R}$$

$$\begin{aligned} D_2(E) &= 2 - \alpha^2 > 0 \\ \Rightarrow& \alpha^2 < 2 \\ \Rightarrow& (*) - \sqrt{2} < \alpha < \sqrt{2} \end{aligned}$$

$$D_3(E) = 1 \cdot \left(4 - \alpha^2\right) - \alpha \cdot (2\alpha - 0) = 4 - \alpha^2 - 2\alpha^2 = 4 - 3\alpha^2 > 0$$

$$\Rightarrow 3\alpha^2 < 4$$

$$\Rightarrow \alpha^2 < \frac{4}{3}$$

$$\Rightarrow (**) - \frac{2}{\sqrt{3}} < \alpha < \frac{2}{\sqrt{3}}$$

$$\begin{aligned} D_4(E) = 1\left(2\left(2 - \alpha^2\right) - \alpha \cdot \alpha\right) - \alpha\left(\alpha\left(2 - \alpha^2\right)\right) = \\ = 4 - 2\alpha^2 - \alpha^2 - 2\alpha^2 + \alpha^4 = \alpha^4 - 5\alpha^2 + 4 > 0 \end{aligned}$$

let's mark $t = \alpha^2$:

$$t^2 - 5t + 4 > 0$$

$$(t - 4)(t - 1) > 0$$

$$t < 1 \cup t > 4$$

now go back to α :

$$\alpha^2 < 1 \cup \alpha^2 > 4$$

$$(***) - 1 < \alpha < 1 \cup \alpha < -2 \cup \alpha > 2$$

the intersection of $(*)$, $(**)$ and $(***)$ is:

$$-1 < \alpha < 1$$

hence for $|\alpha| < 1$, E is **positive definite**.

Positive semi-definite

for $\alpha = \pm 1$, E is diagonally dominant (not strictly) with positive diagonal elements hence, it is positive semi-definite.

Indefinite

let's take a vector $x \in R^4$ ant try to classify by definition:

$$\begin{aligned} x^T E x &= x_1^2 + 2x_2^2 + 2x_3^2 + x_4^2 + 2\alpha x_1 x_2 + 2\alpha x_2 x_3 + 2\alpha x_3 x_4 \\ &= \left(x_1^2 + 2\alpha x_1 x_2 + x_2^2\right) + \left(x_2^2 + 2\alpha x_2 x_3 + x_3^2\right) + \left(x_3^2 + 2\alpha x_3 x_4 + x_4^2\right) \end{aligned}$$

let's pick the vector:

$$x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

for this vector we get:

$$x^T E x = 6 + 6\alpha$$

notice that for $\alpha < -1$:

$$x^T E x = 6 + 6\alpha < 0$$

hence E is not positive semi-definite for $\alpha < -1$

we can also pick:

$$x = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

for this vector we get:

$$x^T E x = 6 - 6\alpha$$

notice that for $\alpha > 1$:

$$x^T E x = 6 - 6\alpha < 0$$

hence E is not positive semi-definite for $\alpha > 1$ neither.

conclusion: for $|\alpha| > 1$, E is indefinite.

let's summarize:

$$E = \begin{cases} PD & |\alpha| < 1 \\ PSD & |\alpha| = 1 \\ ID & |\alpha| > 1 \end{cases}$$

Problem 3:

a)

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$$f(x_1, x_2) = 2x_1^3 - 3x_1^2 - 6x_1x_2(x_1 - x_2 - 1) = 2x_1^3 - 3x_1^2 - 6x_1^2x_2 + 6x_1x_2^2 + 6x_1x_2$$

First of all, f is a continuous and differentiable function over all $\mathbb{R}^2$ therefore any interior optimum point must be a stationery point. In addition, every point in $\mathbb{R}^2$ is an interior point thus, every optimum point must be a stationery point.

Let's find the stationery points:

The gradient of f:

$$\begin{aligned}\frac{\partial f(x_1, x_2)}{\partial x_1} &= 6x_1^2 - 6x_1 - 12x_1x_2 + 6x_2^2 + 6x_2 = \\ &= 6\left(x_1^2 - 2x_1x_2 + x_2^2 - (x_1 - x_2)\right) = \\ &= 6\left((x_1 - x_2)^2 - (x_1 - x_2)\right) = 6(x_1 - x_2)((x_1 - x_2) - 1)\end{aligned}$$

$$\frac{\partial f(x_1, x_2)}{\partial x_2} = -6x_1^2 + 12x_1x_2 + 6x_1 = -6x_1(x_1 - 2x_2 - 1)$$

Compare the gradient to zero:

$$6(x_1 - x_2)(x_1 - x_2 - 1) = 0 \longrightarrow (x_1 = x_2) \cup (x_1 = x_2 + 1)$$

$$-6x_1(x_1 - 2x_2 - 1) = 0 \longrightarrow (x_1 = 0) \cup (x_1 = 2x_2 + 1)$$

To find the stationary points, we need to find all the possible combinations that zeros the gradient.

Thus, the stationary points are:

$$(0, 0), (0, -1), (-1, -1), (1, 0)$$

To classify the points, let's compute the hessian:

$$\frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} = 12x_1 - 6 - 12x_2$$

$$\frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} = 12x_1$$

$$\frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2} = -12x_1 + 12x_2 + 6$$

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

let's look at the hessian at each of the stationery points:

$$\nabla^2 f(0,0) = 6 \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$$

It's enough to look at (multiplication by positive scalar does not change any inequalities):

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

$$Tr(A) = -1 < 0$$

$$det(A) = -1 < 0$$

Since it is a $2x2$ matrix with negative trace and determinant, A is negative definite (as seen in class). Hence, $(0,0)$ is a strict local maximum point.

$$\nabla^2 f(0,-1) = 6 \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$$

It's enough to look at (same reason as before):

$$B = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$$

$$Tr(B) = 1 > 0$$

$$det(B) = -1 < 0$$

Since it is a $2x2$ matrix with positive trace and negative determinant, B is indefinite. Hence $(0,-1)$ is a saddle point.

$$\nabla^2 f(1,0) = 6 \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

It is enough to look at:

$$C = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

$$Tr(C) = 3 > 0$$

$$det(B) = 2 - 1 = 1 > 0$$

Since it is a $2x2$ matrix with positive trace and determinant, C is positive definite, hence $(1,0)$ is a strict local minimum point.

$$\nabla^2 f(-1,-1) = 6\begin{pmatrix} -1 & 1 \\ 1 & -2 \end{pmatrix}$$

It is enough to look at:

$$D = \begin{pmatrix} -1 & 1 \\ 1 & -2 \end{pmatrix}$$

$$Tr(D) = -3 < 0$$

$$det(B) = 2 - 1 = 1 > 0$$

Since it is a $2x2$ matrix with negative trace and positive determinant, D is indefinite, hence $(0,-1)$ is a saddle point.

to find global optimum let's look at the function:

$$f(x_1,x_2) = 2x_1^3 - 3x_1^2 - 6x_1x_2(x_1 - x_2 - 1) = 2x_1^3 - 3x_1^2 - 6x_1^2x_2 + 6x_1x_2^2 + 6x_1x_2$$

Let's look at the direction $(x,0)$:

$$\lim_{x\to\infty} f(x,0) = \lim_{x\to\infty} 2x^3 - 3x^2 = \infty$$

Hence, $f(x_1,x_2)$ doesn't have a global maximum.

Let's look at the direction $(-x,0)$:

$$\lim_{x\to\infty} f(-x,0) = \lim_{x\to\infty} -2x^3 - 3x^2 = -\infty$$

Hence $f(x_1,x_2)$ doesn't have a global minimum.

To summarize:

The stationery points are:

$$\begin{cases} (0,0) & \text{strict local maximum} \\ (0,-1) & \text{saddle point} \\ (1,0) & \text{strict local minimum} \\ (-1,-1) & \text{saddle point} \end{cases}$$

There are no other local maximum/minimum points and there are no global maximum and minimum.

b)

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$$f(x_1, x_2) = \left(x_1^2 + x_2^2\right) e^{-x_1^2 - x_2^2}$$

First of all, f is a continuous and differentiable function over all $\mathbb{R}^2$ therefore any interior optimum point must be a stationery point. In addition, every point in $\mathbb{R}^2$ is an interior point thus every optimum point must be a stationery point. Let's find the stationery points:
Denote (only for convenient) $r^2 = x_1^2 + x_2^2$:

$$f(x_1, x_2) = \left(x_1^2 + x_2^2\right) e^{-x_1^2 - x_2^2} = r^2 e^{-r^2}$$

Let's calculate the gradient:

$$\frac{\partial f}{\partial x_1} = 2x_1 e^{-r^2} - 2r^2 x_1 e^{-r^2} = 2x_1 e^{-r^2}\left(1 - r^2\right)$$

$$\frac{\partial f}{\partial x_2} = 2x_2 e^{-r^2}\left(1 - r^2\right)$$

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 e^{-r^2}\left(1 - r^2\right) \\ 2x_2 e^{-r^2}\left(1 - r^2\right) \end{pmatrix} = \begin{pmatrix} 2x_1 e^{-\left(x_1^2 + x_2^2\right)}\left(1 - \left(x_1^2 + x_2^2\right)\right) \\ 2x_2 e^{-\left(x_1^2 + x_2^2\right)}\left(1 - \left(x_1^2 + x_2^2\right)\right) \end{pmatrix}$$

Stationary points are attained at two cases:

$$\boxed{\begin{gathered} x_1 = x_2 = 0 \Rightarrow (0, 0) \\ r^2 = x_1^2 + x_2^2 = 1 \Rightarrow \left\{(x_1, x_2) \,|\, x_1^2 + x_2^2 = 1\right\} \end{gathered}}$$

Let's calculate the hessian to classify the points:

$$\begin{aligned} \frac{\partial^2 f}{\partial x_1^2} &= 2\left[e^{-r^2}\left(1 - r^2\right) - 2x_1^2 e^{-r^2}\left(1 - r^2\right) - 2x_1^2 e^{-r^2}\right] \\ &= 2e^{-r^2}\left(1 - r^2 - 2x_1^2 + 2x_1^2 r^2 - 2x_1^2\right) \\ &= 2e^{-r^2}\left(1 - r^2 - 4x_1^2 + 2x_1^2 r^2\right) \end{aligned}$$

$$\frac{\partial^2 f}{\partial x_2^2} = 2e^{-r^2}\left(1 - r^2 - 4x_2^2 + 2x_2^2 r^2\right)$$

$$\begin{aligned}
\frac{\partial^2 f}{\partial x_1 \partial x_2} &= 2x_1\left[-2x_2 e^{-r^2}\left(1 - r^2\right) - 2x_2 e^{-r^2}\right] \\
&= -4x_1 x_2 e^{-r^2}\left(1 - r^2 + 1\right) \\
&= -4x_1 x_2 e^{-r^2}\left(2 - r^2\right)
\end{aligned}$$

$$\nabla^2 f\left(x_1, x_2\right) = \begin{pmatrix} 2e^{-r^2}\left(1 - r^2 - 4x_1^2 + 2x_1^2 r^2\right) & -4x_1 x_2 e^{-r^2}\left(2 - r^2\right) \\ -4x_1 x_2 e^{-r^2}\left(2 - r^2\right) & 2e^{-r^2}\left(1 - r^2 - 4x_2^2 + 2x_2^2 r^2\right) \end{pmatrix}$$

$$\nabla^2 f\left(0, 0\right) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$det\left(\nabla^2 f\left(0, 0\right)\right) = Tr\left(\nabla^2 f\left(0, 0\right)\right) = 4 > 0$$

$$\Rightarrow \nabla^2 f\left(0, 0\right) \succ 0 \Rightarrow \left(0, 0\right) \text{ is a strict local minimum with the value:}$$

$$f\left(0, 0\right) = 0$$

$$\begin{aligned}
\nabla^2 f\left(x_1^2 + x_2^2 = 1\right) &= \begin{pmatrix} 2e^{-1}\left(1 - 1 - 4x_1^2 + 2x_1^2\right) & -4x_1 x_2 e^{-1}\left(2 - 1\right) \\ -4x_1 x_2 e^{-1}\left(2 - 1\right) & 2e^{-1}\left(1 - 1 - 4x_2^2 + 2x_2^2\right) \end{pmatrix} \\
&= \begin{pmatrix} -4e^{-1} x_1^2 & -4x_1 x_2 e^{-1} \\ -4x_1 x_2 e^{-1} & -4e^{-1} x_2^2 \end{pmatrix}
\end{aligned}$$

$$det\left(\nabla^2 f\left(x_1^2 + x_2^2 = 1\right)\right) = 16e^{-2} x_1^2 x_2^2 - 16e^{-2} x_1^2 x_2^2 = 0 \leq 0$$

$$Tr\left(\nabla^2 f\left(x_1^2 + x_2^2 = 1\right)\right) = -2e^{-1}\left(x_1^2 + x_2^2\right) = -2e^{-1} \leq 0$$

$$\Rightarrow \nabla^2 f\left(x_1^2 + x_2^2 = 1\right) \preceq 0 \Rightarrow \left\{\left(x_1, x_2\right) | x_1^2 + x_2^2 = 1\right\} \text{ are non-strict local maximums with the value:}$$

$$f\left(x_1^2+x_2^2=1\right)=1e^{-1}=e^{-1}$$

To find global optimum let's look at the function:

$$f\left(x_1,x_2\right)=\left(x_1^2+x_2^2\right)e^{-x_1^2-x_2^2}$$

$$\forall x_1,x_2\in\mathbb{R}:f\left(x_1,x_2\right)\geq 0$$

and:

$$f\left(x_1,x_2\right)=0\iff x_1=x_2=0$$

Therefore the stationery point we have found before $(0,0)$ is actually a strict global minimum point.

In addition:

$$\lim_{|x_1,x_2|_2\to\infty}f\left(x_1,x_2\right)=\lim_{|x_1,x_2|_2\to\infty}\left(x_1^2+x_2^2\right)e^{-x_1^2-x_2^2}=\lim_{\left|(x_1,x_2)^T\right|_2\to\infty}\left|(x_1,x_2)^T\right|_2^2e^{-\left|(x_1,x_2)^T\right|_2^2}=0$$

and the local maximum points we have found:

$$x_1^2+x_2^2=1$$

achieves a higher value:

$$f\left(x_1^2+x_2^2=1\right)=e^{-1}>0$$

Thus, the points on the circle are actually global (non strict) maximum points.

To summarize:

The stationery points are:

$$\begin{cases}(0,0) & \text{strict local minimum}\\ \{(x_1,x_2)\in\mathbb{R}^2|x_1^2+x_2^2=1\} & \text{local maximum}\end{cases}$$

There are no other local maximum/minimum points.

$(0,0)$ is a strict global minimum point.

$\{x_1,x_2\in\mathbb{R}:x_1^2+x_2^2=1\}$ are global maximum points

Problem 4:

Let $f:\mathbb{R}^n\to\mathbb{R}$ be a twice continuously differentiable function such that:

$$\forall x\in\mathbb{R}^n:\nabla^2 f(x)\succ 0$$

Let $x\in\mathbb{R}^n$ be a stationary point of f, meaning:

$$\nabla f(x)=0_n$$

Let $y\in R^n$ such that:

$$y\neq x$$

Denote $r=1.1\cdot\|y-x\|_2$ and the ball $B(x,r)$. We know that:

- $B\subset\mathbb{R}^n$
- $y\in B$ (since $\|y-x\|_2<1.1\,\|y-x\|_2=r$)

In addition, $f(x)$ is twice continuously differentiable over $\mathbb{R}^n$,hence, by a the linear approximation theorem:

$$\begin{aligned}&\exists\xi\in[x,y]:\\&\qquad f(y)=f(x)+\underbrace{\nabla f(x)^T}_{0_n}(y-x)+\frac{1}{2}(y-x)^T\nabla^2 f(\xi)(y-x)\\&\qquad=f(x)+\frac{1}{2}\underbrace{(y-x)^T\nabla^2 f(\xi)(y-x)}_{(*)>0}\\&\qquad>f(x)\end{aligned}$$

hence:

$$\boxed{\forall y\in\mathbb{R}^n:f(x)<f(y)\Rightarrow x\text{ is a strict global minimum}}$$

$(*)$Since $\nabla^2 f(x)\succ 0$ for all $x\in\mathbb{R}^n$:

$$\forall c,x\in\mathbb{R}^n:c^T\nabla^2 f(x)\,c>0$$

In particular, for $c=y-x\in\mathbb{R}^n$,we get:

$$\forall y,x\in\mathbb{R}^n:(y-x)^T\nabla^2 f(x)(y-x)>0$$

Problem 5:

$f(x)$ is twice continuously differentiable function over U ,hence all combined partial derivatives are continuous over U, meaning that each cell in the hessian, $\nabla^2 f(x)$ is a continuous function of x over U.

Let's take an arbitrary vector $y \in R^n$ and define the function:

$$g(x) : U \longrightarrow \mathbb{R}$$

$$g(x) = y^T \nabla^2 f(x) y = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\nabla^2 f(x)\right)_{i,j} y_i y_j$$

As we can see, the function $g(x)$ is a linear combination of the cells of the hessian, meaning it is a linear combination of continuous functions over U and hence continuous by itself over U.

In addition, in the point $x^0 \in U$:

$$g\left(x^0\right) = y^T \nabla^2 f\left(x^0\right) y > 0$$

because:

$$\nabla^2 f\left(x^0\right) \succ 0$$

Since $g(x)$ is a continuous function over U , and $g(x^0) > 0$:

$$\exists r > 0 : \forall x \in B\left(x^0, r\right) \quad g(x) > 0$$

therefore:

$$\exists r > 0 : \forall x \in B\left(x^0, r\right) \quad y^T \nabla^2 f(x) y > 0$$

Since y was chosen arbitrary, this statement is true for every vector $y \in \mathbb{R}^n$ thus:

$$\boxed{\exists r > 0 : \forall x \in B\left(x^0, r\right) \quad \nabla^2 f(x) \succ 0}$$