MA 576 Optimization for Data Science
Homework 2

Nicolas Jorquera

02/19/2023

Problem 1

Prove: $Q \in \mathbf{R}^{nxn}$ symmetric. Then,

$$\lambda_1 x^T x \leq x^T Q x \leq \lambda_n x^T x, \quad \forall x \in \mathbf{R}^n$$

where $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$ its evas.

We know symmetric matrices are orthogonally diagonalizable, meaning that for each eigenvalue λ_i we can find a corresponding eigenvector v_i such that $v_i \cdot v_j = 0$ for $i \neq j$ and $v_i \cdot v_i = 1$

Given that $x \in \mathbf{R}^n$; we can say that for a vector $\vec{v}$ which is a basis for $\mathbf{R}^n$, and x can be written as a linear combination of these values: $\sum_{i=1}^n c_i v_i$. Therefore; we can rewrite the expressions above knowing that $x^T x = \|x\|^2 = \sum_{i=1}^n c_i^2$

$$\begin{aligned}
x^T Q x &= (\sum_{i=1}^n c_i v_i)^T Q (\sum_{i=1}^n c_i v_i) \\
&= (\sum_{i=1}^n c_i v_i^T)(\sum_{i=1}^n c_i Q v_i) \\
&= (\sum_{i=1}^n c_i v_i^T)(\sum_{i=1}^n c_i \lambda_i v_i) = c_1^2 \lambda_1 + c_2^2 \lambda_1 + ...
\end{aligned}$$

We also know $\lambda_i \leq \lambda_n$ implied by $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$.; and can then conclude $c_i^2 \lambda_i \leq c_n^2 \lambda_n$
Therefore, we can show show that:

$$(\sum_{i=1}^n c_i^2)\lambda_1 \leq c_1^2 \lambda_1 + c_2^2 \lambda_1 + ... \leq (\sum_{i=1}^n c_i^2)\lambda_1$$

Therefore we can prove the statement above, as we know that $x^T x = \sum_{i=1}^n c_i^2$

Problem 2

Prove: $Q \in \mathbf{R}^{nxn}$ and $R \in \mathbf{R}^{mxm}$ symmetric. Prove that

$$Q, R \text{ are psd} \quad \iff \quad A = \begin{pmatrix} Q & 0 \\ 0^T & R \end{pmatrix} \text{ is psd,}$$

where 0 denotes the zero matrix of dimension nxm

A matrix is **Positive Symmetric Definite** if the eigenvalues are non-negative. In order to prove A is Positive Symmetric Definite, we will show that A is symmetric and it's eigenvalues are positive:

(1) $A = A^T$, therefore the matrix is symmetric

(2). The eigenvalues of A can be found by solving $\|A - XI\| = 0$

$$\begin{aligned}
\|A - XI\| &= \begin{pmatrix} Q & 0 \\ 0^T & R \end{pmatrix} - X \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} Q - \lambda I_n & 0 \\ 0 & R - \lambda I_m \end{pmatrix} \\
&= |Q - \lambda I_n||R - \lambda I_m|
\end{aligned}$$

Because $|A - \lambda I_n|$ and $|Q - \lambda I_n|$ have the same form, we can conclude that the eigenvalues of A is the set that includes all eigenvalues of Q and R. **Since we know Q and R are psd, then we know the eigenvalues are positive and A is also a psd.**

Problem 3

Find all stationary points of the following $\mathbf{R}^2$ functions and classify them (only local):

Here we will use the **Second Partial Derivative Test**; which states:

$$H = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2$$

- If $H > 0$, the function has a local maximum/ minimum at the point (x_0, y_0).
  - if $f_{xx}(x_0, y_0) > 0$ it is a minimum
  - if $f_{xx}(x_0, y_0) < 0$ it is a maximum
- If $H < 0$, the function has a saddle point at (x_0, y_0).
- If $H = 0$, there is not enough information to tell

1. $f(x, y) = 1 - y^2 - x^4$

   First; we must find the partial derivative f_x andf_y

   $$f_x = -4x^3 \quad \text{and} \quad f_y = -2y$$

   To find the stationary points, we need to find all points (x, y) for which $f_x = f_y = 0$

   $$\begin{aligned} -4x^3 &= 0 \\ x &= 0 \end{aligned} \qquad\qquad \begin{aligned} -2y &= 0 \\ y &= 0 \end{aligned}$$

   Therefore the stationary point only exists at $(x, y) = (0, 0)$ Now we can try identify the stationary point by finding the second derivative:

   $$f_{xx} = -12x^2 \quad \text{and} \quad f_{yy} = -2 \quad \text{and} \quad f_{xy} = 0$$

   Then using the Second Partial Derivative Test:

   $$\begin{aligned} H &= f_{xx} f_{yy} - f_{xy}^2 \\ &= -12(0)^2 \cdot -2 - (0)^2 \\ &= 0 \end{aligned}$$

   Because $H = 0$, we cannot identify the stationary point at $(0, 0)$.

2. $f(x, y) = x^4 + y^4 - 2x^2 + 4xy - 2y^2$

   First; we must find the partial derivative f_x andf_y

   $$f_x = 4x^3 - 4x + 4y \quad \text{and} \quad f_y = 4y^3 - 4y + 4x$$

To find the stationary points, we need to find all points (x, y) for which $f_x = f_y = 0$

$$\begin{aligned} 4x^3 - 4x + 4y &= 0 & 4y^3 - 4y + 4x &= 0 \quad (1) \\ 4x^3 &= 4x - 4y & -4y^3 &= 4x - 4y \quad (2) \\ 4x^3 &= -4y^3 & & \quad (3) \\ x &= -y & & \quad (4) \end{aligned}$$

We can solve Equation 1 by using Equation 4:

$$\begin{aligned} 0 &= 4x^3 - 4x + 4(-x) \\ &= 4x^3 - 8x \\ &= 4x(x^2 - 2) \end{aligned}$$

Therefore the stationary point exists at $(x, y) = (0, 0), (\sqrt{2}, -\sqrt{2}), (-\sqrt{2}, \sqrt{2})$ Now we can try identify the stationary point by finding the second derivative:

$$f_{xx} = 12x^2 - 4 \quad \text{and} \quad f_{yy} = 12y^2 - 4 \quad \text{and} \quad f_{xy} = 4$$

Then using the Second Partial Derivative Test for $(x, y) = (0, 0)$:

$$\begin{aligned} H &= f_{xx} f_{yy} - f_{xy}^2 \Big|_{(x,y)=(0,0)} \\ &= (12(0)^2 - 4) \cdot (12(0)^2 - 4) - (4)^2 \\ &= 16 - 16 \\ &= 0 \end{aligned}$$

Because $H = 0$, we cannot identify the stationary point at $(0, 0)$. Then using the Second Partial Derivative Test for $(x, y) = (\sqrt{2}, -\sqrt{2})$:

$$\begin{aligned} H &= f_{xx} f_{yy} - f_{xy}^2 \Big|_{(x,y)=(\sqrt{2},-\sqrt{2})} \\ &= (12(\sqrt{2})^2 - 4) \cdot (12(-\sqrt{2})^2 - 4) - (4)^2 \\ &= 384 \end{aligned}$$

Because $H > 0$, and $f_{xx} > 0$; we can conclude the stationary point is a local minimum at $(\sqrt{2}, -\sqrt{2})$. Then using the Second Partial Derivative Test for $(x, y) = (-\sqrt{2}, \sqrt{2})$:

$$\begin{aligned} H &= f_{xx} f_{yy} - f_{xy}^2 \Big|_{(x,y)=(-\sqrt{2},\sqrt{2})} \\ &= (12(-\sqrt{2})^2 - 4) \cdot (12(\sqrt{2})^2 - 4) - (4)^2 \\ &= 384 \end{aligned}$$

Because $H > 0$, and $f_{xx} > 0$; we can conclude the stationary point is a local minimum at $(-\sqrt{2}, \sqrt{2})$.

3. $f(x, y) = (ax^2 + by^2)e^{-x^2+y^2}$, where $a, b \in \mathbf{R}$

First; we must find the partial derivative f_x andf_y using *Chain Rule*: $\frac{d}{dx}\left(f\left(x\right)g\left(x\right)\right) = f\left(x\right)\frac{d}{dx}g\left(x\right) + \frac{d}{dx}f\left(x\right)g\left(x\right)$

$$f_x = (ax^2 + by^2)(-2e^{-x^2+y^2}) + (2ax)(e^{-x^2+y^2})$$

$$f_x = (e^{-x^2+y^2})(-2x(ax^2 + by^2)) + 2ax$$

$$f_x = (e^{-x^2+y^2})(-2x)(ax^2 + by^2 - a) \qquad f_y = (e^{-x^2+y^2})(2y)(ax^2 + by^2 + b)$$

To find the stationary points, we need to find all points (x, y) for which $f_x = f_y = 0$

$$\begin{aligned} f_x = (e^{-x^2+y^2})(-2x)(ax^2 + by^2 - a) = 0 \\ (-2x)(ax^2 + by^2 - a) = 0 \end{aligned}$$

$$x = 0 \qquad \text{and} \qquad ax^2 + by^2 - a = 0 \longrightarrow \qquad y = \sqrt{\frac{a}{b}}$$

$$\begin{aligned} f_y = (e^{-x^2+y^2})(2y)(ax^2 + by^2 + b) = 0 \\ (2y)(ax^2 + by^2 + b) = 0 \end{aligned}$$

$$y = 0 \qquad \text{and} \qquad ax^2 + by^2 + b = 0 \longrightarrow \qquad x = \sqrt{\frac{-b}{a}}$$

Therefore the stationary point exists at $(x, y) = (0, \sqrt{\frac{a}{b}}, (\sqrt{\frac{-b}{a}}, 0)$

$$\begin{aligned} f_{xx} &= (-2xe^{-x^2+y^2})(-2ax^3 - 2bxy^2 + 2ax) + (e^{-x^2+y^2})(-6ax^2 - 2by^2 + 2a) \\ f_{yy} &= (-2ye^{-x^2+y^2})(2ax^2y + 2by^3 + 2by) + (e^{-x^2+y^2})(2ax^2 + 6by + 2b) \\ f_{xy} &= (-2ye^{-x^2+y^2})(-2ax^3 - 2bxy^2 + 2ax) + (e^{-x^2+y^2}) \end{aligned}$$

By following the same procedures from above, we can use H to solve for a and b to set the criteria for the function to have a minimum, maximum or saddle point.

Problem 4

Let h be a real function with continuous and positive second derivative such that $h'(0) = 0$. Let the following function be defined as $f(x, y) := h(x + y)$

1. Determine the stationary points of f.

   We know $h'(0) = 0$ and $f(x, y) = h(x + y)$:

   $$\begin{aligned} f_x = h'(x + y)\frac{\partial f}{\partial x} \\ x + y = 0 \end{aligned}$$

   Therefore the stationary points are $\{(x, y) \in \mathbf{R}^2 \mid x + y = 0\}$

2. Classify them (only local). Hint: note h has a minimum at $x = 0$.

   We know $h''(x + y) > 0$

   $$f_{xx} = h''(x + y)\frac{\partial^2 f}{\partial x^2} \geq 0$$

   Let the function h denote the even power operator, as all criteria: (1) Continuous Positive Second Derivative and (2) $h'(0) = 0$ hold. Therefore, the second partial derivative test would give us a positive values. And as shown above, **there would be a minimum on the line x=-y**

Problem 5

Let $b \neq 0 \in \mathbf{R}^n$. Show that the maximum of $f(x) = b^T x$ over $X = \{x \in \mathbf{R}^n : \|x\| \leq 1\}$ is attained at $x^* = \frac{b}{\|b\|}$ and $f(x^*) = \|b\|$. Hint: Use Cauchy-Shwartz inequality to bound f.

Here we will use the **Cauchy-Schwarz Inequality**; which states that for all all $\vec{x}$ and $\vec{y}$ of an inner product space it is true that

$$|\vec{x} \bullet \vec{y}| \leq \|\vec{x}\|\|\vec{y}\|.$$

We know that $|f(x)| = |b^T x| = |\langle bx \rangle| \implies |\langle bx \rangle| = \|b\| \cdot \|x\|$. So by using Caushy-Schwarz Inequality we can bound f

$$|f(x)| \leq \|b\| \cdot \|x\|$$

Given that $\|x\| \leq 1$ and $b \neq 0 \in \mathbf{R}^n$

$$|f(x)| \leq \|b\|$$

Let $x^* = \frac{b}{\|b\|}$, then:

$$|f(x)| \leq \|b\| = b^T(\frac{b}{\|b\|}) = \frac{b^T b}{\|b\|} = \frac{\|b\|^2}{\|b\|} = \|b\|$$

Because we proved that f is bounded by $\|b\|$ above and must be less than $\|b\|$. Therefore, using this reasoning we can conclude that the maximum is attained at $x^* = \frac{b}{\|b\|}$ and $f(x^*) = \|b\|$

Problem 6

Show if the following functions are coercive or not.

A function is **coercive** if it grows rapidly at the extremes of the space on which it is defined. *Completing the Square* was used to simplify the function below. A function f is coercive if

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

1. $f(x, y) = 4x^2 + 2xy + 2y^2$.

$$\begin{aligned} f(x, y) &= 4x^2 + 2xy + 2y^2 \\ &= (x + y)^2 + 3x^2 + y^2 \end{aligned}$$

$$\begin{aligned} \lim_{\|x\| \to \infty} f(x) &= \lim_{\|x\| \to \infty} (x + y)^2 + 3x^2 + y^2 \\ &= \infty \end{aligned}$$

Because $\lim_{\|x\| \to \infty} f(x) = \infty$, this function is coercive

2. $f(x, y) = 2x^2 - 8xy + y^2$

$$\begin{aligned} f(x, y) &= 2x^2 - 8xy + y^2 \\ &= 2(x^2 - 4xy = 4y^2 - 4y^2) + y^2 \\ &= 2(x - 2y)^2 - 7y^2 \end{aligned}$$

$$\begin{aligned} \lim_{\|x\| \to \infty} f(x) &= \lim_{\|x\| \to \infty} 2(x - 2y)^2 - 7y^2 \\ &= \infty \end{aligned}$$

Because $\lim_{\|x\| \to \infty} f(x) = \infty$, this function is coercive

3. $f(x, y, z) = x^3 + y^3 + z^3$

$$f(x, y, z) = x^3 + y^3 + z^3$$

$$\begin{aligned} \lim_{\|x\| \to \infty} f(x) &= \lim_{\|x\| \to \infty} x^3 + y^3 + z^3 \\ &= \infty \end{aligned}$$

$$\begin{aligned} \lim_{\|x\| \to -\infty} f(x) &= \lim_{\|x\| \to -\infty} x^3 + y^3 + z^3 \\ &= -\infty \end{aligned}$$

Because $\lim_{\|x\| \to -\infty} f(x) = -\infty$, this function is not coercive

4. $f(x,y) = x^2 - 2xy^2 + y^4$

$$\begin{aligned} f(x,y) &= x^2 - 2xy^2 + y^4 \\ &= (x-y^2)^2 \end{aligned}$$

$$\begin{aligned} \lim_{\|x\|\to\infty} f(x) &= \lim_{\|x\|\to\infty} (x-y^2)^2 \\ &= \infty \end{aligned}$$

Because $\lim_{\|x\|\to\infty} f(x) = \infty$, this function is coercive

Problem 7

Find all stationary points and classify them (local and global).

1. $f(x,y) = (4x^2 - y)^2$

   First we must simplify the equation above.

$$\begin{aligned} f_x = (4x^2-y)^2 &= (4x^2-y)(4x^2-y) \\ &= 16x^4 - 8x^2y + y^2 \end{aligned}$$

   Then; we must find the partial derivative f_x andf_y and to find the stationary points, we need to find all points (x,y) for which $f_x = f_y = 0$

$$64x^3 - 16xy = f_x \qquad\qquad 2y - 8x^2 = f_y$$
$$16x(4x-4) = 0 \qquad\qquad 2(y-4x^2)$$

   Therefore the stationary point only exists at $(x,y) = (0,0)$ Now we can try identify the stationary point by finding the second derivative:

$$f_{xx} = 192x^2 - 16y \quad \text{and} \quad f_{yy} = 2 \quad \text{and} \quad f_{xy} = -16x$$

   Then using the Second Partial Derivative Test:

$$\begin{aligned} H &= f_{xx}f_{yy} - f_{xy}^2 \\ &= [192(0)^2 - 16(0)] \cdot 2 - (-16(0))^2 \\ &= 0 \end{aligned}$$

   Because $H = 0$, we cannot identify the stationary point at $(0,0)$.

2. $f(x,y) = 2x^2 + 3y^2 - 2xy + 2x - 3y$.

   First we must find the partial derivative f_x andf_y and to find the stationary points, we need to find all points (x,y) for which $f_x = f_y = 0$

$$4x - 2y + 2 = f_x \qquad\qquad 4x - 2y + 2 = f_y$$
$$4x - 2y = -2 \qquad\qquad 4x - 2y = 2$$

   Therefore the stationary point only exists at $(x,y) = (\frac{-3}{10}, \frac{4}{10})$ Now we can try identify the stationary point by finding the second derivative:

$$f_{xx} = 4 \quad \text{and} \quad f_{yy} = 6 \quad \text{and} \quad f_{xy} = -2$$

   Then using the Second Partial Derivative Test:

$$\begin{aligned} H &= f_{xx}f_{yy} - f_{xy}^2 \\ &= (4)(6) - (2)^2 \\ &= 0 \end{aligned}$$

   Because $H > 0$, and $f_{xx} > 0$; we can conclude the stationary point is a global minimum at $(\frac{-3}{10}, \frac{4}{10})$.

3. $f(x, y, z) = x^4 - 2x^2 + y^2 + 2yz + 2z^2$.

   First we must find the partial derivative f_x andf_y and to find the stationary points, we need to find all points (x, y) for which $f_x = f_y = 0$

$$\begin{aligned} 4x^3 - 4x &= f_x & 2y + 2z &= f_y & 4z &= f_z \\ x(x^2 - 1) &= 0 & 2(y + z) &= 0 & z &= 0 \end{aligned}$$

   Therefore the stationary point exists at $(x, y, z) = (1, 0, 0), (0, 0, 0), (-1, 0, 0)$. By using the Second Partial Derivative Test we can also conclude that **there is a local minimum at** $(1, 0, 0)$ **and** $(-1, 0, 0)$**, with a saddle point at** $(0, 0, 0)$**.**