**Manish Kumar (21044)**
**Sep 09, 2023**

**E0 230: Computational Methods of Optimization**
**Assignment 01**

**Q2. Critical point analysis of** $p(x,y) = x^4y^2 + x^2y^4 - 9x^2y^2$

# solution

Given, $p(x,y) = x^4y^2 + x^2y^4 - 9x^2y^2$

To get critical point, we use the relation, $\nabla p(x,y) = 0$

$$\nabla p(x,y) = \begin{bmatrix} 4x^3y^2 + 2xy^4 - 18xy^2 \\ 2x^2y4 + 4x^2y^3 - 18x^2y \end{bmatrix}$$

We will find simultaneous solution of $4x^3y^2 + 2xy^4 - 18xy^2 = 0$ and $2x^2y^4 + 4x^2y^3 - 18x^2y = 0$,

They simplify as,
$4x^3y^2 + 2xy^4 - 18xy^2 = 2xy^2(-9 + 2x^2 + y^2) = 0$, and
$2x^2y^4 + 4x^2y^3 - 18x^2y = 2x^2y(-9 + x^2 + 2y^2) = 0$

The simultenous solution of the above equations: $(x,y) = (0,c),\ (c,0),\ (\sqrt{3},\sqrt{3}),\ (-\sqrt{3},\sqrt{3}),\ (\sqrt{3},-\sqrt{3}),\ (-\sqrt{3},-\sqrt{3})$

To find minima among the above critical points, we calculate Hessian of $p(x,y)$:
$H[p(x,y)] = \begin{bmatrix} 12x^2y^2 + 2y^4 - 18y^2 & 8yx^3 + 8y^3x - 36yx \\ 8yx^3 + 8y^3x - 36yx & 2x^4 + 12y^2x^2 - 18x^2 \end{bmatrix}$

Value of Hessian at $x = [0,0]^T \implies H[p(0,0)] = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

• Hence, $x = [0,0]^T$ is a stationary point but not a (global) minima.

Value of Hessian at $(\sqrt{3},\sqrt{3}) \implies H[p(\sqrt{3},\sqrt{3})] = \begin{bmatrix} 72 & -18 \\ -18 & 72 \end{bmatrix}$

Its eigenvalue is 54 and 90 (see the Python script for eigenvalue estimation). Hence $H[p(\sqrt{3},\sqrt{3})]$ is psd.
$\implies$ it is local maxima.

Value of Hessian at $(-\sqrt{3},\sqrt{3}) \implies H[p(-\sqrt{3},\sqrt{3})] = \begin{bmatrix} 72 & 18 \\ 18 & 72 \end{bmatrix}$

Its eigenvalue is 54 and 90 (see the Python script for eigenvalue estimation). Hence $H[p(\sqrt{3},\sqrt{3})]$ is psd.
$\implies$ it is local maxima.

Value of Hessian at $(\sqrt{3},-\sqrt{3}) \implies H[(\sqrt{3},-\sqrt{3})] = \begin{bmatrix} 72 & 18 \\ 18 & 72 \end{bmatrix}$

Its eigenvalue is 54 and 90 (see the Python script for eigenvalue estimation). Hence $H[p(\sqrt{3},\sqrt{3})]$ is psd.
$\implies$ it is local maxima.

Value of Hessian at $(-\sqrt{3},-\sqrt{3}) \implies H[p(-\sqrt{3},-\sqrt{3})] = \begin{bmatrix} 72 & -18 \\ -18 & 72 \end{bmatrix}$

Its eigenvalue is 54 and 90 (see the Python script for eigenvalue estimation). Hence $H[p(\sqrt{3},\sqrt{3})]$ is psd.
$\implies$ it is local maxima.

**Conclusion**: • Hence, $x = [0,0]^T$ is a stationary point but not a (global) minima.

• Global minima is not unique. There are four global minima. Because, $p(\sqrt{3},\sqrt{3}) = p(-\sqrt{3},\sqrt{3}) = p(\sqrt{3},-\sqrt{3}) = p(-\sqrt{3},-\sqrt{3})$

**Q3. Estimate bound on eigenvalues**

## solution:

Given, $f(x) = e^{x^T A x} \frac{e^{-x^T(B+C)x}}{1+e^{-x^T(C-B)x}}$

Equivalently, $f(x) = \frac{e^{(x^T Ax)-(x^T(B+C)x)}}{1+e^{-x^T(C-B)x}} = \frac{e^{x^T(A-B-C)x}}{1+e^{-x^T(C-B)x}}$

Coercivity requires: $\lim_{(x^T x)\to\infty} f(x) \to \infty$.

Using the below inequalities to get upper bound on the function-

$\frac{e^{x^T(A-B-C)x}}{1+e^{-x^T(C-B)x}} \leq \frac{e^{x^T(A-B-C)x}}{e^{-x^T(C-B)x}} = e^{x^T(A-2B)x}$

Now we need to find the condition on matrix A and B such that:

$\lim_{(x^T x)\to\infty} e^{x^T(A-2B)x} \to \infty$.

Since, $\lim_{(t)\to\infty} e^{f(t)} \to \infty$, if $\lim_{(t)\to\infty} f(t) \to \infty$. Hence, we need $\lim_{(x^T x)\to\infty} x^T(A-2B)x \to \infty$.

This is possible if $A - 2B$ is positive-definite.

The sum of two positive definite matrices is positive definite.

**Q4. Least Square fit for Linera function**

# solution:

- Part-I: estimated $w \in \mathbb{R}^5 = [-0.20, -0.42, -0.42, -0.09, -0.55]$

- Part-II: Closed-form solution for m data points and $x \in \mathbb{R}^n$:

Error/cost function to be minimized over $w \in \mathbb{R}^n,\ b \in \mathbb{R}$ is:-

$f(w, b) = \frac{1}{m} \sum_{i=0}^{m} (w^T x_i + b - y_i)^2 = \frac{1}{m} \sum_{i=0}^{m} (\sum_{j=0}^{n} w_j x_{ij} + b - y_i)^2$

To minimize the cost function, we take partial derivatives as:

$\frac{\partial f}{\partial w_l} = \frac{1}{2m} \sum_{i=0}^{m} x_{li} (\sum_{j=0}^{n} w_j x_{ij} + b - y_i)$

$\frac{\partial f}{\partial b} = \frac{1}{2m} \sum_{i=0}^{m} (\sum_{j=0}^{n} w_j x_{ij} + b - y_i)$

Using, $\frac{\partial f}{\partial w_l} = \frac{\partial f}{\partial b} = 0$ and simplifying calculation yields,

$\frac{\partial f}{\partial w_l} = \frac{1}{2m} \sum_{i=0}^{m} x_{li} (\sum_{j=0}^{n} w_j x_{ji} + b - y_i) = 0$

$\frac{\partial f}{\partial w_l} = \frac{1}{2m} (\sum_{j=0}^{n} (\sum_{i=0}^{m} x_{li} x_{ji}) w_j) - \frac{1}{2m} (\sum_{1}^{m} x_{li} y_i) = 0$

Solving it is equivalent to solving the m-linear equation in n-variables.

$\boxed{\text{or, taking } \sum_{i=1}^{m} x_{li} x_{ji} = A_{lj}, \text{ and } \sum_{i=1}^{m} x_{li} y_i = c_l \implies \sum_{j=1}^{n} A_{lj} w_j - c_l; \forall l \in \{1, ..., n\}}$

We can solve the above system of equations using a linear equation solver in Python.

$\boxed{w = (A^T A)^{-1} A^T c.}$

**Conclusion-1**: Generally, a system with the same number of equations and unknowns has a single unique solution.

- If the number of data points 'm' is less than the number of variable 'n'

**Conclusion-2**: In this case $m > n$, also called undetermined system of equation. In general, a system with fewer equations than unknowns has infinitely many solutions, but it may have no solution, too.

Way to solve under-determines system: We can use the Moore-Penrose inverse ($A^+$) technique. It generalizes the idea of matrix inverse for non-full-rank matrixes, which is the current case.