Pre-class Assignment-23

Reading

① Quadratic Optimization

A quadratic optimization problem is an optimization problem of the form:

minimize $f(x) = \frac{1}{2} x^T Q x + c^T x$; $x \in \mathbb{R}^n$

here, x is the vector of variables we are optimizing
Q is the Symmetric +ve defintie matrix
c is a vector

The objective is to minimize the quadratic form $f(x)$.

② * Quadratic optimization is advantageous due to its computational efficiency, often having efficient algorithm for optimal solutions.

* Analytical solutions are often attainable, providing clear expressions for optimal outcomes.
* The convex nature of quadratic function ensures a Straight forward optimization process with a unique global minimum.

* Widely applicable across various domains, quadratic optimization finds use in finance, engineering, statistics & machine learning, making it an ideal choice for diverse problem solving scenarios.

③ The method of least squares is a technique used to approximate the solution of an overdetermined system of linear equations by minimizing the sum of squares of the residuals. The least squares soln. is obtained by solving a quadratic optimization problem.

given an overdetermined system $Ax = b$

where, A is a matrix with more rows than columns

the least squares soln. $\hat{x}$ is obtained by minimizing the following quadratic objective function:

$$f(x) = \frac{1}{2} \underbrace{\|Ax - b\|}_{\text{Euclidean Norm}}^2$$

The soln. $\hat{x}$ minimizes the sum of squared diff b/w the observed values 'b' & values predicted by linear model 'Ax'.

Exercises

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

i) The gradient of f is given by vector of partial derivatives.

$$\nabla f(x,y) = \begin{bmatrix} \partial f/\partial x \\ \partial f/\partial y \end{bmatrix}$$

$$\frac{\partial f}{\partial x} = 12x^2 - 3y^2 - 1 \qquad \frac{\partial f}{\partial y} = -6xy + 2y$$

At origin $(0,0)$: $\frac{\partial f}{\partial x} = 0 - 0 - 1 = -1$

$$\frac{\partial f}{\partial y} = 0$$

hence, $\nabla f(x,y) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$

Hessian matrix (H_f) is given by:

$$H_f = \begin{bmatrix} \partial^2 f/\partial x^2 & \partial^2 f/\partial x \partial y \\ \partial^2 f/\partial y \partial x & \partial^2 f/\partial y^2 \end{bmatrix}$$

$$\frac{\partial^2 f}{\partial x^2} = 24x \qquad \frac{\partial^2 f}{\partial y^2} = -6x+2$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -6y$$

At origin $(0,0)$:

$$\frac{\partial^2 f}{\partial x^2} = 24(0) = 0 \qquad \frac{\partial^2 f}{\partial y^2} = -6(0)+2 = 2$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -6(0) = 0$$

therefore, the Hessian Matrix (H_f) is given by:

$$\Rightarrow H_f = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

b) Quadratic approximation of 'f' near the origin. i.e.,

$$f(x) \cong f(0) + \nabla f(0)^T \cdot x + \frac{1}{2} x^T H_f(0)\, x.$$

from a) part we know, $\nabla f(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$

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

$$f(0) = f(0,0) = 4(0) - 3(0) + 0 - 0 + 3 = 3$$

$$f(x,y) \approx 3 + \begin{bmatrix} -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$f(x,y) \approx 3 - x + \frac{1}{2} \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$f(x,y) \approx 3 - x + \frac{1}{2} \begin{bmatrix} 0 & 2y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$f(x,y) \approx 3 - x + \frac{1}{2}(2y^2)$$

$$\Rightarrow \boxed{f(x,y) \approx y^2 - x + 3}$$

c) The quadratic optimization problem involves minimizing the quadratic form:

$$q(x) = \frac{1}{2} x^T H x + \nabla f^T x$$

here, $H = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$ & $\nabla f = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$

Minimize $q(x) = \frac{1}{2} \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

on Simplifying,

$$q(x) = \frac{1}{2} \begin{bmatrix} 0 \cdot 2y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - x$$

$$q(x) = \frac{1}{2}\left[2y^2\right] - x$$

$$q(x) = y^2 - x$$

$$\nabla q(x) = \begin{bmatrix} \partial q/\partial x \\ \partial q/\partial y \end{bmatrix} = \begin{bmatrix} -1 \\ 2y \end{bmatrix}$$

when $\nabla q(x) = \begin{bmatrix} -1 \\ 2y \end{bmatrix} = 0$

$y = 0 \Rightarrow$ substitute in $f(x, y)$

hence, $q(x) = -x$

$$f(x,y) = 4x^3 - 3xy^2 + y^2 - x + 3$$

$$f(x,0) = 4x^3 - x + 3$$

At $f(0,0) = 0 - 0 + 3$

$= 3$

Hence, the minimum value is 3 at $(0,0)$

(2) Here is an example of a matrix that has non-negative eigen values but its not positive semidefinite:

Consider the matrix Q

$$Q = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$$

the eigen values of Q are $\lambda_1 = 1$ & $\lambda_2 = 2$
both are non-negative

However, Q is not +ve semidefinite because there exists a vector $x = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$ for which $x^T Q x < 0$

Specifically:

$$x^T Q x = [2 \ -3] \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \end{bmatrix} = 4 - 9 = -5 < 0$$

Even though the eigen values of Q are non-negative, the matrix itself does not satisfy the +ve semidefinte condition for all vector 'x'. Requirement for positive semidefiniteness is $x^T Q x \geq 0$ for all vectors 'x'.

(3) Let 'Q' be a real Symmetric matrix.

Consider the optimization problem

Maximize $f(x) = x^T Q(x)$

given, $\|x\| = 1$; where $\|x\| = \sqrt{x^T x}$ denotes euclidean norm of 'x'.

'Q' posses 'n' orthogonal eigen vectors $v_1, v_2, \ldots, v_n$ with corresponding real eigen values $\lambda_1, \lambda_2, \ldots, \lambda_n$

Let $\lambda_{max} = \max_i \lambda_i$ be the max eigen value of 'Q'
with corresponding eigen vector V_{max}

Now, express Q in terms of eigen decomposition

$$x^T Q x = x^T \left(\sum_{i=1}^{n} \lambda_i v_i v_i^T \right) x$$

$$= \sum_{i=1}^{n} \lambda_i (x^T v_i)^2$$

Applying Cauchy-schwarz inequality.

$$x^T Q x \leq \lambda_{max} \sum_{i=1}^{n} |\langle x, v_i \rangle|^2$$

using pythagoras theorem.

$$x^T Q x = \lambda_{max} ||x||^2$$

$$= \lambda_{max}$$

Equality is achieved when $x = V_{max}$.

Therefore, the max value of $f(x)$ subject to $||x|| = 1$ is exactly λ_{max}, the max eigenvalue of Q.