Optimization using Calculus

Optimization of Functions of Multiple Variables: Unconstrained Optimization

D Nagesh Kumar, IISc

Objectives

- To study functions of multiple variables, which are more difficult to analyze owing to the difficulty in graphical representation and tedious calculations involved in mathematical analysis for unconstrained optimization.
- To study the above with the aid of the gradient vector and the Hessian matrix.
- To discuss the implementation of the technique through examples

Unconstrained optimization

- If a convex function is to be minimized, the stationary point is the global minimum and analysis is relatively straightforward as discussed earlier.
- A similar situation exists for maximizing a concave variable function.
- The necessary and sufficient conditions for the optimization of unconstrained function of several variables are discussed.

Necessary condition

- In case of multivariable functions a necessary condition for a stationary point of the function *f*(**X**) is that each partial derivative is equal to zero. In other words, each element of the gradient vector $\Delta_x f$ defined below must be equal to zero. i.e. the gradient vector of *f*(**X**), at **X=X***, defined as follows, must be equal to zero:

$$\Delta_x f = \begin{bmatrix} \frac{\partial f}{\partial x_1}(X^*) \\ \frac{\partial f}{\partial x_2}(X^*) \\ \vdots \\ \vdots \\ \frac{\partial f}{\partial dx_n}(X^*) \end{bmatrix} = 0$$

Sufficient condition

- For a stationary point **X*** to be an extreme point, the matrix of second partial derivatives (Hessian matrix) of *f*(**X**) evaluated at **X*** must be:
  - positive definite when **X*** is a point of relative minimum, and
  - negative definite when **X*** is a relative maximum point.

- When all eigen values are negative for all possible values of **X**, then **X*** is a global maximum, and when all eigen values are positive for all possible values of **X**, then **X*** is a global minimum.
- If some of the eigen values of the Hessian at **X*** are positive and some negative, or if some are zero, the stationary point, **X***, is neither a local maximum nor a local minimum.

Example

Analyze the function $f(x) = -x_1^2 - x_2^2 - x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_1 - 5x_3 + 2$ and classify the stationary points as maxima, minima and points of inflection

Solution

$$\Delta_x f = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathrm{X}^*) \\ \frac{\partial f}{\partial x_2}(\mathrm{X}^*) \\ \frac{\partial f}{\partial x_3}(\mathrm{X}^*) \end{bmatrix} = \begin{bmatrix} -2x_1 + 2x_2 + 2x_3 + 4 \\ -2x_2 + 2x_1 \\ -2x_3 + 2x_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Example …contd.

Solving these simultaneous equations we get **X***=[1/2, 1/2, -2]

$$\frac{\partial^2 f}{\partial x_1^{\;2}} = -2; \frac{\partial^2 f}{\partial x_2^{\;2}} = -2; \frac{\partial^2 f}{\partial x_3^{\;2}} = -2$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = 2$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_3} = \frac{\partial^2 f}{\partial x_3 \partial x_2} = 0$$

$$\frac{\partial^2 f}{\partial x_3 \partial x_1} = \frac{\partial^2 f}{\partial x_1 \partial x_3} = 2$$

Example …contd.

Hessian of $f(\mathbf{X})$ is

$$\mathbf{H} = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right]$$

$$\mathbf{H} = \begin{bmatrix} -2 & 2 & 2 \\ 2 & -2 & 0 \\ 2 & 0 & -2 \end{bmatrix}$$

$$\left| \lambda \mathbf{I} - \mathbf{H} \right| = \begin{vmatrix} \lambda + 2 & -2 & -2 \\ -2 & \lambda + 2 & 0 \\ -2 & 0 & \lambda + 2 \end{vmatrix} = 0$$

Example …contd.

or $$(\lambda+2)(\lambda+2)(\lambda+2)-2(\lambda+2)(2)+2(2)(\lambda+2)=0$$

$$(\lambda+2)[\lambda^2+4\lambda+4-4+4]=0$$

$$(\lambda+2)^3=0$$

or $$\lambda_1=\lambda_2=\lambda_3=-2$$

Since all eigenvalues are negative the function attains a maximum at the point $\mathbf{X}^*$=[1/2, 1/2, -2]

Thank you