Module – 8 Lecture Notes – 4

Direct and Indirect Search Methods

Introduction

Most of the real world system models involve nonlinear optimization with complicated objective functions or constraints for which analytical solutions (solutions using quadratic programming, geometric programming, etc.) are not available. In such cases one of the possible solutions is the search algorithm in which, the objective function is first computed with a trial solution and then the solution is sequentially improved based on the corresponding objective function value till convergence. A generalized flowchart of the search algorithm in solving a nonlinear optimization with decision variable $\mathbf{X_i}$**,** is presented in Fig. 1.

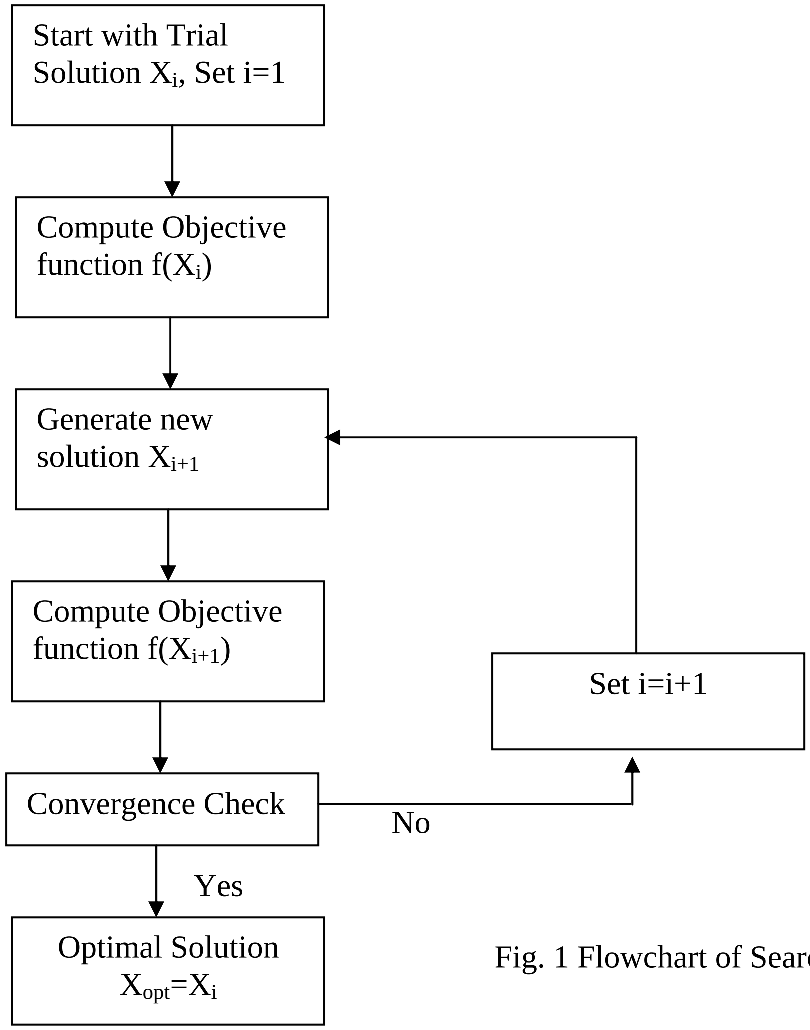

Fig. 1 Flowchart of Search Algorithm

The search algorithms can be broadly classified into two types: (1) direct search algorithm and (2) indirect search algorithm. A direct search algorithm for numerical search optimization depends on the objective function only through ranking a countable set of function values. It does not involve the partial derivatives of the function and hence it is also called nongradient or zeroth order method. Indirect search algorithm, also called the descent method, depends on the first (first-order methods) and often second derivatives (second-order methods) of the objective function. A brief overview of the direct search algorithm is presented.

Direct Search Algorithm

Some of the direct search algorithms for solving nonlinear optimization, which requires objective functions, are described below:

A) Random Search Method: This method generates trial solutions for the optimization model using random number generators for the decision variables. Random search method includes random jump method, random walk method and random walk method with direction exploitation. Random jump method generates huge number of data points for the decision variable assuming a uniform distribution for them and finds out the best solution by comparing the corresponding objective function values. Random walk method generates trial solution with sequential improvements which is governed by a scalar step length and a unit random vector. The random walk method with direct exploitation is an improved version of random walk method, in which, first the successful direction of generating trial solutions is found out and then maximum possible steps are taken along this successful direction.

B) Grid Search Method: This methodology involves setting up of grids in the decision space and evaluating the values of the objective function at each grid point. The point which corresponds to the best value of the objective function is considered to be the optimum solution. A major drawback of this methodology is that the number of grid points increases exponentially with the number of decision variables, which makes the method computationally costlier.

C) Univariate Method: This procedure involves generation of trial solutions for one decision variable at a time, keeping all the others fixed. Thus the best solution for a decision variable keeping others constant can be obtained. After completion of the process with all the decision variables, the algorithm is repeated till convergence.

D) Pattern Directions: In univariate method the search direction is along the direction of co-ordinate axis which makes the rate of convergence very slow. To overcome this drawback, the method of pattern direction is used, in which, the search is performed not along the direction of the co-ordinate axes but along the direction towards the best solution. This can be achieved with Hooke and Jeeves' method or Powell's method. In the Hooke and Jeeves' method, a sequential technique is used consisting of two moves: exploratory move and the pattern move. Exploratory move is used to explore the local behavior of the objective function, and the pattern move is used to take advantage of the pattern direction. Powell's method is a direct search method with conjugate gradient, which minimizes the quadratic function in a finite number of steps. Since a general nonlinear function can be approximated reasonably well by a quadratic function, conjugate gradient minimizes the computational time to convergence.

E)Rosen Brock's Method of Rotating Coordinates: This is a modified version of Hooke and Jeeves' method, in which, the coordinate system is rotated in such a way that the first axis always orients to the locally estimated direction of the best solution and all the axes are made mutually orthogonal and normal to the first one.

F) Simplex Method: Simplex method is a conventional direct search algorithm where the best solution lies on the vertices of a geometric figure in N-dimensional space made of a set of N+1 points. The method compares the objective function values at the N+1 vertices and moves towards the optimum point iteratively. The movement of the simplex algorithm is achieved by reflection, contraction and expansion.

Indirect Search Algorithm

The indirect search algorithms are based on the derivatives or gradients of the objective function. The gradient of a function in N-dimensional space is given by:

$$\nabla f = \begin{bmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \cdot \\ \cdot \\ \cdot \\ \partial f / \partial x_N \end{bmatrix} \tag{1}$$

Indirect search algorithms include:

A) Steepest Descent (Cauchy) Method: In this method, the search starts from an initial trial point X1, and iteratively moves along the steepest descent directions until the optimum point is found. Although, the method is straightforward, it is not applicable to the problems having multiple local optima. In such cases the solution may get stuck at local optimum points.

B) Conjugate Gradient (Fletcher-Reeves) Method: The convergence technique of the steepest descent method can be greatly improved by using the concept of conjugate gradient with the use of the property of quadratic convergence.

C) Newton's Method: Newton's method is a very popular method which is based on Taylor's series expansion. The Taylor's series expansion of a function $f(\boldsymbol{X})$ at $\boldsymbol{X}=\boldsymbol{X_i}$ is given by:

$$f(X) = f(X_i) + \nabla f_i^T (X - X_i) + \frac{1}{2}(X - X_i)^T [J_i](X - X_i) \tag{2}$$

where, $[J_i]=[J]|x_i$, is the Hessian matrix of f evaluated at the point $\mathbf{X_i}$. Setting the partial derivatives of Eq. (2), to zero, the minimum value of $f(\boldsymbol{X})$ can be obtained.

$$\frac{\partial f(X)}{\partial x_j} = 0, \qquad j = 1,2,...,N \tag{3}$$

From Eq. (2) and (3)

$$\nabla f = \nabla f_i + [J_i](X - X_i) = 0 \tag{4}$$

Eq. (4) can be solved to obtain an improved solution $\boldsymbol{X}_{i+1}$

$$X_{i+1} = X_i - [J_i]^{-1} \nabla f_i \tag{5}$$

The procedure is repeated till convergence for finding out the optimal solution.

D) Marquardt Method: Marquardt method is a combination method of both the steepest descent algorithm and Newton’s method, which has the advantages of both the methods, movement of function value towards optimum point and fast convergence rate. By modifying the diagonal elements of the Hessian matrix iteratively, the optimum solution is obtained in this method.

E) Quasi-Newton Method: Quasi-Newton methods are well-known algorithms for finding maxima and minima of nonlinear functions. They are based on Newton's method, but they approximate the Hessian matrix, or its inverse, in order to reduce the amount of computation per iteration. The Hessian matrix is updated using the secant equation, a generalization of the secant method for multidimensional problems.

It should be noted that the above mentioned algorithms can be used for solving only unconstrained optimization. For solving constrained optimization, a common procedure is the use of a penalty function to convert the constrained optimization problem into an unconstrained optimization problem. Let us assume that for a point $\mathbf{X_i}$, the amount of violation of a constraint is δ. In such cases the objective function is given by:

$$f(X_i) = f(X_i) + \lambda \times M \times \delta^2 \tag{6}$$

where, λ=1(for minimization problem) and -1 (for maximization problem), M=dummy variable with a very high value. The penalty function automatically makes the solution inferior where there is a violation of constraint.

Summary

Various methods for direct and indirect search algorithms are discussed briefly in the present class. The models are useful when no analytical solution is available for an optimization problem. It should be noted that when there is availability of an analytical solution, the search algorithms should not be used, because analytical solution gives a global optima whereas, there is always a possibility that the numerical solution may get stuck at local optima.