MINIMIZATION OF THE PERFORMANCE FUNCTION

For minimizing the performance function, the conjugate gradient method [85] is employed. The conjugate gradient method is a modification of the steepest descent method [84]. Hence the steepest descent method is first discussed below.

The Method of Steepest Descent

The gradient vector $\nabla f$ represents the direction of greatest increase in the function $f$. Hence the negative of the gradient vector represents the direction of greatest decrease, also called the direction of steepest descent.

In the steepest descent method for function minimization, a trial point $X_i$ is chosen and moved iteratively towards the optimum point according to the rule

$$X_{i+1} = X_i + \lambda^*_i S_i, \quad i = 1, 2, 3 \ldots (B.1)$$

where $\lambda^*_i$ is the optimum step length in the search direction $S_i$. In the case of the steepest descent method
\[ S = -\nabla f \]  \hspace{1cm} \text{(B.2)}

where \( i \) stands for the iteration number.

The two important things in the minimization procedure are thus:

i) Calculating the gradient vector \( \nabla f \).

ii) Finding the optimal step length \( \lambda^* \) in the search direction \( S \), i.e., in the direction of \( -\nabla f \).

Thus the method of steepest descent reduces the search for the minimum of the performance function \( f \), to a series of searches for the minimum along a given direction \( S \). This search is one-dimensional since only the step size \( \lambda \), for which \( f(X + \lambda S) \) is minimized, must be found.

The convergence characteristic of the steepest descent method can be greatly improved by modifying it into a conjugate gradient method.

Conjugate Gradient Method (Fletcher-Reeves Method)

The conjugate gradient method is a quadratically convergent method, meaning that for quadratic function the minimum is found in \( N \) steps or less, where \( N \) is the number of adjustable variables of the function. For general functions, as the function approaches the minimum, the function becomes more closely quadratic and so convergence is more nearly assured.
In the conjugate gradient method each new search direction is taken as a linear combination of the previous search directions and the newly determined gradient. Thus the method takes into account the accumulated knowledge of the local behaviour of the function and hence is more powerful than the steepest descent method which merely uses the negative of the newly determined gradient as the search direction.

The function \( f(X) \) may be expanded in the neighbourhood of the required minimum \( h \) as:

\[
 f(X) = f(h) + \frac{1}{2} (X-h)^t A(X-h) + \ldots \quad \text{(B.3)}
\]

where \( A \) is the matrix of second order partial derivative and is symmetric and positive definite.

Now let us consider the minimization of the quadratic,

\[
 f(X) = f(h) + \frac{1}{2} (X-h)^t A(X-h) \quad \text{(B.4)}
\]

The search directions \( S_1, S_2, \ldots, S_l \) are chosen so that they are mutually conjugate with respect to \( A \), i.e. they satisfy the relation

\[
 S_i^t A S_j = 0 \quad \text{for} \quad i \neq j \quad \text{(B.5)}
\]

These requirements are satisfied if the search directions are set up as given below

\[
 S_i = -\nabla f_i \quad \text{(B.6)}
\]

\[
 S_i = -\nabla f_i + \beta S_{i-1}, \quad i = 2, 3, \ldots \quad \text{(B.7)}
\]
where

$$\beta_i = \frac{\nabla f_i^T \nabla f_i}{\nabla f_{i-1}^T \nabla f_{i-1}}$$ ... (B.8)

**Linear Search**

After finding the search direction \( S_i \), the next step in the minimization process is the determination of the optimal step length \( \lambda^*_i \) along the search direction. In other words, it is required to find \( \lambda^*_i \) such that the function \( f(X_i + \lambda_i S_i) \) attains a local minimum. The method adopted in the present case is the one based on cubic interpolation.

The method consists of three stages.

In the first stage the initial step size is estimated as:

$$h = k \text{ if } 0 < k < (S_i^T S_i)^{-1} \quad \text{...(B.9)}$$

$$= (S_i^T S_i)^{-1} \quad \text{otherwise.} \quad \text{...(B.10)}$$

The value of \( k \) is obtained from the estimated value \( \text{EST} \) of the function at the unconstrained minimum as:

$$k = 2 \left( \text{EST} - f(X_i) \right) / (S_i^T \nabla f_i) \quad \text{...(B.11)}$$

In the second stage the lower and upper bounds on the optimal step size \( \lambda^*_i \) are established. For this
purpose the slope $df/d\lambda$ is estimated at points

$$= 0, h, 2h, 4h, 8h, \ldots, b_1, b_2, \ldots$$

where $\lambda$ is doubled each time and where $b_2$ is first of
which of these values at which either the slope $df/d\lambda$
becomes non-negative or the value of the function $f$
stops decreasing. If then follows that $\lambda^*$ is bounded in
the interval

$$b_1 < \lambda^* < b_2$$

The third stage uses the cubic interpolation to
find the estimated value $\lambda^*_e$ of the optimal step size.

The function $f(\lambda)$ is approximated by a cubic
equation in the interval $b_1, b_2$. The estimated optimum
step size $\lambda^*_e$ is obtained from the following relations.

Setting

$$f_{b_1} = f(\lambda = b_1) \quad \ldots (B.12)$$

$$f'_{b_1} = \frac{df(\lambda = b_1)}{d\lambda} \quad \ldots (B.13)$$

$$f_{b_2} = f(\lambda = b_2) \quad \ldots (B.14)$$

$$f'_{b_2} = \frac{df(\lambda = b_2)}{d\lambda} \quad \ldots (B.15)$$

$$Z = \frac{3(f_{b_1} - f_{b_2})}{b_2 - b_1} + f'_{b_1} + f'_{b_2} \quad (B.16)$$

$$W = (Z^2 - f_{b_1} f_{b_2})^{1/2} \quad \ldots (B.17)$$
The estimated optimum step size is given by

\[ \lambda^*_e = b_2 - \frac{(f_{b_2} + w - z)(b_2 - b_1)}{(f_{b_2} - f_{b_1} + 2w)} \]  \hspace{1cm} \text{(B.18)}

If both \( f_{b_1} \) and \( f_{b_2} \) are greater than \( f(\lambda^*_e) \) then \( \lambda^*_e \) is accepted as the estimate of the optimal step size \( \lambda^*_e \). Otherwise the interpolation is repeated over the subinterval \( b_1, \lambda^*_e \) or \( \lambda^*_e, b_2 \) according as \( f(\lambda^*_e) \) is positive or negative.

The conjugate gradient algorithm for minimization process has been illustrated in section 3.5.

The process locates the minimum of a quadratic function of \( N \) arguments in at most, \( N \) iterations. For non-quadratic functions it may take more than \( N \)-cycles for convergence. The rate of convergence is adversely affected if the iterations are continued in this way due to accumulation of errors in the computation of the search direction. The method is therefore periodically restarted with the steepest descent direction as the search direction after every \( N+1 \) iterations. This procedure retains the quadratic convergence property when applied to quadratic functions and at the same time overcomes the ill effects when applied to non-quadratic functions.
BIODATA

NAME : Abani Mohan Panda

FATHER'S NAME : Late Dibakar Panda

DATE OF BIRTH : 12.3.1952

NATIONALITY : Indian

ACADEMIC QUALIFICATIONS :

<table>
<thead>
<tr>
<th>Name of Exam.</th>
<th>Year of Passing</th>
<th>University</th>
<th>Division</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.Sc (Engg) (Electrical)</td>
<td>1977</td>
<td>Sambalpur</td>
<td>First</td>
</tr>
<tr>
<td>P.G.Dip (W.R.D.)</td>
<td>1983</td>
<td>Roorkee</td>
<td>First</td>
</tr>
<tr>
<td>M.E. (W.R.D.)</td>
<td>1984</td>
<td>Roorkee</td>
<td>First</td>
</tr>
</tbody>
</table>

EXPERIENCE :

(a) Served as Junior Engineer and Assistant Engineer in government of Orissa during 1978-1986.

(b) Lecturer in electrical Engineering Department, Indira Gandhi Institute of Technology, Sarang, Orissa, India during 1986-1992.

(c) Continuing as Assistant Professor in Electrical Engineering Department, Indira Gandhi Institute of Technology, Sarang, Orissa, India since 1992.