Computer-Aided Engineering

Adapted from

C.N.Nightingle and J.K.Fidler ., Computer-Aided Circuit and System Design.,U.K +IEEE papers

Lecture 3

Prof .Dr.Eng.ARN

Classical Theory of Minima and Maxima

- The use of computers has led to the development of the theory of optimization and a number of optimization algorithms are now available.
- Some of these will be discussed later.

Classical Theory of Minima and Maxima

Of all possible solutions that satisfy the constraints of a given problem, an optimal solution is one which yields the maximum (minimum) value of the given performance measure. The mathematical optimization problem becomes that of maximizing or minimizing a scalar function, say f ,of n real scalar independent variables $X_1, X_2, \ldots\ldots\ldots\ldots X_n$, often assembled for abbreviation into an n-component column vector:-

Classical Theory of Minima and Maxima

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, x_2, \ldots\ldots, x_n)^T$$

The definition of transpose is to mix two things up or cause them to change places with each other, or to rewrite equation in a different way, or to change something to a new form or place it in a new context,i.e column to be changed by a row

Classical Theory of Minima and Maxima

In the following, vectors are always underlined, and the superscript T denotes transposition. The fact that f is a function depending on the vector $\underline{x}$ is expressed by writing $f(\underline{x})$. These scalar variables, i.e.:- $X_1, X_2, \cdots . X_n$ should be adjusted to obtain the minimum, or the maximum, required. Minimizing a function is the same as maximizing the negative of the function, so there is no loss of generality in the following sections which are concerned exclusively with function minimization.

Classical Theory of Minima and Maxima

- The use of computers has led to the development of the theory of optimization and a number of optimization algorithms are now available.
- Some of these will be discussed later.

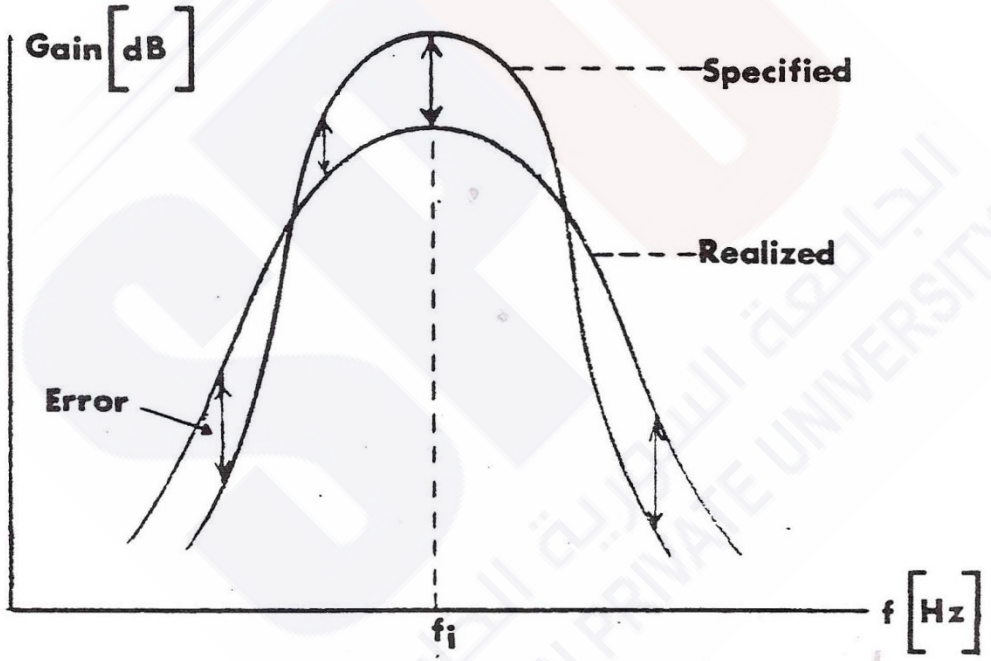
Objective Function and Function Minimization

Before discussing optimization algorithms themselves, it is instructive to consider the measure of the difference between the required and the actual performances. This measure is known in optimization theory as an objective or an error function $f(\underline{x})$, where $f(\underline{x})$ is a scalar non-linear function of n variables.

Objective Function and Function Minimization

Let the realized and corresponding specified performance measures be expressed as a set of m discrete quantities Pri and Psi respectively, where i = 1, 2, , m . The m quantities Psi could be for example, the coefficients of some network functions or a response specified at m discrete time or frequency points, as in the case of the amplitude characteristics as shown in the following Figure .

Objective Function and Function Minimization



Typical frequency domain design problem showing the formulation of amplitude matching error function

Objective Function and Function Minimization

Pri in turn will usually be a function of a set of variables $\underline{x}_j$. which are to be adjusted to minimize f(x) . In particular, f($\underline{x}$) will be a measure of error between Psi and Pri .Error functions used in network design problems tend to fall into two major classes :-

Objective Function and Function Minimization

(a)Least squares and

(a)Chebyshev.

Objective Function and Function Minimization

In least squares, the error function f(x) is assumed as the sum of the residual functions, which are squared to eliminate the occurrence (ظهور) of zero error values for non-zero residual values. Thus the problem is to minimize the least squares error:-

Classical Theory of Minima and Maxima

$$f(\underline{x}) \triangleq \sum_{i=1}^{m} (P_{si} - P_{ri})^2 \cdot W_i^2$$

Objective Function and Function Minimization

where Wi is a penalty or weighting function. The weighting function is introduced to influence the significance of individual residuals according to the requirement of the problem.

Objective Function and Function Minimization

In the Chebyshev approach, the error function takes the form:-

$f(\underline{x}) \triangleq \max_i [|w_i (Psi-Pri)|]$

- The maximum modulus error function results in an error that is minimum in the Chebyshev sense (minimax) in that the resulting residuals are equal and of alternating sign.

Objective Function and Function Minimization

- The disadvantage of the maximum modulus criterion is that discontinuous derivatives are generated when the maximum deviation jumps from one peak error to another as a function of $\underline{x}$

Objective Function and Function Minimization

It has been shown in some literature that the least squares error formulation usually leads to a better behaved function of $\underline{x}$. than does the Chebyshev. However, the choice of error function formulation depends largely on the application .

Objective Function and Function Minimization

The fundamental problem of optimization is to arrive at the best possible decision in any given set of circumstances, i.e. to obtain an optimum set of the independent variables corresponding to zero error.

Objective Function and Function Minimization

If there is no point of zero error on the functional surface then the point of lowest possible error is achieved. Most of the optimization methods described in the literature are methods for obtaining local rather than global minima. The local minimum represents the relative minimum, while the global minimum is the absolute minimum .

Objective Function and Function Minimization

In general, there will be constraints that must be satisfied either during the optimization or by the optimization solution for a minimum of the n variable function, f($\underline{x}$). The general non-linearly constrained minimization problem can be expressed mathematically as:

Objective Function and Function Minimization

$$\underline{x}_{\ell i} \leq \underline{x}_i \leq \underline{x}_{ui} \quad \text{for} \quad i = 1, 2, \ldots\ldots, n$$

where $\underline{\mathbf{X}}_{1i}$ and $\underline{\mathbf{X}}_{ui}$. are lower and upper bounds respectively subject to the m inequality constraints:-

$$C_j(\underline{x}) \geq 0 \quad \text{for} \quad j = 1, 2, \ldots\ldots, m$$

and k equality constraints:-

$$E_k(\underline{x}) = 0 \quad \text{for} \quad k = 1, 2, \ldots\ldots, k$$

Objective Function and Function Minimization

A problems for which there are no bounds or constraints is said to be unconstrained. The required conditions for a minimum may be derived by considering the multi-dimensional Taylor series expansion of the function about the minimum point X^* in the error space of $f(\underline{x})$ .Let a small perturbation Δx be made so that :-

Objective Function and Function Minimization

$$\underline{x} = \underline{x}^{*} + \Delta\underline{x}$$

- The first three terms of the Taylor series (which, requires that $f(\underline{x})$ be continuous and infinitely differentiable) so that:-

$$f(\underline{x}) = f(x^{*}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \underline{\Delta x}_i + {}^{1}/_{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} \underline{\Delta x}_j \underline{\Delta x}_i + \cdots\cdots$$

This equation may be written in matrix form as:-

$$f(\underline{x}) = f(\underline{x}^*) + \underline{g}^T(\underline{x}^*)\,\Delta\underline{x} + {}^1/2\,\Delta\underline{x}^T H(\underline{x}^*)\,\Delta\underline{x} + \ldots\ldots$$

where $\underline{x}$ is the minimum, and the transpose of Δx is given by:-

$$\Delta\underline{x}^T = [\Delta x_1 \;\; \Delta x_2 \;\; \Delta x_3 \;\; \ldots\ldots \;\; \Delta x_n]$$

The gradient vector $g(\underline{x})$ evaluated at $\underline{x}$ can be written as:-

Objective Function and Function Minimization

$$\underline{g}(\underline{x}) = \begin{Bmatrix} \vdots \\ \frac{\partial f}{\partial x_1} \\ \cdots \\ \frac{\partial f}{\partial x_2} \\ \cdots \\ \vdots \\ \vdots \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{Bmatrix}$$

where $\underline{g}^T$ is the transpose of the gradient vector, with n element.

Objective Function and Function Minimization

The nxn symmetric matrix of second order partial derivatives of f($\underline{x}$) is known as the Hessian matrix and is denoted by H where:-

$$H = \begin{Bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \\ \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{Bmatrix}$$

$$= \begin{Bmatrix} h_{11} & h_{12} & & h_{1n} \\ h_{21} & h_{22} & & h_{2n} \\ \\ \\ h_{n1} & h_{n2} & & h_{nn} \end{Bmatrix}$$

Objective Function and Function Minimization

A necessary condition for a minimum of $f(\underline{x})$ is that:-

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \ldots\ldots = \frac{\partial f}{\partial x_n} = 0 \qquad (3.14)$$

or

$$\underline{g}^T(\underline{x}^*) = 0 \qquad (3.15)$$

Objective Function and Function Minimization

A sufficient condition for a point satisfying equation (3.15) to be a minimum is that all the second partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}$ (where i and $j = 1, 2, 3, \ldots\ldots, n$) exist at this point and that $D_i > 0$ for $i = 1, 2, 3, \ldots\ldots, n$ where:-

$$D_1 = h_{11},\ D_2 = \begin{vmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{vmatrix},\ \ldots\ldots,\ D_n = |H| \qquad (3.16)$$

Objective Function and Function Minimization

i.e., the matrix of second partial derivatives (Hessian matrix H) must be positive definite. If H is positive semi-definite, then higher order terms of the Taylor expansion must be examined.

Objective Function and Function Minimization

Note that a solution of equation (3.14) does not necessarily represent the required minimum (global minimum), it could represent a local minimum, a maximum, or a saddle point. These situations are illustrated for various types of minimum for a two variable function $f(\underline{x}_1, \underline{x}_2)$ in Figure 3.3.

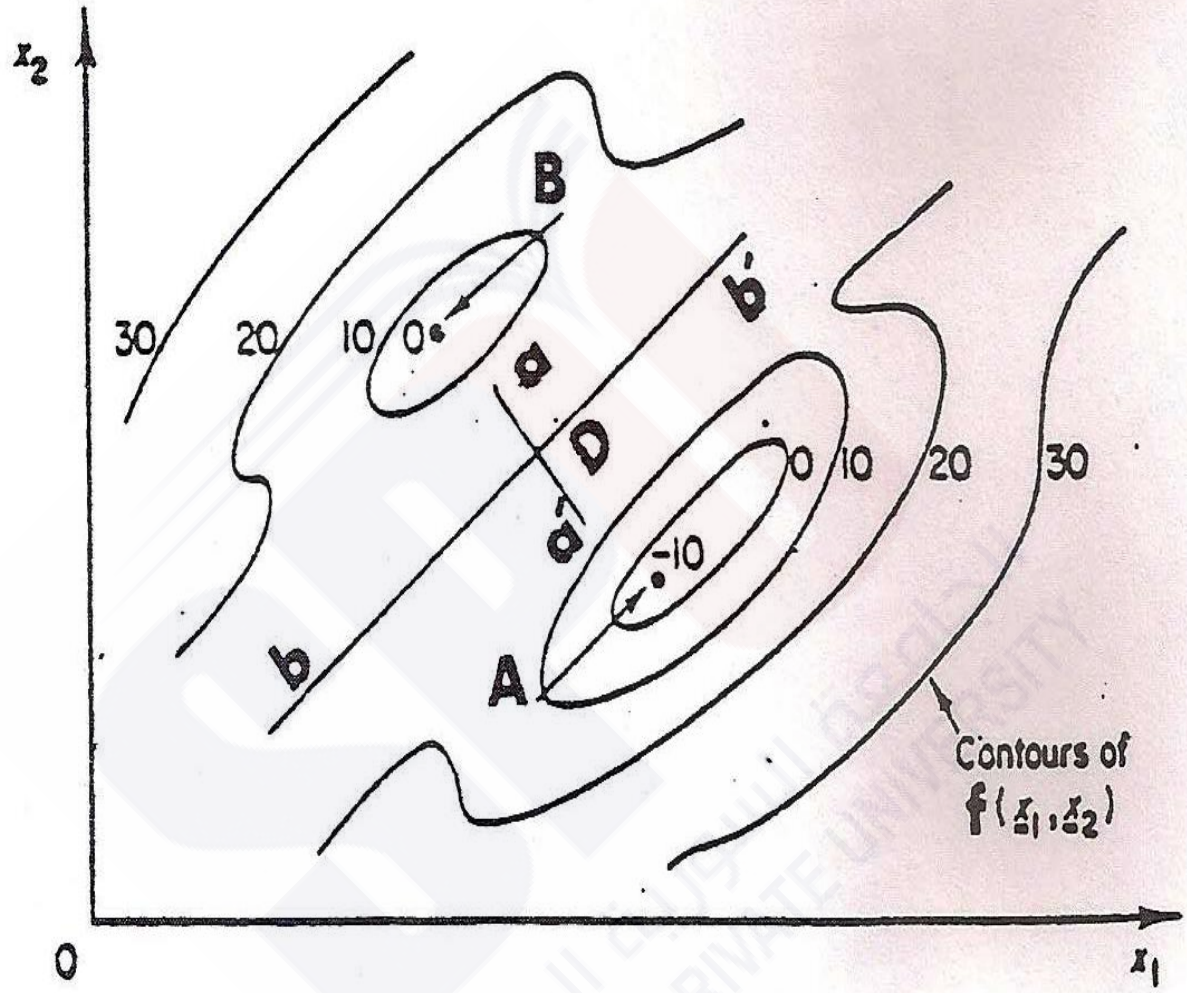


Figure 3.3 Contours of a function of two variables

In the Figure 3.3, A is a global minimum, i.e. the required overall minimum of the function. B is a local minimum, i.e. f(B) is less than $f(\underline{x})$ for all points x in the immediate neighbourhood of B . D is a saddle point, since along the direction aa', D is a maximum of f , whilst along bb' , D is a minimum of f .

In practice it is very difficult to determine whether the minimum obtained by a numerical process is a global minimum or not. In most circumstances it can only be said that the minimum obtained is a minimum within a local area of research. Therefore, the local minimum will be referred to simply as minimum hereafter.