HÁSKÓLI ÍSLANDS
VERKFRÆÐI- OG NÁTTÚRUVÍSINDASVIÐ
IÐNAÐARVERKFRÆÐI-, VÉLAVERKFRÆÐI-
OG TÖLVUNARFRÆÐIDEILD

VÉL 113F Design and Optimization Classical Methods

Magnus Thor Jonsson

Contents:

Summary:

What: Systematic search for the "best" design.

Why: Help you to stay competitive

How:

- Build a model of the system of interest (usually the hard part)
- Model has several parameters
  - Fixed
  - Others can be varied (<u>design variables</u>)
- Design goals are represented by an <u>objective function.</u>
- Design constraints are represented by <u>constraint functions.</u>
- Find values of the design parameters which <u>minimize</u> (or maximize) the objective function while satisfying all the constraints.

Introduction

Methods:

Mathematical programming or optimization techniques	Stochastic process techniques	Statistical methods
Calculus methods	Statistical decision theory	Regression analysis
Calculus of variations	Markov processes	Cluster analysis, pattern recognition
Nonlinear programming	Queueing theory	Design of experiments
Geometric programming	Renewal theory	Discriminate analysis (factor analysis)
Quadratic programming	Simulation methods	
Linear programming	Reliability theory	
Dynamic programming		
Integer programming		
Stochastic programming		
Separable programming		
Multiobjective programming		
Network methods: CPM and PERT		
Game theory		

Modern or nontraditional optimization techniques

Genetic algorithms
Simulated annealing
Ant colony optimization
Particle swarm optimization
Neural networks
Fuzzy optimization

Classification of optimization problems

Classification Based on the Nature of the Equations Involved

Classification Based on the Permissible Values of the Design Variables

Classification Based on the Deterministic Nature of the Variables

Classification Based on the Existence of Constraints

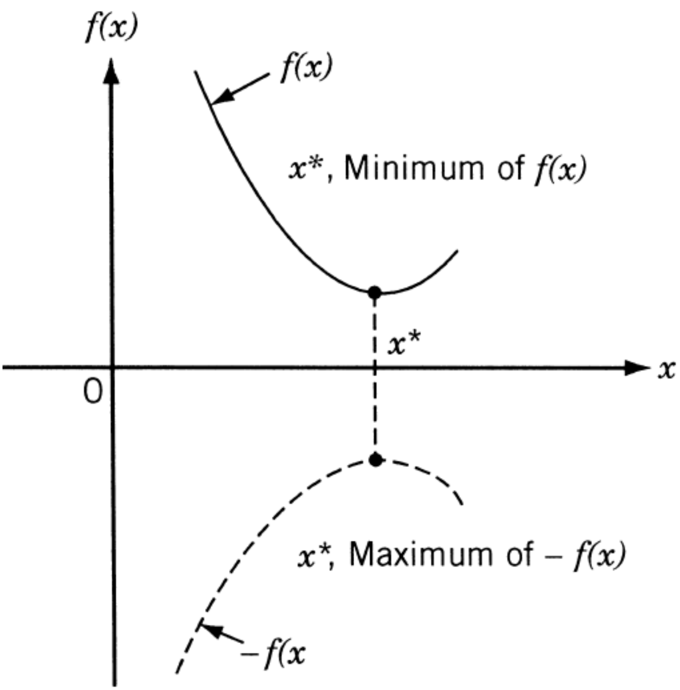
Classification Based on the Nature of the Design Variables

Classification Based on the Physical Structure of the Problem

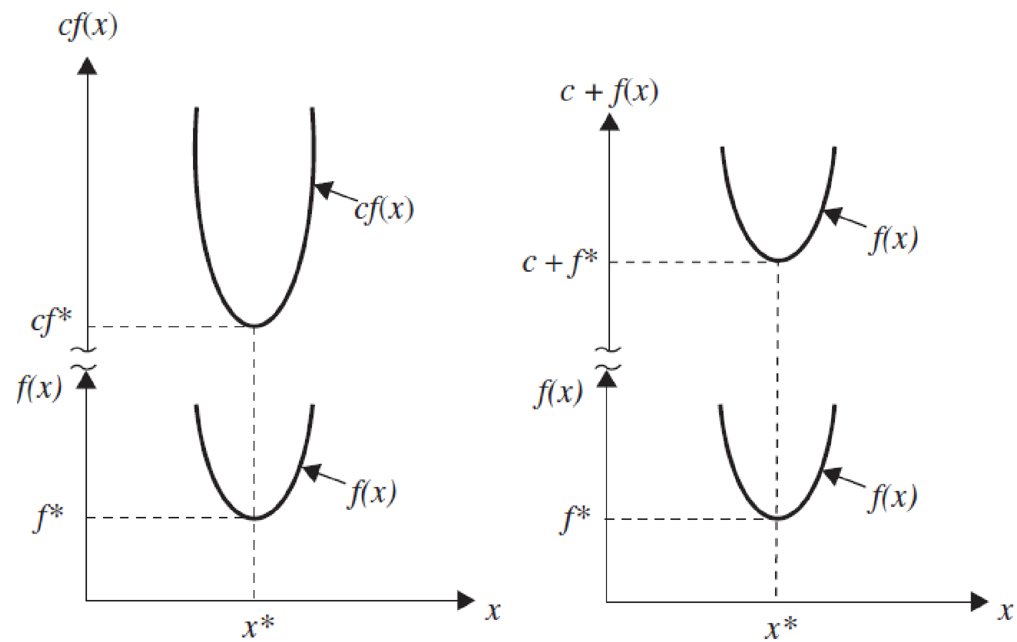
Classification Based on the Separability of the Functions

Classification Based on the Number of Objective Functions

Objective function max or min



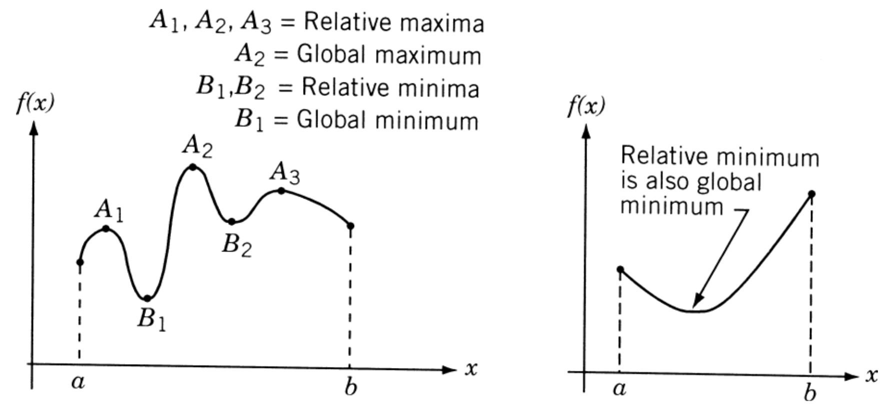
Minimum of $f(x)$ is same as maximum of $-f(x)$.



Optimum solution of $cf(x)$ or $c + f(x)$ same as that of $f(x)$.

Single variable optimization with no constraints:

A *single-variable optimization problem* is one in which the value of $x = x^*$ is to be found in the interval $[a, b]$ such that x^* minimizes $f(x)$.



Necessary Condition

If a function $f\ (x)$ is defined in the interval $a\ \leq\ x\ \leq\ b$ and has a relative minimum at $x = x^*$, where $a\ <\ x * <\ b$, and if the derivative $df(x)/dx\ =\ f'(x)$ exists as a finite number at $x = x^*$, then $f\ '(x\ *)\ =\ 0$.

Notes:

This can be proved even if x^* is a relative maximum.

This does not say what happens if a minimum or maximum occurs at a point x^* where the derivative fails to exist.

This does not say what happens if a minimum or maximum occurs at an endpoint of the interval of definition of the function.

This does not say that the function necessarily will have a minimum or maximum at every point where the derivative is zero.

Sufficient Condition

Let $f'(x^*) \ = \ f''(x^*) \ = \cdots = \ f^{n-1}(x^*) \ = \ 0,$

but $f^n(x^*) \ = \ 0.$

Then $f \ (x^*)$ is

(i) a minimum value of $f(x)$ if $f^n(x^*) > 0$ and n is even;

(ii) a maximum value of $f(x)$ if $f^n(x^*) \ < \ 0$ and n is even;

(iii) neither a maximum nor a minimum if n is odd.

Multivariable optimization with no constraints

Necessary Condition

If $f(\boldsymbol{x})$ has an extreme point (maximum or minimum) at $\boldsymbol{x} = \boldsymbol{x}^*$ and if the first partial derivatives of $f(\boldsymbol{x})$ exist at $\boldsymbol{x}^*$, then

$$\frac{\delta f}{\delta x^1}(\boldsymbol{x}^*) = \frac{\delta f}{\delta x^2}(\boldsymbol{x}^*) = \cdots = \frac{\delta f}{\delta x^n} = 0$$

Multivariable optimization with no constraints

Sufficient Condition

A sufficient condition for a stationary point $\boldsymbol{x}^*$ to be an extreme point is that the matrix of second partial derivatives (Hessian matrix) of $f(\boldsymbol{x})$ evaluated at $\boldsymbol{x}^*$ is:

(i) positive definite when $\boldsymbol{x}^*$ is a relative minimum point,

and

(ii) negative definite when $\boldsymbol{x}^*$ is a relative maximum point.

Multivariable optimization with no constraints

Hessian matrix

$$\mathbf{J}|_{\mathbf{X}=\mathbf{X}^*} = \left[\left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*} \right]$$

A matrix $\boldsymbol{A}$ will be positive definite if all its eigenvalues are positive; that is, all the values of λ that satisfy the determinantal equation $|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0$ should be positive.

Similarly, the matrix $[\boldsymbol{A}]$ will be negative definite if its eigenvalues are negative.

Multivariable optimization with no constraints

To find the positive definiteness of a matrix $\boldsymbol{A}$ of order n involves evaluation of the determinants

$$A = \left| a_{11} \right|,$$

$$A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},$$

$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{32} \end{vmatrix}, \ldots,$$

$$A_n = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & & & & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}$$

The matrix $\boldsymbol{A}$ will be positive definite if and only if all the values $A_1, A_2, A_3, \ldots, A_n$ are positive. The matrix $\boldsymbol{A}$ will be negative definite if and only if the sign of A_j is $(-1)^j$ for $j = 1, 2, \ldots, n$.

Multivariable optimization with equality constraints

$$\text{Minimize } f = f(\mathbf{X})$$

subject to

$$g_j(\mathbf{X}) = 0, \quad j = 1, 2, \ldots, m$$

$$\mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix}$$

Here m is less than or equal to n; otherwise (if m>n), the problem becomes overdefined and, in general, there will be no solution.

Multivariable optimization with equality constraints

Method of Lagrange Multipliers:

Problem with two variables and one constraint.

$$\text{Minimize } f\,(x_1, x_2)$$

subject to

$$g(x_1, x_2) = 0$$

For this problem, the necessary condition for the existence of an extreme point at

$$\boldsymbol{x} = \boldsymbol{x}^*$$

is

$$\left.\left(\frac{\partial f}{\partial x_1} - \frac{\partial f/\partial x_2}{\partial g/\partial x_2}\frac{\partial g}{\partial x_1}\right)\right|_{(x_1^*,\, x_2^*)} = 0$$

Method of Lagrange Multipliers:

By defining a quantity λ, called the *Lagrange multiplier*, as

$$\lambda = -\left(\frac{\partial f/\partial x_2}{\partial g/\partial x_2}\right)\Bigg|_{(x_1^*,\, x_2^*)}$$

Equations can be expressed as

$$\left(\frac{\partial f}{\partial x_1} + \lambda \frac{\partial g}{\partial x_1}\right)\Bigg|_{(x_1^*,\, x_2^*)} = 0$$

$$\left(\frac{\partial f}{\partial x_2} + \lambda \frac{\partial g}{\partial x_2}\right)\Bigg|_{(x_1^*,\, x_2^*)} = 0$$

Method of Lagrange Multipliers:

In addition, the constraint equation has to be satisfied at the extreme point, that is,

$$g(x_1, x_2)|_{(x_1^*, x_2^*)} = 0$$

Notice that the partial derivative

has to be nonzero to be able to define λ

Method of Lagrange Function:

The necessary conditions are more commonly generated by constructing a function L, known as the Lagrange function, as

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

By treating L as a function of the three variables x1, x2, and , the necessary conditions for its extremum are given by

$$\frac{\partial L}{\partial x_1}(x_1, x_2, \lambda) = \frac{\partial f}{\partial x_1}(x_1, x_2) + \lambda \frac{\partial g}{\partial x_1}(x_1, x_2) = 0$$
$$\frac{\partial L}{\partial x_2}(x_1, x_2, \lambda) = \frac{\partial f}{\partial x_2}(x_1, x_2) + \lambda \frac{\partial g}{\partial x_2}(x_1, x_2) = 0$$
$$\frac{\partial L}{\partial \lambda}(x_1, x_2, \lambda) = g(x_1, x_2) = 0$$

Method of Lagrange Function:

The necessary conditions are more commonly generated by constructing a function L, known as the Lagrange function, as

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

By treating L as a function of the three variables x1, x2, and , the necessary conditions for its extremum are given by

$$\frac{\partial L}{\partial x_1}(x_1, x_2, \lambda) = \frac{\partial f}{\partial x_1}(x_1, x_2) + \lambda \frac{\partial g}{\partial x_1}(x_1, x_2) = 0$$
$$\frac{\partial L}{\partial x_2}(x_1, x_2, \lambda) = \frac{\partial f}{\partial x_2}(x_1, x_2) + \lambda \frac{\partial g}{\partial x_2}(x_1, x_2) = 0$$
$$\frac{\partial L}{\partial \lambda}(x_1, x_2, \lambda) = g(x_1, x_2) = 0$$

Method of Lagrange Function:

The necessary conditions for general problem

$$\text{Minimize } f(\boldsymbol{x})$$

subject to

$$g_j(\boldsymbol{x}) = 0, \qquad j = 1, 2, \ldots, m$$

The Lagrange function, L, in this case is defined by introducing one Lagrange multiplier λ_j for each constraint $g_j(\boldsymbol{x})$ as

$$L(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m)$$
$$= f(\boldsymbol{x}) + \lambda_1 g_1(\boldsymbol{x}) + \lambda_2 g_2(\boldsymbol{x}) + \cdots + \lambda_m g_m(\boldsymbol{x})$$

Method of Lagrange Function:

By treating L as a function of the $n + m$ unknowns,

$$x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m,$$

the necessary conditions for the extremum of L, which also correspond to the solution of the original problem are given by

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \qquad i = 1, 2, \ldots, n$$

$$\frac{\partial L}{\partial \lambda_j} = g_j(\mathbf{X}) = 0, \qquad j = 1, 2, \ldots, m$$

Equations represent $n + m$ equations in terms of the $n + m$ unknowns, x_i and λ_j .

Method of Lagrange Function:

The solution gives

$$\mathbf{X}^* = \begin{Bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{Bmatrix} \quad \text{and} \quad \lambda^* = \begin{Bmatrix} \lambda_1^* \\ \lambda_2^* \\ \vdots \\ \lambda_m^* \end{Bmatrix}$$

The vector $\boldsymbol{X}^*$ corresponds to the relative constrained minimum of $f(\boldsymbol{X})$ (sufficient conditions are to be verified) while the vector λ^* provides the sensitivity information.

Method of Lagrange Functions

Sufficient Condition

A sufficient condition for $f(\boldsymbol{X})$ to have a relative minimum at $\boldsymbol{X}^*$ is that the quadratic, Q, defined by

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L}{\partial x_i \, \partial x_j} dx_i \, dx_j$$

evaluated at $\boldsymbol{X} = \boldsymbol{X}^*$ must be positive definite for all values of dX for which the constraints are satisfied.