1-General Introduction to optimization

The general constrained minimization(maximization) problem is defined by:

$$\underset{x\in R^n}{\text{minimize}} \text{ or maximize } f(x)$$

$$s.t \quad c_i(x) \geq b \;, 1 \leq i \leq m \qquad (1)$$

$$h_i(x) = b \;, m+1 \leq i \leq l$$

where $x \in R^n$ is real vector and $f : R^n \to R$ is a smooth function, and its as $c_i(x)$ is inequality constrained as $h_i(x)$ is equality constrained usually assumed to possess continuous second partial derivatives.

Note:

1)Of course, if f, $c1$, $c2$, . . . , cm *and* $h1$, $h2$, . . . , hl are all linear functions, then (1) is a linear programming problem and may be solved by the simplex algorithm.

There are basically two different kinds of constrained optimization approaches:

Indirect Method: changes the constrained optimization into unconstrained
optimization to be solved.
(*S*equential *U*nconstrained *M*inimization *T*echnique, SUMT)

Direct Method: deals with the constraints directly in the search for the Optimum.

Algorithms may be distinguished with respect to two main categories
- what kind of problem do they solve: unconstrained or constrained.
- which kind of data do they use:

function values only (zero order methods or direct search methods), or additionally gradient (gradient or first order methods), or even second order information (second order or Newton type methods).
As many strategies for constrained problems generate unconstrained sub-problems algorithms for the latter kind are important even for structural optimization. In the class of unconstrained problems we distinguish again

between minimization methods made for one unknown only, one-dimensional methods, and those made for many variables, n-dimensional methods. Methods for constrained problems are again subdivided with respect to the kind of variables (primal and/or dual variables) they prefer.

The subdivision is divided in to four sub classes.

-***Primal methods***, working directly in the n-dimensional space of the optimization variables **x**. They make no use of Lagrange multipliers and of the KKT necessary conditions. The most simple algorithms are primal, direct search methods. Evolutionary strategies and genetic algorithms belong to these species. They are preferably useful to handle discrete variables. Successful primal methods which make use of gradient information are known as method of feasible directions or general reduced gradient methods.

-***Penalty and Barrier function methods***, working also in the n-dimensional space of the optimization variables. They, however, transform the constrained problem into an unconstrained. The approach is principally simple and quite robust. An old methodology is known as SUMT (Sequential Unconstrained Minimization Technique) generates a series of unconstrained sub-problems to finally get a solution near to but not exactly at the optimum. It was one of the reasons why it became unpopular in the meantime. The basic idea, however, came back just recently as interior point method. The mathematical basis has been improved and SUMT has now been put into the frame of so called continuation methods.

-***Dual methods***, working primarily in the dual m-dimensional space of Lagrange multipliers ω, ϑ . The primal optimization variables **x** are determined by back substitution. Dual methods split the original optimization problem into two partial problems which have to be solved sequentially. One is unconstrained and formulated in terms of **x**, the other is formulated in terms of ω, ϑ and is only constrained by simple bounds. In the case of equality constraints, it is unbounded too. Because of the simply structured sub-problems methods for unconstrained problems can be applied directly or with minor modifications to handle bounds on variables.

-***Lagrange methods***, working in the full (n+m)-dimensional space of primal and dual variables **x,** ω, ϑ . They directly tackle the Kuhn-Tucker necessary conditions by solving a sequence of linearized sub-problems. Consistently derived these sub problems are characterized by a quadratic objective and linear constraints. That’s why these kinds of methods are called SQP- or Sequential Quadratic Programming methods. A simplified variant uses a linear approximation for the objective also and is called SLP- or Sequential Linear Programming. SQP methods are considered to be one of the most or even the most sophisticated methods from the mathematical point of view. They have been successfully applied for many structural optimization tasks and are available in almost every structural optimization package. However, they appear to be not robust enough for very large problems. Research on the field is still in progress.

A point $x^{\wedge}$ which satisfies all the functional constraints is said to be a feasible point. A fundamental concept, that provides a great deal of insight as well, as simplifies the theoretical development, is that of an active constraint. An inequality constraint $c\,(x) \le 0_{\,i}$ is said to be active at feasible point $x^{\wedge}$ if $c\,(x) = 0_{\,i}$ and inactive at x if $c\,(x) \prec 0_{\,i}$. By convention, we refer to any equality constraint $h\,(x) = 0_{\,i}$ as active at any feasible point

2 Basic Definitions:

This section contains a collection of some basic definitions, properties and results about constant real –valued matrices.

Definition 1:

Optimization means finding the best solution in some sense to a given problem. Mathematically, this means finding the minimum or maximum value of a function of n variables,

Definition 2:

Nonlinear programming means non-linear objective function or nonlinear constraints or both

Definition (3)

A set $S \subseteq R^n$ is *convex*, if for any $x_1, x_2 \in S$ and for any $\nu \in [0,1]$, we have

$$\nu x_1 + (1-\nu)x_2 \in S,$$

Definition (4)

Let $f : S \to R$, where S is a nonempty convex set in R^n. The function f is said to be *convex* on S if:

$$f(\nu x_1 + (1-\nu)x_2) \le \nu f(x_1) + (1-\nu) f(x_2),$$

for each $x_1, x_2 \in S$ and for each $\nu \in (0,1)$, and we said the function f is *strictly convex* on S, if

$$f(\nu x_1 + (1-\nu)x_2) < \nu f(x_1) + (1-\nu) f(x_2),$$

for all $x_1 \neq x_2$ in S and for each $\nu \in (0,1)$,

Definition (5)

Let $f : S \to R$, where S is a nonempty convex set in R^n. The function f is said to be *uniformly convex* on S, if there is a constant t >0 such that for any $x_1, x_2 \in S$,

$$f(\nu x_1 + (1-\nu)x_2) < \nu f(x_1) + (1-\nu) f(x_2) - \frac{1}{2} t\nu(1-\nu)\|x_1 - x_2\|^2,$$

Definition (6)

A point $x^* \in R^n$ is a *local minimizer* of $f : R^n \to R$, if there exists a neighborhood $U = \{x \in R^n : \|x - x^*\| < \varepsilon\}$,

such that f attains its minimum value in U at x^*, i.e.

$$f(x^*) \leq f(x), \ \forall x \in U,$$

Definition (7)

A point $x^* \in R^n$ is a *global minimizer* of $f : R^n \to R$, if f attains there is smallest value in R^n, i.e.

$$f(x^*) \leq f(x), \ \forall x \in R^n,$$

Definition (8)

The *Euclidean norm* of a vector x is defined as:

$$\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2} \quad ,$$

3 Review of Optimality Condition
First-Order Optimality Condition
Second-Order Optimality Condition