

ADVANCE METHOD FOR SOLUTION OF CONSTRAINT OPTIMIZATION PROBLEMS

Dr. Binay Kumar

Assistant Professor, Department of Electrical & Electronics Engineering

Dronacharya College of Engineering, Gurgaon, India

ABSTRACT:

Constraint optimization problems are encountered in numerous applications. There are different areas like Engineering design, structural optimization, VLSI design, economics & allocation problem can be applicable constraint optimization problem approach. In this present paper we have developed advance solution approach through extended Saddle points, Lagrange multipliers and penalty methods for solving constrained-optimization problems. Here studies some new theorems have been stated and simple proofs have been given. The method can be directly used to solve practical problems.

Keywords: *Optimization, Constraint, Lagrange multiplier, Saddle points & Penalty method*

INTRODUCTION

1.1 Constraint optimization problem can be defined as a regular constraint satisfaction problem augmented with a number of "local" cost functions. The aim of constraint optimization is to find a solution to the problem whose cost, evaluated as the sum of the cost functions, is maximized or minimized. The regular constraints are called hard constraints, while the cost functions are called soft constraints. These names illustrate that hard constraints are to be necessarily satisfied, while soft constraints only express a preference of some solutions (those having a high or low cost) over other ones (those having lower/higher cost).

A general constrained optimization problem may be written as follows:

$$\begin{array}{ll} \max & f(\mathbf{x}) \\ \text{subject to} & g_i(\mathbf{x}) = c_i \quad \text{for } i = 1, \dots, n \quad \text{Equality constraints} \\ & h_j(\mathbf{x}) \leq d_j \quad \text{for } j = 1, \dots, m \quad \text{Inequality constraints} \end{array}$$

Where \mathbf{X} is a vector residing in a n-dimensional space, $f(\mathbf{x})$ is a scalar valued objective function, $g_i(\mathbf{x}) = c_i$ for $i = 1, \dots, n$ and $h_j(\mathbf{x}) \leq d_j$ for $j = 1, \dots, m$ are constraint functions that need to be satisfied.

Constrained optimization problems: which are subject to one or more constraints.

Unconstrained optimization problems: in which no constraints exist.

1.2 Classification of Optimization Problem

1.2.1 In the first category the objective is to find a set of design parameters that makes a prescribed function of these parameters minimum or maximum subject to certain constraints. For example to find the minimum weight design of a strip footing with two loads shown in Fig 1 (a) subject to a limitation on the maximum settlement of the structure can be stated as follows.

Find $\mathbf{X} = \begin{Bmatrix} b \\ d \end{Bmatrix}$ which minimizes

$$f(\mathbf{X}) = h(b, d)$$

Subject to the constraints $\delta_s(\mathbf{X}) \leq \delta_{\max}; b \geq 0; d \geq 0$

where δ_s is the settlement of the footing. Such problems are called parameter or static optimization problems.

It may be noted that, for this particular example, the length of the footing (l), the loads P_1 and P_2 and the distance between the loads are assumed to be constant and the required optimization is achieved by varying b and d .

1.2.2 In the second category of problems, the objective is to find a set of design parameters, which are all continuous functions of some other parameter that minimizes an objective function subject to a set of constraints. If the cross sectional dimensions of the rectangular footings are allowed to vary along its length as shown in Fig 1 (b), the optimization problem can be stated as :

Find $\mathbf{X}(t) = \begin{Bmatrix} b(t) \\ d(t) \end{Bmatrix}$ which minimizes

$$f(\mathbf{X}) = g(b(t), d(t))$$

Subject to the constraints

$$\delta_s(\mathbf{X}(t)) \leq \delta_{\max} \quad 0 \leq t \leq l$$

$$b(t) \geq 0 \quad 0 \leq t \leq l$$

$$d(t) \geq 0 \quad 0 \leq t \leq l$$

The length of the footing (l) the loads P_1 and P_2 , the distance between the loads are assumed to be constant and the required optimization is achieved by varying b and d along the length l .

Here the design variables are functions of the length parameter t . this type of problem, where each design variable is a function of one or more parameters, is known as trajectory or dynamic optimization problem.

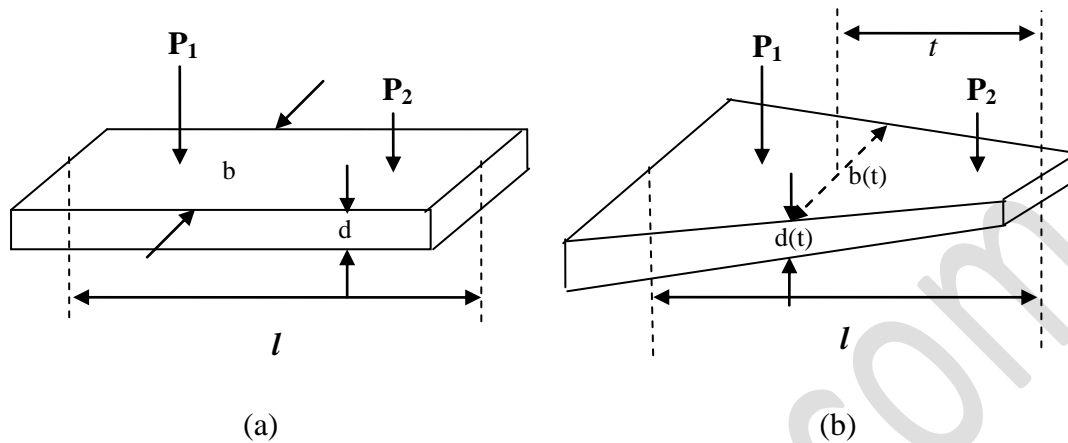


Figure 1

1.3 Classification based on the physical structure of the problem

Based on the physical structure, optimization problems are classified as optimal control and non-optimal control problems.

(i) Optimal control problems

An optimal control (OC) problem is a mathematical programming problem involving a number of stages, where each stage evolves from the preceding stage in a prescribed manner. It is defined by two types of variables: the control or design and state variables. The control variables define the system and controls how one stage evolves into the next. The state variables describe the behavior or status of the system at any stage. The problem is to find a set of control variables such that the total objective function (also known as the performance index, PI) over all stages is minimized, subject to a set of constraints on the control and state variables. An OC problem can be stated as follows:

$$\text{Find } \mathbf{X} \text{ which minimizes } f(\mathbf{X}) = \sum_{i=1}^l f_i(x_i, y_i)$$

Subject to the constraints

$$q_i(x_i, y_i) + y_i = y_{i+1} \quad i = 1, 2, \dots, l$$

$$g_j(x_j) \leq 0, \quad j = 1, 2, \dots, l$$

$$h_k(y_k) \leq 0, \quad k = 1, 2, \dots, l$$

Where x_i is the i th control variable, y_i is the i th state variable, and f_i is the contribution of the i th stage to the total objective function. g_j , h_k , and q_i are the functions of $x_j, y_j; x_k, y_k$ and x_i and y_i , respectively, and l is the total number of states. The control and state variables x_i and y_i can be vectors in some cases.

(ii) Problems which are not optimal control problems are called non-optimal control problems.

VARIOUS SOLUTION METHOD

2.1 Branch and bound Method

2.1.1 Constraint optimization can be solved by [branch and bound](#) algorithms. These are backtracking algorithms storing the cost of the best solution found during execution and use it for avoiding part of the search [1]. More precisely, whenever the algorithm encounters a partial solution that cannot be extended to form a solution of better cost than the stored best cost, the algorithm backtracks, instead of trying to extend this solution.

2.1.2 Assuming that cost is to be maximized, the efficiency of these algorithms depends on how the cost that can be obtained from extending a partial solution is evaluated. Indeed, if the algorithm can backtrack from a partial solution, part of the search is skipped. The lower the estimated cost, the better the algorithm, as a lower estimated cost is more likely to be lower than the best cost of solution found so far.

On the other hand, this estimated cost cannot be lower than the effective cost that can be obtained by extending the solution, as otherwise the algorithm could backtrack while a solution better than the best found so far exists. As a result, the algorithm requires an upper bound on the cost that can be obtained from extending a partial solution, and this upper bound should be as small as possible.

2.2 First-choice bounding functions

One way for evaluating this upper bound for a partial solution is to consider each soft constraint separately. For each soft constraint, the maximal possible value for any assignment to the unassigned variables is assumed. The sum of these values is an upper bound because the soft constraints cannot assume a higher value. It is exact because the maximal values of soft constraints may derive from different evaluations: a soft constraint may be maximal for $x = a$ while another constraint is maximal for $x = b$.

2.3 Russian doll search

This method runs a branch-and-bound algorithm on n problems, where n is the number of variables. Each such problem is the sub problem obtained by dropping a sequence of variables x_1, \dots, x_i from the original problem, along with the constraints containing them. After the problem on variables x_{i+1}, \dots, x_n is solved, its optimal cost can be used as an upper bound while solving the other problems,

In particular, the cost estimate of a solution having x_{i+1}, \dots, x_n as unassigned variables is added to the cost that derives from the evaluated variables [2]. Virtually, this corresponds on ignoring the evaluated variables and solving the problem on the unassigned ones, except that the latter problem has already been solved. More precisely, the cost of soft constraints containing both assigned and unassigned variables is estimated as above (or using an arbitrary other method); the cost of soft constraints containing only unassigned variables is instead estimated using the optimal solution of the corresponding problem, which is already known at this point.

2.4 Bucket elimination

2.4.1 The [bucket elimination](#) algorithm can be adapted for constraint optimization. A given variable can be indeed removed from the problem by replacing all soft constraints containing it with a new soft constraint. The cost of this new constraint is computed assuming a maximal value for every value of the removed variable. Formally, if x is the variable to be removed, C_1, \dots, C_n are the soft constraints containing it, and y_1, \dots, y_m are their variables except x , the new soft constraint is defined by:

$$C(y_1 = a_1, \dots, y_n = a_n) = \max_a \sum_i C_i(x = a, y_1 = a_1, \dots, y_n = a_n)$$

2.4.2 Bucket elimination [3] works with an (arbitrary) ordering of the variables. Every variable is associated a bucket of constraints; the bucket of a variable contains all constraints having the variable has the highest in the order. Bucket elimination proceed from the last variable to the first. For each variable, all constraints of the bucket are replaced as above to remove the variable. The resulting constraint is then placed in the appropriate bucket.

The [bucket elimination](#) algorithm can be adapted for constraint optimization. A given variable can be indeed removed from the problem by replacing all soft constraints containing it with a new soft constraint. The cost of this new constraint is computed assuming a maximal value for every value of the removed variable. Formally, if x is the variable to be

removed, C_1, \dots, C_n are the soft constraints containing it, and y_1, \dots, y_m are their variables except x , the new soft constraint is defined by:

$$C(y_1 = a_1, \dots, y_n = a_n) = \max_a \sum_i C_i(x = a, y_1 = a_1, \dots, y_n = a_n)$$

2.4.3 Bucket elimination works with an (arbitrary) ordering of the variables. Every variable is associated a bucket of constraints; the bucket of a variable contains all constraints having the variable has the highest in the order. Bucket elimination proceed from the last variable to the first. For each variable, all constraints of the bucket are replaced as above to remove the variable. The resulting constraint is then placed in the appropriate bucket.

2.5 Distributed constraint optimization

(DCOP or DisCOP) is the [distributed](#) analogue to [constraint optimization](#). A DCOP is a problem in which a group of agents must distributed choose values for a set of variables such that the cost of a set of constraints over the variables is either minimized or maximized. Distributed Constraint Satisfaction is a framework for describing a problem in terms of constraints that are known and enforced by distinct participants (agents). The constraints are described on some variables with predefined domains, and have to be assigned to the same values by the different agents.

2.6 Penalty function methods

Power methods are a certain class of [algorithms](#) for solving [constrained optimization](#) problems [4]. A penalty method replaces a constrained optimization problem by a series of unconstrained problems whose solutions ideally converge to the solution of the original constrained problem. The unconstrained problems are formed by adding a term to the [objective function](#) that consists of a penalty parameter and a measure of violation of the constraints. The measure of violation is nonzero when the constraints are violated and is zero in the region where constraints are not violated. Maximize

$$Z = f(x)$$

$$\text{s.t. the constraint } g(x) = c$$

The objective function to be maximized becomes

$$W = Z - p(g(x) - c)^2 \text{ with } p \rightarrow \infty$$

If $g(x) \neq c$ then $(g(x) - c)^2$ would be positive and with large p value, the objective function value drops precipitously. Therefore, it would be the decision maker's advantage to stick to the given constraint.

Similarly, the minimization problem

$$\text{Minimize } Z = f(x)$$

$$\text{s.t. the constraint } g(x) = c$$

assumes the form

$$W = Z + p(g(x) - c)^2 \text{ with } p \rightarrow \infty$$

After such formulation, we would obtain our results via

$$\frac{\partial W}{\partial x} = \mathbf{0}, \quad \frac{\partial W}{\partial p} = \mathbf{0}$$

Example. Find the point on the parabola $y^2 = 4x$ that is closest to the point (1,0).

In this case, the objective function that we need to minimize is

$$\text{Min } Z = (x - 1)^2 + y^2$$

$$\text{s.t. } y^2 = 4x$$

Using the penalty p , we seek to minimize the objective function

$$W = (x - 1)^2 + y^2 + p(y^2 - 4x)^2$$

This yields
$$\frac{\partial W}{\partial x} = 2(x - 1) - 8p(y^2 - 4x) = \mathbf{0}$$

$$\frac{\partial W}{\partial y} = 2y + 4p(y^2 - 4x) = \mathbf{0}$$

and,
$$\frac{\partial W}{\partial p} = y^2 - 4x = \mathbf{0}$$

From the first two, we get $x + 2y = 1$. Substituting this into the constraint, we get

$$y^2 + 8y - 4 = \mathbf{0} \text{ which gives the optimum } y^* \text{ as}$$

$$y^* = \frac{-8 \pm \sqrt{80}}{2} = -4 + 2\sqrt{5} = \mathbf{0.47214} \text{ and } x^* = \mathbf{0.055728}$$

Interior point methods (also referred to as barrier methods)

2.7.1 Barrier methods constitute an alternative class of algorithms for constrained optimization. These methods also add a penalty-like term to the objective function, but in this case the iterates are forced to remain interior to the feasible domain and the barrier is in place to bias the iterates to remain away from the boundary of the feasible region [5]. Interior point methods are a certain class of [algorithms](#) to solve linear and nonlinear [convex optimization](#) problems.

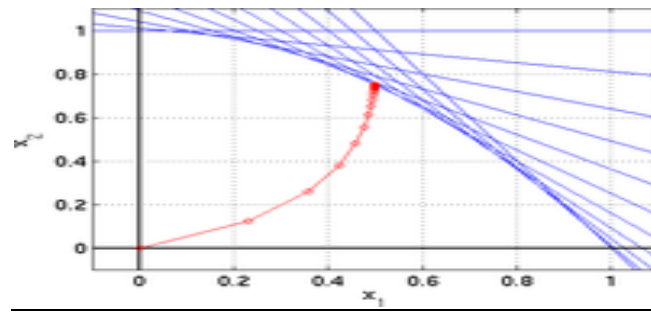


Figure 2

2.7.2 The interior point method was invented by [John von Neumann](#). Von Neumann suggested a new method of linear programming, using the homogeneous linear system of Gordan which was later popularized by [Karmarkar's algorithm](#), developed by [Narendra Karmarkar](#) in 1984 for [linear programming](#) [6]. The method consists of a [self-concordant barrier function](#) used to encode the [convex set](#). Contrary to the [simplex method](#), it reaches an optimal solution by traversing the interior of the feasible region. Any convex optimization problem can be transformed into minimizing (or maximizing) a [linear function](#) over a convex set by converting to the [epigraph form](#). The idea of encoding the [feasible set](#) using a barrier and designing barrier methods was studied in the early 1960s by, amongst others, Anthony V. Fiacco and Garth P. McCormick. These ideas were mainly developed for general [nonlinear programming](#), but they were later abandoned due to the presence of more competitive methods for this class of problems (e.g. [sequential quadratic programming](#)).

2.7.3 [Yurii Nesterov](#) and [Arkadi Nemirovski](#) came up with a special class of such barriers that can be used to encode any convex set. They guarantee that the number of [iterations](#) of the algorithm is bounded by a polynomial in the dimension and accuracy of the solution. Karmarkar's breakthrough revitalized the study of interior point methods and barrier problems, showing that it was possible to create an algorithm for linear programming characterized by [polynomial complexity](#) and, moreover, that was competitive with the simplex method. Already [Khachiyan's ellipsoid method](#) was a polynomial time algorithm; however, in practice it was too slow to be of practical interest.

The class of primal-dual path-following interior point methods is considered the most successful. [Mehrotra's predictor-corrector algorithm](#) provides the basis for most implementations of this class of methods. Primal-dual method's idea is easy to demonstrate for constrained [nonlinear optimization](#) [7].

2.8 Karush–Kuhn–Tucker (KKT) conditions (also known as the Kuhn–Tucker conditions)

2.8.1 Kuhn–Tucker conditions method are first order [necessary conditions](#) for a solution in [nonlinear programming](#) to be [optimal](#), provided that some [regularity conditions](#) are satisfied. Allowing inequality constraints, the KKT approach to nonlinear programming generalizes the method of [Lagrange multipliers](#), which allows only equality constraints. The system of equations corresponding to the KKT conditions is usually not solved directly, except in the few special cases where a [closed-form](#) solution can be derived analytically. In general, many optimization algorithms can be interpreted as methods for numerically solving the KKT system of equations.

2.8.2 The KKT conditions were originally named after [Harold W. Kuhn](#), and [Albert W. Tucker](#), who first published the conditions in 1951. Later scholars discovered that the necessary conditions for this problem had been stated by [William Karush](#) in his master's thesis in 1939

Example: Optimise the objective function $x^{0.25}y^{0.25}$ subject to the constraint $24 = x/10 + y$

The Lagrange multiplier method:

Step 1: The Lagrangian:

$$L = x^{0.25}y^{0.25} + \lambda(24 - x/10 - y)$$

Step 2: $\frac{\partial L}{\partial x} = 0.25x^{-0.75}y^{0.25} - \lambda/10 = 0$

$$\frac{\partial L}{\partial y} = 0.25x^{0.25}y^{-0.75} - \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = 24 - x/10 - y = 0$$

Step 3: Solve the 3 simultaneous equations:

$$0.25x^{0.25}y^{-0.75} - \lambda = 0$$

$$0.25x^{0.25}y^{-0.75} = \lambda$$

$$0.25x^{-0.75}y^{0.25} - (0.25x^{0.25}y^{-0.75})/10 = 0$$

$$2.5x^{-0.75}y^{0.25} = 0.25x^{0.25}y^{-0.75}$$

$$2.5y^{0.25}/y^{-0.75} = 0.25x^{0.25}/x^{-0.75}$$

$$2.5y = 0.25x$$

$$10y = x$$

$$24 - x/10 - y = 0$$

$$24 - 10y/10 - y = 0$$

$$24 - 2y = 0$$

$$y = 12$$

$$10y = x$$

$$10(12) = x$$

$$120 = x$$

Optimum point at: $x = 120, y = 12$

EXTENDED SADDLE POINT & LAGRANGE MULTIPLIERS APPROACH

3.1 Variational problems arise in constrained minimization problems we seek a minimum of F subject to the constraint that the minimizers lie in a convex set K . The specification of the constraint on F is equivalent to specifying the constraint set K . As it is well known, constraint minimization problem can be reformulated as saddle point problems using the method of Lagrange multipliers. Such formulations sometimes may make it possible to seek minima of functional in linear spaces rather than closed convex sets.

Let U and V be Banach spaces.

K and M be non-empty close convex subsets of

U and V , respectively and $L: K \times M \rightarrow \mathbb{R}$ a real functional defined on $K \times M$.

We recall that a pair $(u, p) \in K \times M$ is a saddle point of L if and only if

$$L(u, q) \leq L(u, p) \leq L(v, p), \text{ for all } v \in K, q \in M \quad (1)$$

$$\text{The functional } L \text{ possesses a saddle point if} \quad (2)$$

$$\max_{q \in M} \inf_{v \in K} L(v, q) = \min_{v \in K} \sup_{q \in M} L(v, q)$$

$$q \in M \quad v \in K \quad q \in M$$

We will be primarily concerned with cases in which the following conditions hold:

$$\text{For all } q \in M, v \rightarrow L(v, q) \text{ is } G\text{-differentiable} \quad 3 \text{ (a)}$$

$$\text{For all } v \in K, q \rightarrow L(v, q) \text{ is } G\text{-differentiable} \quad 3 \text{ (b)}$$

$$\text{For all } q \in M, v \rightarrow L(v, q) \text{ is strictly convex} \quad 3 \text{ (c)}$$

(and therefore weakly lower semi continuous)

$$\text{For all } v \in K, q \rightarrow L(v, q) \text{ is concave and upper semi continuous.} \quad 3 \text{ (d)}$$

We have the following theorems on the existence and characterization of saddle point.

3.2 Theorem (1) : Let condition (3a) hold. In addition suppose that either K and M are bounded or L is coercive in the following sense:

$\exists q_0 \in M$ such that

Lim

$$\| \| v \| \| u \rightarrow \infty L(v, q_0) = +\infty \quad (4)$$

and $v_0 \in K$ such that

Lim

$$\| \| q \| \| v \rightarrow \infty L(v_0, q) = -\infty \quad (5)$$

Then there exists a saddle point $(u, p) \in K \times M$ of L .

Moreover,

$$L(U, P) = \min_{v \in K} \sup_{q \in M} L(v, q) = \max_{q \in M} \inf_{v \in K} L(v, q) \quad (6)$$

3.3 Theorem (2): Let (3a) and (3b) hold and also let

$\phi : K \times M \rightarrow \mathbb{R}$ be such that for all

$\forall v \in K, q \in M, \phi(v, q) \leq L(v, q)$

ϕ is concave and for all $q \in M, \phi(v, q)$ is convex. Then, If (u, p) is a saddle point of the functional L , it is characterized by the pair of variational inequalities.

$$\begin{aligned} \frac{\delta L(u, p)}{\delta v}, v - u > \phi(v, p) - \phi(u, p) &\geq 0 \text{ for all } v \in K \\ \frac{\delta L(u, p)}{\delta q}, q - p > \phi(u, q) - \phi(u, p) &\leq 0 \text{ for all } q \in M \end{aligned} \quad (7)$$

Where $\langle \cdot, \cdot \rangle_u$ and $\langle \cdot, \cdot \rangle_v$ denote duality pairing on $U' \times U$ and $V' \times V$, respectively, and

$\frac{\delta L(u, p)}{\delta v}$ and $\frac{\delta L(u, p)}{\delta q}$ denote the gradients of L with respect to v and q for fixed q and v , respectively.

Remark (1): Let us now return to the minimization problem of finding u in a non-empty closed convex set K Such that

$$\begin{aligned} \inf_{v \in K} F(v) &= F(u) \\ v &\in K \end{aligned} \quad (8)$$

We will reformulate this problem using Lagrange multipliers for the case in which

$$K = \{v \in U : B(v) = 0\} \quad (9)$$

Where B is an operator mapping U into the dual V' of reflexive Banach space V . Towards this end, we introduce the Lagrangian $L: U \times V \rightarrow \mathbb{R}$ defined by

$$L(v, q) = F(v) + \langle B(v), q \rangle_v \quad (10)$$

We will make the following assumptions on F and B :

$F: U \rightarrow \mathbb{R}$ is G -differentiable, coercive, strictly convex and its gradient DF is bounded,

$B: U \rightarrow V'$ is weakly sequentially continuous

$V \rightarrow \langle B(u), q \rangle_v$ is G -differentiable and

$$\lim_{\epsilon \rightarrow 0} \frac{\delta}{\epsilon} \langle B(u + v), q \rangle_v = \langle c(u)q, v \rangle_v = \langle c(u)v, q \rangle_v$$

$$\text{Where } C: U \times V \rightarrow U', C^*: U \times U \rightarrow V' \quad (11)$$

Remark (2): Returning to theorem (1) and particularly conditions (3),

we see that all of the conditions of that theorem are met except (5)

To ensure that (5) be also satisfied, it is customary to introduce a Perturbed lagrangian

$$L(v, q) = L(v, q) - \epsilon \|q\|_v \quad (12)$$

Where ϵ is an arbitrary positive number, we easily verify that all of the conditions of theorem (7) are met by L_ϵ .

Thus, for each $\epsilon > 0$, there exists a saddle point $(u_\epsilon, p_\epsilon) \in U \times V$ i.e.,

$$L_\epsilon(u_\epsilon, q) \leq L_\epsilon(u_\epsilon, p_\epsilon) \leq L_\epsilon(v, p_\epsilon), \text{ for all } v \in U, q \in V \quad (13)$$

Moreover, since F is coercive, the sequence $\|u_\epsilon\|$ is uniformly bounded in ϵ ,

Since U is reflexive, there exists a subsequence, also denoted by u_ϵ and an element $u \in U$ such that

$$u_\epsilon \rightarrow \text{weakly in } U \quad (14)$$

In addition, it can be also shown that the limit u is a solution of the original minimization problem (8).

An additional condition is needed to guarantee the existence of a p_ϵ such that $p_\epsilon \rightarrow P$ weakly in V . To arrive at a suitable condition on P_ϵ , let us denote

$$\lim_{\lambda \rightarrow \delta_+} \frac{\delta}{\lambda} \|p + \lambda q\|_v = E(p), q \in V$$

Then saddle points of the perturbed Lagrangian L_ϵ of (12) are characterized by

$$DF(u_\epsilon), \langle B(u_\epsilon), q \rangle_v - \epsilon \langle P_\epsilon, q \rangle_v = 0, \text{ for all } q \in V \quad (15)$$

Let us suppose that a constant $\alpha_0 > 0$ exists independent of ϵ , such that

$$\frac{\sup_{V \in U} | \langle C(u_\epsilon)q, V \rangle_U |}{\|V\|} > \alpha_0 \|q\| \quad (16)$$

Since DF is assumed to be bounded, we have from (15)

$$| \langle C(u_\epsilon) P_\epsilon, V \rangle_U | = | \langle DF(u_\epsilon), V \rangle_U | \leq C \|V\|_U \quad (C = \text{constant})$$

Hence, if (16) holds,

$$\alpha_0 \|P_\epsilon\| \leq C$$

i.e. the sequence P_ϵ is uniformly bounded in ϵ and, therefore had a sequence P_{ϵ_k} which converges weakly to an element p in V as $\epsilon \rightarrow 0$. The limit (u, p) of the subsequence $(u_{\epsilon_k}, p_{\epsilon_k})$ of solutions of (15) is a saddle point of the original functional L of (10)

3.4 Advance Penalty Methods

Penalty methods provide an alternative approach to constrained optimization problems without the necessity of introducing additional unknowns in the form of Lagrange multipliers. Suppose that we wish to minimize $F: U \rightarrow \mathbb{R}$ subject to the constraint that the minimizers u belong to a convex set $K \subset U$. The idea behind penalty methods is roughly speaking to append to F a penalty functional P which increases in magnitude according to how severely the constraint is violated. In other words, the more a candidate (minimize) $v \in U$ violates the constraint, the greater the penalty we must pay. Let F be a coercive weakly lower semi-continuous functional defined on a reflexive Banach space U and again denoted by K a non-empty closed convex subset of U . We seek minimizers of F in K . The penalty method for this problem consists of introducing a new functional, F_ϵ , depending on a real parameter $\epsilon > 0$, of the form.

$$F_\epsilon(v) = F(v) + \frac{1}{\epsilon} P(v) \quad (17)$$

where $P: U \rightarrow \mathbb{R}$ is a penalty functional satisfying the condition

$$P: U \rightarrow \mathbb{R} \text{ is weakly lower semi-continuous} \quad (18)$$

$$P(v) > 0, p(v) = 0 \text{ iff } v \in K$$

In most instances, we also expect P to satisfy

$$P \text{ is } G\text{-differentiable on } U \quad (19)$$

The functional F is defined on all of U , in view of the properties of F and P , F_ϵ is coercive and weakly lower semi-continuous. Hence, in accordance with theorem

for each $\epsilon > 0$ there exists a $u_\epsilon \in U$ which minimizes F_ϵ

$$\inf_{v \in U} F_\epsilon(v) = F(u_\epsilon) \quad (20)$$

Moreover, if F and P are G -differentiable, u_ϵ is characterized and by

$$\langle DF(u_\epsilon), v \rangle + 1/\epsilon \leq DP(u_\epsilon), v \rangle = 0, \text{ for all } v \in U \quad (21)$$

The question, of course, is whether or not the solution U_ϵ of (20) (or 21) generate to a sequence U_ϵ which converges to a solution u of the original minimization problem. This is easily resolved. Since u_ϵ minimizes F_ϵ

$$F(u_\epsilon) + 1/\epsilon - P(u_\epsilon) \leq F(v) + 1/\epsilon - P(v), \text{ for all } v \in U.$$

If $v \in K$, $P(v) = 0$ and since $(1/\epsilon) p(u_\epsilon) \geq 0$ we have

$$F(u_\epsilon) \leq F(v) \quad (22)$$

If (u_ϵ) denotes a sequence of solutions of (20) obtained as $\epsilon \rightarrow 0$ the fact that F is coercive implies that a constant $C > 0$ exists, independent of ϵ , such that $\|u_\epsilon\| < C$. Since U is reflexive, this guarantee the existence, of a subsequence m also denoted u_ϵ , such that $u_\epsilon \rightarrow u$ weakly in U . Finally, we use the weak lower semi continuity of F and P to obtain

$$\liminf F(u_\epsilon) \geq \liminf F(u_\epsilon) \geq F(u).$$

In summary, we have proved.

3.5 Theorem (3) : Let $F : U \rightarrow \mathbb{R}$ be coercive and weakly lower semicontinuous, K be a non-empty, closed convex subset of the reflexive Banach space U , and $P : U \rightarrow \mathbb{R}$ a penalty functional satisfying (18). Then for every $\epsilon > 0$, there exists a solution $\{ U_\epsilon \}$ to (20), in addition, there exists a subsequence $\{ u_\epsilon \}$ of such solution which converges weakly to $u \in U$, where

$$F(u) \leq \inf F(V)$$

Moreover, if F and P are G -differentiable, then U satisfy (21) for each $\epsilon > 0$ and the limit u satisfies the variational inequality.

$$\langle DF(u), v - u \rangle \geq 0, \text{ for all } v \in K \quad (23)$$

Remarks (3) : Suppose that the penalty functional P satisfies (18) and (19) and that P is the composition of the operators, $P = j \circ B$, $B : U \rightarrow V$, $J : R(B) \rightarrow \mathbb{R}$ where B is the operator defining the constraint set.

CONCLUSION

This paper presents the study of advance method for solution of constraint optimization problem. We discuss different optimization methods Branch & bound method, Distributed constraint optimization, Penalty function method, [Barrier methods](#) & Karush–Kuhn–Tucker (KKT) conditions for equality & inequality constraints.

Saddle points, Lagrange Multipliers and Penalty methods for solving Constrained Optimization problems have been analyzed with different examples. It is observed that

penalty function methods are more acceptable for analyzing problems of almost all types appearing in operational research.

REFERENCES

- [1] Andreani R, Martínez, J. M, Schuverdt ,M. L, On the relation between constant positive linear dependence condition and quasinormality constraint qualification. Journal of Optimization Theory and Applications, vol. 125, no2, 473–485, (2005).
- [2] Avriel, Mordecai Nonlinear Programming: Analysis and Methods. Dover Publishing. [ISBN 0-486-43227-0](#). (2003).
- [3] Boyd .S, Vandenberghe .L, [Convex Optimization](#). Cambridge University Press. [ISBN 0-521-83378-7](#), (2004)
- [4] Nocedal J., S. J. Wright, Numerical Optimization. Springer Publishing. [ISBN 978-0-387-30303-1](#)(1999)
- [5] Dechter, Rina Constraint Processing. Morgan Kaufmann. ISBN 1-55860-890-7 (2003).
- [6] Jeter MIW mathematical programming pure and applied mathematics, Inc. (1986).
- [7] Himm elblau D.M. applied non linear programming Mc-Graw Hill, (1972).