Combinatorial Optimization and Nonlinear Optimization

Abbas Musleh Salman¹*, Dr. Ahmed Sabah Al-Jilawi²
¹² University of Babylon /College of Education for Pure Sciences

Emails: abbas.mmm2019@gmail.com
aljelawy2000@yahoo.com

Abstract: The objective of this paper is to locate a superior method that merges quicker of maximal independent set problem (MIS). One methodology is to the correlation between The Penalty and Augmented Lagrangian Methods. We have likewise built up the hypothetical combination properties of these methods.

Keywords: Optimization, Semidefinite programming, Nonlinear Optimization, Combinatorial, maximal, minimum, Independent set.

1. Introduction:
Optimization implies finding the ideal solution to a specific issue. Mathematically, it implies finding the minimum or maximum value of the objectivist function of n variables [1]. The way toward tackling an optimization problem where a portion of the constraints or the goal functions is nonlinear programming (NLP) in mathematics. An optimization problem is one of calculating of the extreme (maximal, minimal or fixed points) of an objectivist function on a bunch of obscure genuine factors and the condition as per the general inclination of a system of equalities and not equalities, by and large named Restrictions. It is the subfield of mathematical optimization that deals with problems that are nonlinear [3], [5]. "The subfield of mathematical optimization that is Link to operations research, algorithm theory, and computational complexity theory is combinatorial optimization. It has important applications in many fields; include artificial cleverness, E-learning, auction theory [6], hardware engineering", applied mathematics and theoretical computer skills. It does on the area of those optimization problems in which the arrangement of feasible solutions is separated or can be decreased to separated, and in which the goal is to identify the best solution. One of the most common problems is the travelling salesman problem (TSP) [7], [8], [9].

2. Definitions and basic Concepts:
Definition (2.1): [5]
Let \( X \subseteq \mathbb{R}^n \), n, m and p ∈ \( Z^+ \), where \( Z^+ \) is positive integers. Let \( f, g_i \) and \( k_j \) be real valued function on, \( \forall i \in \{1, \ldots, n\} \) and each \( j \in \{1, \ldots, m\} \), with in any least one of \( f, g_i \) and \( k_j \) being nonlinear. A nonlinear minimize issue is an optimization issue of the structure.
Min \[ f(x) \]
Subject to \[ g_i(x) \leq 0, \forall \ i \in \{1, \ldots, n\} \]
\[ k_j(x) = 0, \forall \ j \in \{1, \ldots, m\} \]
\[ x \in X \] 

\[ \text{...... (1)} \]

And the feasible set of (1) is denoted by: \( S = \{x \in X \mid g(x) = 0\} \).

**Example (1):**
Two-dimensional model
A simple issue (Figure (1)). It can be characterized more simply by restrictions.

![Figure (1): Two-dimensional](image-url)

In figure (1) the blue area is the feasible area. The intersection of the line with the feasible area the solution. The line is the best achievable shape line (locus with a given value of the goal function) [4], [5].

\[
\begin{align*}
    z_1 &\geq 0 \\
    z_2 &\geq 0 \\
    z_1^2 + z_2^2 &\geq 1 \\
    z_1^2 + z_2^2 &\leq 2 
\end{align*}
\]

With an goal function to be maximal
\[ f(z_1) = z_1 + z_2 \]
When \( z = (z_1, z_2) \).

Three- dimensional model.

![Figure (2): Three-dimensional](image-url)
The intersection of the top surface with the constrained space in the center the solution. (See the Figure (2)). It can be defined more simply by restrictions.
\[ x_1^2 - x_2^2 + x_3^2 \leq 2 \]
\[ x_1^2 + x_2^2 + x_3^2 \leq 10 \]

With an goal function to be maximal
\[ f(x) = x_1x_2 + x_2x_3 \]
When \( x = (x_1, x_2, x_3) \).

**Definition (2.2): [9]**
Let \( f : \mathbb{R}^n \to \mathbb{R} \) and \( x^* \in \mathbb{R}^n \). Then \( x^* \) is said to be "a local minimizer" of \( f \) if there is a scalar \( t > 0 \) such that \( f(x^*) \leq f(x) \), \( \forall x \in N(x^*, t) = \{ x \in \mathbb{R}^n; \|x - x^*\| \leq t \} \) we call \( x^* \) "a strict local minimizer" of \( f \) if there is a scalar \( t > 0 \) such that \( f(x^*) < f(x) \), \( \forall x \neq x^* \) such that \( x \in N(x^*, t) \).

We call \( x^* \) "a global minimizer" of \( f \) if \( f(x^*) \leq f(x) \) for all \( x \in \mathbb{R}^n \).
Finally, we say that \( x^* \) is "a strict global minimizer" of \( f \) if \( f(x^*) < f(x) \), \( \forall x \neq x^* \).

**Definition (2.3): [3]**
Let \( S \subseteq \mathbb{R}^n \). If the line segment between any two points in \( S \) lies in \( S \), i.e.
\[ \partial x_1 + (1 - \partial)x_2 \in S, \quad \forall x_1, x_2 \in S, 0 \leq \partial \leq 1 \]

Then \( S \) is said to be convex. It can be indicated that a set \( S \subseteq \mathbb{R}^n \) is convex iff for any \( x_1, \ldots, x_n \in S \), the convex combination \( \sum_{i=0}^{n} \partial_i x_i \)
Where \( \sum_{i=0}^{n} \partial_i = 1 \), and \( \partial_i \geq 0, i = \{1, \ldots, n\} \in S \)

See the figure (3):

![Convex set vs Non-convex set](image)

**Figure (3): convex set and nonconvex set**

**Definition (2.4): [3]**
Let \( S \subseteq \mathbb{R}^n \) be a convex set and \( S \neq \emptyset \). If \( f : S \to \mathbb{R} \) satisfies:
\[ f(\partial x_1 + (1 - \partial)x_2) \leq f(x_1) + (1 - \partial)f(x_2), \forall x_1, x_2 \in S, 0 \leq \partial \leq 1 \]

Then \( f \) is said to be "a convex function " on \( S \). If the inequality above is true as a strict inequality for all\( x_1 \neq x_2 \) and for all \( 0 < \partial < 1 \), then \( f \) we say is "a strictly convex function" on \( S \). If there is a constant \( a > 0; \forall x_1, x_2 \in S \) and for all \( 0 \leq \alpha \leq 1 \)
\[ f(\partial x_1 + (1 - \partial)x_2) \leq f(x_1) + (1 - \partial)f(x_2) - \frac{\alpha}{2} a(1 - \partial)\|x_1 - x_2\|^2 \]

--------

3
Definition (2.5): [2]
The matrices encountered most frequently in numerical optimization are symmetric. Let $\mathcal{B} = \{a_{ij}\}$ be a square matrix. The matrix $\mathcal{B}$ is said to be "symmetric" i.e. $a_{ij} = a_{ji}$, if $\mathcal{B}^T = \mathcal{B}$.

Example (2):
Let $\mathcal{B}$ be any matrix such that:

\[
\mathcal{B} = \begin{bmatrix}
3 & 1 & 0 \\
1 & 4 & 2 \\
0 & 2 & 5
\end{bmatrix}
\]

Then $\mathcal{B}^T = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 4 & 2 \\ 0 & 2 & 5 \end{bmatrix} = \mathcal{B}$.

### 3. Semidefinite Programming: [10]

The standard general semidefinite programming (SDP) problem is given by:

\[
\text{min } \langle c, x \rangle \\
\text{subject to } A_i x \geq b_i \text{, } \forall i = 1, ..., m, \\
x \geq 0 \text{, } x \in s^n
\]

---

The dual problem semidefinite programming of SDP problem is given by:

\[
\text{max } \langle b, y \rangle \\
\text{subject to } \sum_{i=1}^{m} y_i A_i + S = C \\
S \geq 0
\]

We define the optimal values of the primal and it's dual as:

\[
p^* = \inf \{ \langle c, x \rangle; (A_i, x) \geq b_i, \forall i = 1, ..., m \text{ and } x \geq 0 \}
\]

\[
d^* = \sup \{ \langle b, y \rangle; \sum_{i=1}^{m} y_i A_i + S = C \text{ and } S \geq 0 \}.
\]

---

Example (3):

\[
A_1 = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 3 & 7 \\ 1 & 7 & 5 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 2 & 8 \\ 2 & 6 & 0 \\ 8 & 0 & 4 \end{bmatrix}, b = \begin{bmatrix} 11 \\ 9 \end{bmatrix}, \text{ and } C = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 9 & 0 \\ 3 & 0 & 7 \end{bmatrix}
\]

The variable $X$ will be the $3 \times 3$ symmetric matrix:

\[
X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\
           x_{21} & x_{22} & x_{23} \\
           x_{31} & x_{32} & x_{33} \end{bmatrix}
\]

Find (SDP) and (SDD).
Solution:

SDP: \[
\begin{align*}
\text{Min} & \quad x_{11} + 4x_{12} + 6x_{13} + 9x_{22} + 0x_{23} + 7x_{33} \\
\text{Subject to} & \quad x_{11} + 0x_{12} + 2x_{13} + 3x_{22} + 14x_{23} + 5x_{33} = 11 \\
& \quad 0x_{11} + 4x_{12} + 16x_{13} + 6x_{22} + 0x_{23} + 4x_{33} = 19 \\
& \quad x_{21} x_{22} x_{23} + x_{31} x_{32} x_{33} \geq 0 \\
\end{align*}
\]

SDD: \[
\begin{align*}
\text{Max} & \quad 11y_1 + 19y_2 \\
\text{Subject to} & \quad \begin{bmatrix} 1 & 0 & 3 \\ 0 & 3 & 7 \\ 1 & 7 & 5 \end{bmatrix} + y_2 \begin{bmatrix} 0 & 2 & 8 \\ 2 & 6 & 0 \\ 8 & 0 & 4 \end{bmatrix} + S = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 9 & 0 \\ 3 & 0 & 7 \end{bmatrix} \\
& \quad S \geq 0 \\
\end{align*}
\]

SDD: Max \[
\begin{align*}
11y_1 + 19y_2 \\
\begin{bmatrix} 1 & -1 & y_1 - 0y_2 \\ 2 & -0 & y_1 - 2y_2 \\ 3 & -1 & y_1 - 3y_2 \end{bmatrix} \begin{bmatrix} 2 & -2 & y_2 \\ 9 & 6 & y_2 \\ 0 & 7 & y_2 \end{bmatrix} \geq 0 \\
\end{align*}
\]

4. Basic concept of (SDP) and (SDD):

(4.1) Theorem (Weak duality):[11]
If \( n \) is feasible solution for SDP and \( m \) is feasible solution for SDD, then \( \langle c, n \rangle \geq \langle b, m \rangle \)

Proof:
\[
\langle c, n \rangle - \langle b, m \rangle = \langle A^*m + s, n \rangle - \langle b, m \rangle \\
= \langle A^*m, n \rangle + \langle s, n \rangle - \langle b, m \rangle \\
= \langle An, m \rangle + \langle s, n \rangle - \langle b, m \rangle \\
= \langle b, m \rangle + \langle s, n \rangle - \langle b, m \rangle \\
= \langle s, n \rangle \\
\]
Where \( \langle s, n \rangle \geq 0 \). It follows \( \langle c, n \rangle - \langle b, m \rangle \geq 0 \)
Implies \( \langle c, n \rangle \geq \langle b, m \rangle \).

(4.2) Theorem (Strong duality for the primal):
If there exists \( x_0 \) strictly feasible for (SDP), then there exists an optimal Solution \( y^* \) of (SDD) and \( (b, y^*) = p^* \).

(4.3) Theorem (Strong duality for the dual):
If there exists \( y_0 \) strictly feasible for (SDD), then there exists an optimal solution \( x^* \) for (SDP) and \( (c, x^*) = d^* \).

(4.4) Theorem (Primal and dual strong duality):
If there exists \( x_0 \) which is feasible for (SDP) and \( (y_0, S_0) \) which is feasible for (SDD) such that \( x_0 > 0 \) and \( S_0 > 0 \), then there exists an optimal solution \( x \) for (SDP) and an optimal solution \( (y, S) \) for (SDD) and \( \langle c, x^* \rangle - \langle b, y^* \rangle = \langle S^*, x^* \rangle = 0 \).

5. Maximal independent set problem (MIS):
In graph theory, all a set of vertices, no two of which are adjacent is called an independent set (IS) [11]. "Equivalently, and all edge in the graph has at most one endpoint in \( S \), and the size of an independent set is the number of vertices it included, and independent sets is called internally stable sets". Also in graph theory, very independent set that is not a subset of any other independent set is called a maximal independent set (MIS) as in the figure (5). And shown in example (4) simplified battery. In the second sense, there is not a vertex outside the independent set that may correlate it
because it is maximal With regard to the characteristics of the independent set. Although the problem of the maximum independent sets is one of the problems of NP-hard, it is difficult to find a definitive solution to it. Therefore, an important question that should be verified is, what is the usefulness of knowledge from the critical independent groups problem in the graph in solving the maximum independent group problem in this graph?

**Definition (6.1): [7]**

In a graph $G = (V, E)$ any independent set $S$ is called "a maximal independent set" if for $v \in V$, one of coming up next is true:

- $v \in S$
- $N(v) \cap S \neq 0$ where $N(v)$ denotes the neighborhood of $v$

"The above can be formulated as either a vertex that belongs to an independent set or that has at least one neighbor vertex that belongs to the independent set. Therefore, each edge contains at least one endpoint in the set $S$. However, it is not possible that each edge in the set has at least one endpoint or even one endpoint in $S".$

![Figure (5): the nine blue vertices form a maximum independent set](image)

**Example (4):**

From the figure below, find (IS) and (MIS)

![Figure (6): The graph $G = (V, E)$](image)

**Solution:**

The independent set (IS) is:

$x_1 = \{E, B\}, x_2 = \{F, B\}, x_3 = \{F, A\}, x_4 = \{E, A\}, x_5 = \{E, B, F\}, x_6 = \{E, B, F, A\}$

Hence the maximum independent set is $x_6 = \{E, B, F, A\}$

In the maximal independent set (MIS) problem, we are given the graph $G = (V, E)$ with vertex weight $x_i$, and the goal is to maximal the total weight of the vertices in an independent set (Such as the $S$ set of vertices having no two vertices at $E$); This problem can be mentioned as follows:

\[
\text{(MIS) Max: } \sum_{i=1}^{n} x_i y_i \\
\text{Subject to: } x_i x_j = 0 \text{ for all } (i, j) \in E \\
x \in \{0,1\}^n.
\]
6. NP- hardness: [11]
- P represents the set of problems that can be answered with a yes or no and that can be solved in polynomial time.
- NP represents a set of problems that can be answered with yes or no with the following property: If the answer is: Yes, the proof of this property can be verified in polynomial time.
- NP-hard is one of the important and difficult problems that cannot be resolved in a definitive way.
- NP-complete problems are from the hardest problems in NP.

Take for example of an NP-hard problem is the optimization problem of finding the least cost cyclic route through each vertices of the graph. This is usually known as a problem as the traveling salesman problem [13]. See the figure (7).

7. Approximation Methods:

(7.1) Augmented Lagrangian Method:[2]
This method began to be used in the 1970s. "At first it was called the multiples method or double method. The aim of this method is to solve the restricted optimization problems. And this is done by substitute a restricted problem with a series of unrestricted problems" [5,7,11]. "The augmented Lagrangian method is undifferentiated from to the Penalty method since in both of them a Penalty term is added to the aim. The only difference in the augmented Lagrangian method is that the double term Lagrange multiplier term is added to it [2,4,11]". The augmented Lagrangian method it was introduced by Hestenes .
The general form for the augmented Lagrangian is:

\[ \mathcal{L}_\alpha(x, y) = f(x) + \langle y, b_i - \langle a_i, x \rangle \rangle + \frac{1}{2\alpha} \| b_i - \langle a_i, x \rangle \|^2 \]  

\\ 

\text{Represents the primal (P) and the dual (D) general linear programming problem:} 

\begin{align*}
(P) & \left\{ \begin{array}{l}
\text{maximize} \\
\text{subjectto}
\end{array} \right. \\
& \langle c, x \rangle \\
& Ax = b \\
& x \geq 0 \\
\end{align*}

and

\begin{align*}
(D) & \left\{ \begin{array}{l}
\text{minimize} \\
\text{subjectto}
\end{array} \right. \\
& \langle b, y \rangle \\
& A^Ty \geq c
\end{align*}

The double method is to update sometimes the Penalty parameter $\gamma$ and the Lagrange Multiplier estimate [11] $\theta$ in all iteration. The method of the double is summarized in the Algorithm [1].

**Algorithm [1] : The Augmented Lagrangian Method**

1) Choose $\delta^0$, and $\gamma^0 > 0$, choose $\theta^0$.
2) Find $\delta^{m+1}$ as

\[ \delta^{m+1} = \arg\min_{\delta} F(\delta, \theta^k, y^m). \]

\[ \text{........... (13)} \]

3) Update $\gamma^k$ and $\theta^k$.
4) Set $m = m + 1$ Then repeat.

(7.2) The Penalty Function Methods:

Penalty method is a certain class of algorithms for solving restricted optimization problems. A penalty method is an approach that uses a sequence of unconstrained problems. This method replaces a restricted problem with unconstrained problems by adding a penalty term to the objective function [5, 7, and 11]. In 1943, the first penalty function was studied by Courant [9]. However, sequential unconstrained minimization approaches began to be widely used for solving optimization problems in the 1960s (see [1]), and such approaches were applied to get the solution of the original constrained problems. Following [13], we consider the standard penalty function (The squared penalty function) of problem (1) given by the figure (10).
We consider the standard penalty function (The squared penalty function) of problem (1) given by:

\[ p_q(x, \beta) = f(x) + \frac{1}{2\beta} \|g(x)\|^2 \]  

(14)

Where \( \beta > 0 \). Let \( x(\beta) \) be a minimizer of \( p_q(x, \beta) \) for \( \beta > 0 \), the penalty methods produce about her a sequence of infeasible solutions. Each iterate \( x(\beta^k) \) is either necessarily infeasible or a local optimal solution of problem (1).

8. Numerical Results:

Is we can see in Figures (11 and 12); we choose the problem (g05_60.5) as in Table (1), we had the result of the two methods augmented Lagrangian method and Penalty method. In the Tables (2, 3) it was shown to us that the augmented Lagrangian method is better and more accurate than the Penalty method, even if we take mixed, this means that the enhanced augmented Lagrangian method. Its performance was better and faster convergence to find the best solutions.

![Figure (10) The Penalty Function](image-url)
Figure (12): Comparison of the penalty and augmented Lagrangian test method

Table (1) Test results between the two methods the penalty and augmented Lagrangian method

| Graphs    | Penalty fcalls | Augmented fcalls | Mixed fcalls |
|-----------|----------------|------------------|--------------|
| g05_60.0  | 704            | 500              | 410          |
| g05_60.1  | 750            | 314              | 422          |
| g05_60.2  | 510            | 463              | 345          |
| g05_60.3  | 891            | 850              | 524          |
| g05_60.4  | 550            | 490              | 330          |
| g05_60.5  | 902            | 443              | 465          |
| g05_60.6  | 614            | 459              | 379          |
| g05_60.7  | 851            | 415              | 472          |
| g05_60.8  | 662            | 250              | 115          |
| g05_60.9  | 515            | 269              | 394          |

Table (2) Test results between the two methods the penalty and augmented Lagrangian method

| Graphs    | Penalty fcalls | Augmented fcalls | Mixed fcalls |
|-----------|----------------|------------------|--------------|
| g05_80.0  | 566            | 319              | 261          |
| g05_80.1  | 943            | 489              | 375          |
| g05_80.2  | 623            | 292              | 333          |
| g05_80.3  | 421            | 165              | 142          |
| g05_80.4  | 666            | 490              | 330          |
| g05_80.5  | 323            | 185              | 200          |
| g05_80.6  | 459            | 256              | 238          |
| g05_80.7  | 353            | 311              | 262          |
| g05_80.8  | 554            | 334              | 181          |
| g05_80.9  | 435            | 316              | 290          |
Table (3) Test results between the two methods the penalty and augmented Lagrangian method

| Graphs      | Penalty fcalls | Augmented fcalls | Mixed fcalls |
|------------|----------------|------------------|--------------|
| g05_100.0  | 401            | 150              | 200          |
| g05_100.1  | 354            | 201              | 197          |
| g05_100.2  | 517            | 450              | 285          |
| g05_100.3  | 371            | 193              | 170          |
| g05_100.4  | 409            | 240              | 270          |
| g05_100.5  | 502            | 202              | 179          |
| g05_100.6  | 441            | 391              | 256          |
| g05_100.7  | 626            | 472              | 391          |
| g05_100.8  | 549            | 390              | 401          |
| g05_100.9  | 375            | 300              | 289          |

9. Conclusion:
The objective of the paper was accomplished, and the following points were explained:
- We tested the comparison and difference between two methods, the penalty method and the Lagrange augmented method.
- We demonstrate the properties of hypothetical approximation and also we studied algorithms of the two methods.
- We also used the graphs provided in the Big Mac library to evaluate the methods. Also, these figures included various features and properties with no small number of edges and vertices.

References:
[1] A. Fiacco and G. McCormick, Sequential unconstrained minimization techniques for nonlinear programming, A primal dual method. Management Sci, 10 (1968).
[2] Angelo Miele, EE Cragg, RR Iyer, and AV Levy. Use of the augmented penalty function in mathematical programming problems, part 1. Journal of Optimization Theory and Applications, 8(2):115–130, 1971.
[3] Andrzej Ruszczynski. Nonlinear optimization. Princeton university press, 2011.
[4] Christian Kanzow, Daniel Steck, and Daniel Wachsmuth. An augmented lagrangian method for optimization problems in banach spaces. SIAM Journal on Control and Optimization, 56(1):272–291, 2018.
[5] Dimitri P Bertsekas and Athena Scientific. Convex optimization algorithms. Athena Scientific Belmont, 2015.
[6] D. P. Bertsekas, Nonlinear programming. Athena scientific Belmont, 1999.
[7] Judah Ben Rosen, Olvi L Mangasarian, and Klaus Ritter. Nonlinear Programming: Proceedings of a Symposium Conducted by the Mathematics Research Center, the University of Wisconsin, Madison, May 4-6, 1970. Number 25. Elsevier, 2014.
[8] JB Hiriart-Urruty and C Lemaréchal. Convex analysis and minimization algorithms. no. 305-306 in grund. der math. wiss, 1993.

[9] M. M. El-Alem, A global convergence theory for a class of trust region algorithms for constrained optimization, PhD thesis, Rice University, 1988.

[10] Michael James Lowe. Nonlinear programming: augmented Lagrangian techniques for constrained minimization. PhD thesis, Montana State University-Bozeman, College of Engineering, 1974.

[11] P Dimitri Bertsekas. Nonlinear programming. 1999. Athena, Scientific, Belmont CA.