

# Harnessing Karush-Kuhn-Tucker (KKT) as Optimality Conditions for Solving Nonlinear Constrained Optimization Problems

Agu, C.M., Unaegbu, E.N., Chikwendu, C.R.

Nnamdi Azikiwe University, Awka.

DOI : <https://doi.org/10.51584/IJRIAS.2024.907043>

Received: 14 June 2024; Revised: 04 July 2024; Accepted: 09 July 2024; Published: 14 August 2024

## ABSTRACT

This research work succinctly investigated the solution of Nonlinear Optimization Problems using Karush-Kuhn-Tucker (KKT) as optimality condition. The historical development of nonlinear optimization was discussed. The nonlinear constrained optimization problems with inequality constraints were solved with the method of Karush-Kuhn-Tucker (KKT).

## INTRODUCTION

Optimization is the act of obtaining the best result under given circumstances. It is also finding solutions from set of admissible or feasible solutions that minimizes (or maximizes) a performance measure or objective, Asim Karim (2003). The technique allows comparison of the different choices for determining which decision might be best. The goal of all such decisions is either to minimize the effort required or maximize the desired benefit. Since the effort required or the benefit desired in any practical situation can be expressed as a function of certain decision variables, optimization can be defined as the process of finding the conditions that give the maximum or minimum value of a function,  $f(x)$ . According to Shang (1997), optimization problems are made up of three components: a set of unknowns (variables), an objective function to be minimized or maximized and a set of constraints that specify feasible values of the variables. These sets of unknown variables are called non-negativity constraints or decision variables.

In Xue Jin (2021), the direct idea is to use mathematical tools such as differential calculus, variational method and Lagrange multiplier method to obtain the solution expression of the problem through logical derivation and analysis. Hence, we succinctly investigated classical optimization methods which are analytical in nature and make use of differential calculus to find optimal value for both unconstrained and constrained objective functions. We shall devote more time on the constrained optimization problems with equality and inequality constraints.

According to Freund (2014), most of the theoretical and computational interest in nonlinear optimization has taken place since 1947. However, it is useful to note that nonlinear optimization was first studied as early as the 1600s. Indeed, both Fermat and Newton studied the 1-dimensional nonlinear optimization problem:

$$\max f(x) \text{ where } x \in R$$

and Newton developed the classical optimality condition:

$$\frac{df(x)}{dx} = 0$$

This was generalized by Euler to multivariable nonlinear optimization:

$$\min f(x_1, \dots, x_n);$$

with the optimality condition:

$$\nabla f(x) = 0$$

In the late 1700s both Euler and Lagrange examined problems in infinite dimensions and developed the calculus of variations. The study of optimization received very little attention thereafter until 1947, when modern linear optimization was developed by Dantzig. In the late 1940s and early 1950s Kuhn and Tucher (preceded by Karush) developed general optimization conditions for nonlinear constrained optimization. At that time, application of nonlinear optimization included problems in chemical engineering, portfolio optimization and a host of decision problems in management and industrial operations. The 1960s saw the development of larger scale nonlinear optimization coincident with the expanding capabilities of computers.

According to Brain Wallace (2004), new possibilities in optimization methods were developed with Karmarkar's interior-point method for linear optimization in 1984, which was extended to nonlinear convex optimization by Nesterov and Nemirovskii in 1994. In 1990s the important fields of semi definite optimization, conic optimization, and "robust optimization" were developed. Since 2000 there has been great interest and progress in "huge scale" nonlinear optimization, which considers problems in thousands or even millions of decision variables and constraints.

Note that Kuhn-Tucker (KT) Theorem was called Karush-Kuhn-Tucker (KKT) Theorem in recognition of the fact that in 1939 William Karush produced the same result of the KT Theorem in his Master of Science degree thesis at the mathematics department of the University of Chicago, according to Richard W. Cottle (2012)

Ejikeme C.L et al (2015) highlighted that Nonlinear Programming is of two types, viz: single-variable optimization and multivariable optimization. Our interest is to study solutions of Nonlinear Optimization Problems which can still be referred to as Nonlinear Programming Problems (NLPP's). There are many methods of solving problems which are in the form of constrained multivariable optimization with equality constraints but in this paper, we investigate Karush-Kuhn-Tucker (KKT) Method that deals with the constrained multivariable optimization with inequality constraints.

## Preliminaries

Karush-Kuhn-Tucker (KKT) conditions are first derivatives test for a solution in nonlinear programming to be optimal. With the inequality constraints the KKT generalizes the method of Lagrange multipliers which allows only equality constraints. According to Mikulas Luptacik (2010), the Kuhn-Tucker conditions are the natural generalization of the Lagrange multiplier approach, from classical differential calculus replacing equality constraints by inequality constraints, to take account of the possibility that the maximum or minimum in question can occur not only at a boundary point but also at an interior point.

As stated by Scott Moura (2014), let us consider the general constrained optimization problem.

$$\min f(x)$$

subject to

$$g_i(x) \leq 0, i = 1, \dots, m$$

$$h_j(x) = 0, j = 1, \dots, l$$

Introduced the "Lagrange multipliers"  $\lambda_j, j = 1, \dots, l$  each associate with equality constraints  $h_j(x), j = 1, \dots, l$  and  $\mu_i, i = 1, \dots, m$ . then we can augment the cost function to form the "Lagrangian"  $L(x)$  as follows

$$L(x) = f(x) + \sum_{i=1}^m \mu_i g_i(x) + \sum_{j=1}^l \lambda_j h_j(x)$$

$$= f(x) + \mu^T g(x) + \lambda^T h(x) \tag{1}$$

As before, when the equality constraints are satisfied,  $h(x) = 0$  then the third term becomes zero. Elements of the second term becomes zero in the two cases:

1. An inequality constraint is active, that is  $g_i(x) = 0$ .
2. The Lagrange multiplier  $\mu_i = 0$

Consequently, the Lagrangian  $L(x)$  can be constructed to have identical values of the cost function  $f(x)$  of the conditions are applied. This motivates the first-order necessary conditions (FONC) for the general constrained optimization problem called the Karush Kuhn Tucker (KKT) conditions.

**Proposition 2.1.** (KKT Conditions) If  $x^*$  is a local minimum, then the following necessary conditions hold:

$$L(x) = \frac{\partial f}{\partial x}(x^*) + \sum_{i=1}^m \mu_i \frac{\partial}{\partial x} g_i(x^*) + \sum_{j=1}^l \lambda_j \frac{\partial}{\partial x} h_j(x^*) = 0 \tag{Stationality}$$

$$g_i(x^*) \leq 0 \quad i = 1, \dots, m \tag{Feasibility}$$

$$h_j(x^*) = 0 \quad j = 1, \dots, l \tag{Feasibility}$$

$$\mu_i \geq 0 \quad i = 1, \dots, m \tag{Non-Negativity}$$

$$\mu_i g_i(x^*) = 0 \quad i = 1, \dots, m \tag{Complementary Slackness}$$

which can also be written in the matrix- vector form as

$$\frac{\partial f}{\partial x}(x^*) + \mu^T \frac{\partial}{\partial x} g(x^*) + \lambda^T \frac{\partial}{\partial x} h(x^*) = 0 \tag{Stationary}$$

$$g(x^*) \leq 0 \tag{Feasibility}$$

$$h(x^*) = 0 \tag{Feasibility}$$

$$\mu \geq 0 \tag{Non-Negativity}$$

$$\mu^T g(x^*) = 0 \tag{Complementary Slackness}$$

## KARUSH-KUHN-TUCKER METHOD

The Karush Kuhn Tucker method of solving nonlinear optimization problem uses the necessary and sufficient conditions for a local optimum.

### Kuhn Tucker Necessary Condition

According to J. K Sharma, 2009, we have

$$\text{Optimize } Z = f(x) \tag{2}$$

subject to the constraints

$$g_i(x) = h_j(x) - b_i \leq 0; \quad i = 1, 2, \dots, n. \quad j = 1, 2, \dots, l \tag{3}$$

where  $x = (x_1, x_2, \dots, x_n)$

Add non-negative slack variables  $s_i (i = 1, 2, \dots, m)$  in each of the constrained to convert them to equality constraints. The problem can then be restated as:

$$\text{Optimize } Z = f(x) \tag{4}$$

subject to the constraints

$$g_i(x) + s_i^2 = 0, \quad i = 1, 2, \dots, m \tag{5}$$

The  $s_i^2$  has only be added to ensure non-negative value (feasibility requirement) of  $s_i$  and to avoid adding  $s_i \leq 0$  as an additional side constraint. Therefore, we now have constrained multivariable optimization problem with equality constraints with  $n + m$  variable. Thus, it can be solved by using the Lagrangian multiplier method. Hence, we form the Lagrangian function:

$$L(x, s, \lambda) = f(x) - \sum_{i=1}^m \lambda_i [g_i(x) + s_i^2] \tag{6}$$

where  $\lambda_i = (\lambda_1, \lambda_2, \dots, \lambda_m)^T$  is the vector of Lagrange multiplier.

The necessary conditions for an extreme point to be local optimum can be obtained by solving the following equations,

$$\frac{\partial L}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial g_i(x)}{\partial x_j} = 0 \quad i = 1, \dots, m \tag{7}$$

$$\frac{\partial L}{\partial \lambda_i} = -[g_i(x) + s_i^2] = 0 \quad i = 1, \dots, m \tag{8}$$

$$\frac{\partial L}{\partial s_i} = -2s_i \lambda_i = 0 \quad i = 1, \dots, m \tag{9}$$

The eqn (8) gives us back the original set of constraints:  $g_i(x) + s_i^2 = 0$ . If a constraint is satisfied with equality sign,  $g_i(x) = 0$  at the optimum point  $x$ , then it is called an active (binding or light) constraint, otherwise it is known as inactive (slack) constraint. The equation (9) provides us the set of rules:  $-2\lambda_i s_i = 0$  or  $\lambda_i s_i = 0$  for finding the unconstrained optimum. The condition  $\lambda_i s_i = 0$  implies that either  $\lambda_i = 0$  or  $s_i = 0$ . If  $s_i = 0$  and  $\lambda_i > 0$ , then eqn (8) gives  $g_i(x) = 0$ . This means either  $\lambda_i = 0$  or  $g_i(x) = 0$ , and therefore we may also write  $\lambda_i g_i(x) = 0$ . Since  $s_i^2$  has been taken to be a non-negative ( $\geq$ ) slack variable, therefore  $g_i(x) < 0$ ,  $\lambda_i = 0$  and when  $g_i(x) = 0$ ,  $\lambda_i > 0$ . However,  $\lambda_i$  is unrestricted in sign corresponding to  $g_i(x) = 0$ . But if  $\lambda_i = 0$  and  $s_i^2 > 0$ , then the  $i$ th constraint is inactive. (i.e., this constraint will not change the optimum value of  $Z$  because  $\lambda = \frac{\partial Z}{\partial b_i} = 0$  and hence can be discarded). Thus, the Kuhn-Tucker necessary conditions (When active constraints are known) to be satisfied at a local optimum (Maximum or Minimum) point can be stated as follows:

$$\frac{\partial F}{\partial x_j} - \sum_{i=1}^n \lambda_i \frac{\partial g_i}{\partial x_j} = 0 \quad i = 1, \dots, n \tag{10}$$

$$\lambda_i g_i(x) = 0 \tag{11}$$

$$g_i(x) \leq 0 \tag{12}$$

$$\lambda_i \geq 0 \quad i = 1, \dots, m \tag{13}$$

### Kuhn Tucker Sufficient Conditions

J. K Sharma (2009) stated the following theorem as the Kuhn Tucker Sufficient conditions for solving nonlinear optimization problem.

**Theorem 3.3.1 (Sufficiency of Kuhn Tucker Conditions).** The Kuhn Tucker necessary conditions for the problem

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

subject to the constraints

$$g_i(x) \leq 0, i = 1, 2, \dots, m$$

are also sufficient conditions if  $f(x)$  is concave and all  $g_i(x)$  are convex functions of  $x$

**Proof.** The Lagrangian function of the problem

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

subject to the constraints

$$g_i(x) \leq 0, \quad i = 1, 2, \dots, m$$

can be written as:

$$L(x, s, \lambda) = f(x) - \sum_{i=1}^m \lambda_i [g_i(x) + s_i^2] \quad (14)$$

If  $\lambda_i \geq 0$ , then  $\lambda_i g_i(x)$  is convex and  $-\lambda_i g_i(x)$  is concave. Further, since  $\lambda_i s_i = 0$ , we get  $g_i(x) + s_i^2 = 0$ . Thus, it follows that  $L(x, s, \lambda)$  is a concave function. We have derived that necessary condition for  $f(x)$  to be a relative maximum at an extreme point is that  $L(x, s, \lambda)$  also have the same extreme point. However, if  $L(x, s, \lambda)$  is concave, its first derivative must be zero at one point and obviously this point must be an absolute maximum for  $f(x)$ .

## Examples

### Using Karush Kuhn Tucker method to solve optimization problems.

Problem 4.1

$$\text{Maximize } U = xy \quad (15)$$

Subject to

$$100 \geq x + y \quad (16)$$

and

$$x \leq 40 \quad (17)$$

Solution

The Lagrange is

$$L(x, y; \lambda_1; \lambda_2) = xy + \lambda_1(100 - x - y) + \lambda_2(40 - x)$$

and the KKT conditions become

$$L_x = y - \lambda_1 - \lambda_2 = 0; \quad x \geq 0$$

$$L_y = x - \lambda_1 = 0; \quad y \geq 0$$

$$L_{\lambda_1} = 100 - x - y \geq 0; \quad \lambda_1 \geq 0$$

$$L_{\lambda_2} = 40 - x \geq 0; \quad \lambda_2 \geq 0$$

Which gives us four equations with four unknowns:  $x, y, \lambda_1, \lambda_2$ . To solve, we typically approach the problem in a stepwise manner. First, ask if any  $\lambda_i$  could be zero. Try  $\lambda_2 = 0$  ( $\lambda_1 = 0$  does not make sense, given the form of the utility function), then

$$x - \lambda_1 = y - \lambda_1$$

$$x = y$$

from the constraint  $100 - x - y$  we have

$$x = y = 50$$

which violates our constraint  $x \leq 40$ . Therefore,

$$x = 40 \text{ and } y = 60$$

also

$$\lambda_1 = 40 \text{ and } \lambda_2 = 20$$

then the maximum value of the function

$$U(40, 60) = 40 \times 60 = 2400$$

#### Problem 4.2

$$\text{Maximize } f(x, y) = xy \tag{18}$$

Subject to

$$x + y^2 \leq 2 \tag{19}$$

and

$$x, y \geq 0 \tag{20}$$

Solution

Note that the feasible region is bounded, so a global maximum must exist: a continuous function on a closed and bounded set has a maximum there. We write the constraints as

$$g_1(x, y) = x + y^2 \leq 2,$$

$$g_2(x, y) = -x \leq 0$$

$$g_3(x, y) = -y \leq 0$$

Thus, the KKT conditions can be written as

$$y - \lambda_1 + \lambda_2 = 0$$

$$x - 2y\lambda_1 + \lambda_3 = 0$$

$$\lambda_1(2 - x - y^2) = 0$$

$$\lambda_2 x = 0$$

$$\lambda_3 y = 0$$

$$x + y^2 \leq 2$$

$$x, y, \lambda_1, \lambda_2, \lambda_3 \geq 0$$

In each of the “complementary slackness” equations  $\lambda_i(b_i - g_i(x_1, \dots, x_n)) = 0$ , at least one of the two factors must be 0. With  $n$  such conditions, there would potentially be  $2^n$  possible cases to consider. However, with some thought we might be able to reduce that considerably.

**Case 1:** Suppose  $\lambda_1 = 0$ , then the first KKT condition says  $y + \lambda_2 = 0$  and the second says  $x + \lambda_3 = 0$ . Since each term is nonnegative, the only way that can happen is if  $x = y = \lambda_2 = \lambda_3 = 0$ . Indeed, KKT conditions are satisfied when

$$x = y = \lambda_1 = \lambda_2 = \lambda_3 = 0$$

(although clearly this is not a local maximum since  $f(0,0) = 0$  while  $f(x, y) > 0$  at points in the interior of the feasible region).

**Case 2:** Suppose  $x + y^2 = 2$ . Now at least one of  $x = 2 - y^2$  and  $y$  must be positive.

**Case 2a:** Suppose  $x > 0$ , then  $\lambda_2 = 0$ . The first KKT condition says  $\lambda_1 = y$ . The second KKT condition then says  $x - 2y\lambda_1 + \lambda_3 = 2 - 3y^2 + \lambda_3 = 0$ , so

$$3y^2 = 2 + \lambda_3 > 0 \text{ and } \lambda_3 = 0$$

Thus

$$y = \sqrt{\frac{2}{3}} \text{ and } x = 2 - \frac{2}{3} = \frac{4}{3}$$

Again, all the KKT conditions are satisfied.

**Case 2b:** Suppose  $x = 0$ , i.e.  $y = \sqrt{2}$ . Since  $y > 0$ , we have  $\lambda_3 = 0$ . From the second KKT condition we must have  $\lambda_1 = 0$ . But that take us back to **Case 1**. We conclude that there are only two conditions for local maximum:  $(0,0)$  and  $(\frac{4}{3}, \sqrt{\frac{2}{3}})$ . The global maximum is at  $(\frac{4}{3}, \sqrt{\frac{2}{3}})$ .

### Problem 4.3

Minimize  $f(x, y) = (x - 3)^2 + (y - 2)^2$  (21)

Subject to

$$x^3 - y - 3 \leq 0 \tag{22}$$

$$y - 1 \leq 0 \tag{23}$$

$$-x \leq 0 \tag{24}$$

Solution

The objective function and the inequality constraints are:

$$f(x, y) = (x - 3)^2 + (y - 2)^2$$

$$g_1(x, y) = x^2 - y - 3$$

$$g_2(x, y) = x - 1$$

$$g_3(x, y) = -x$$

The figure below illustrates the feasible region. The problem then, is to find a point in the feasible region with the smallest possible value of  $(x - 3)^2 + (y - 2)^2$ . Note that point  $(x, y)$  with  $(x - 3)^2 + (y - 2)^2 = c$  are circles with radius  $\sqrt{c}$  and center  $(3, 2)$ . This circle is called the contour of the objective function having the value  $c$ . To minimize  $c$ , we must find the circle with the smallest radius that intersects the feasible region. As shown in the figure below, the smallest circle corresponds to  $c = 2$  and intersects the feasible region at the point  $(2, 1)$ . Hence the optimal solution occurs at the point  $(2, 1)$  and has objective value equal to 2. The graphical approach is the smallest objective value that intersect the feasible region, is only suitable for small problems, it becomes intractable for problems containing more than three variables, as well as for problem having complicated objective and/or constraint function. This is from Benoit. C. Chachuat (2007)

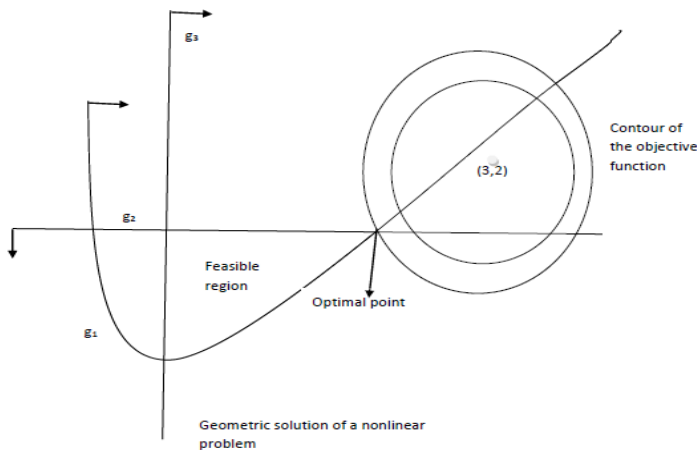


Figure 1.1

## CONCLUSION

We have established that the methods of Karush-Kuhn-Turker are used in making best choice in economic decisions and other vital areas of life. The nonlinear constrained optimization problems with inequality constraints are solved with the method of Karush-Kuhn-Tucker. The nonlinear program must include at least one nonlinear function, which could be the objective function or some or all the constraints. We have also seen that it is necessary to consider optimization approach when it comes to profit maximization and cost minimization. Moreover, Karush-Kuhn-Turker is a useful mathematical tool in dealing with economic and other issues of life.

## REFERENCES

1. Asim Karim. Nonlinear Optimization Theory and Practice. Computer Science Department, Lahoren University of Management Sciences. Presented at the 2nd International Bhurban Conference on Applied Sciences and Technology Control and Technology, Pg. 1-56. June 19-21, 2003.
2. Robert M. Freund. Introduction to Nonlinear Optimization; and Optimality Conditions for Unconstrained Optimization Problems. Massachusetts Institute of Technology, 2014.



3. Xue Jin. Research on The Optimal Solution of Lagrange Multiplier Function Method in Nonlinear Programming; Journal of Physics Conference Series 1952 042075. 2021
4. Richard W. Cottle. William Karush and the KKT Theorem. Documental Mathematica, Extra Volume ISPM , 2012.
5. Mikulas Luptacik. Kuhn-Tucker Conditions. Department of Economic Policy, University of Economics in Bratislava, Slovakia. ResearchGate, DOI: [10.1007/978-0-387-89552-9\\_2](https://doi.org/10.1007/978-0-387-89552-9_2). Sept 2010.
6. Vi Shang. Global Search Methods for Solving Nonlinear Optimization Problems. PhD Dissertation, University of Illinois at Urbana. <http://hdl.handle.net/2142/81906>. 1997.
7. J. K. Sharma. Operation Research; Theory and Applications, 4th Edition, Macmillan India Ltd, New Delhi. 2009.
8. Scott Moura. Chapter 4, Nonlinear Programming; CE 191 -CEE System Analysis. University of California, Berkeley. 2014
9. Brian Wallace. Constrained Optimization: Kuhn-Tucker Conditions. <http://cdn.world>. 2004.
10. Benoit C. Chachuat. Nonlinear and Dynamic Optimization from Theory to Practice. Info science EPFL scientific publication. 2007.
11. Ejikeme, C.L., Unaegbu, E.N. and Okofu, M.B. The Existence of a Periodic Solution of a Boundary Value Problem for a Linear System. IOSR Journal of Mathematics, e-ISSN:2278-5728, p-ISSN: 2319-765X. Volume 11, Issue 4 Ver. 1 (Jul-Aug). PP 73-82. 2015.