International Digital Organization for Scientific Research ISSN: 2550-7931
IDOSR JOURNAL OF APPLIED SCIENCES 2(2) 106-116, 2017.

Roles of Convexity in Optimization Theory

Efor, T. E and Nshi C. E

Department of Mathematics and Computer Science, Ebonyi State University, Abakaliki Nigeria.

E-mail: tesyefor@gmail.com

ABSTRACT

This paper, first, recall some strong implications of convex functions such as continuity and differentiability. It discusses Characterizations of convex functions in terms of first and second derivatives. It considers the roles of convexity in optimization problems in a finite dimensional space. Convex problems with inequality constraints are discussed and Kuhn and Tucker condition was applied to a real life problem in a small scale firm.

Keywords: Convex function, Derivative, continuity, Differentiability, epigraph, Karush-Kuhn Tucker, global optima.

INTRODUCTION

Optimization in general, is a mathematical procedure for determining optimal allocation of scarce resources, according to [1]. Convex optimization problem is a problem in which the objective functions and the constraint sets are convex. It has found practical applications in almost all sectors of life such as Mathematics, engineering, economics etc. This is because most algorithms that are used in computing the minimum point or maximum point if they exist of a function are point to point maps and therefore the solution they generate are extreme point. The attractiveness of convexity for optimization theory arises from the fact that when an optimization problem meets suitable convexity conditions, the same first order necessary optimality conditions we know to be local optima also become sufficient for global optima. Convexity has strong implications for continuity and differentiability such as, it must be continuous everywhere on the interior of its domain and also differentiable.

Convex constrained optimization problems are of two forms, equality and inequality constrained problems. Many authors such as, Efor, in [1], Robert [2], have discussed equality and inequality constrained problem without convexity respectively. Edward and et al in [3] discussed optimization in small scale Business

In this paper, we focus on characterizations of convex functions in terms of their first and second derivatives and the role of convexity to optimization theory. The paper, equally apply Karush-Kuhn and Tucker (KKT) conditions to a convex revenue problem with inequality constraints.

We recall some definitions in convexity.

(i). A set $K \subset \mathbb{R}^n$ is said to be convex if the convex combination of any two points say x, y is in K.

Intuitively, K is convex if the line segment joining two points, x and y is completely contained in K, that is

$[x, y] = \lambda x + (1-\lambda)\, y \subset K, 0 \le \lambda \le 1.$

(ii) A function $f: K \to K \subset \mathbb{R}^n$ is said to be convex if

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)\, f(y), \forall x, y \in K, \lambda \in (0,1).$$

If the inequality in the above definition is strict, and, $x \neq y$, then the function is called a strict convex function.

If the inequality is reversed then f is called a concave function

Geometrically, a function $f: K \subset \mathbb{R}^n \to K$ is said to be convex if the epigraph of f is a convex set,

$epi\, f = \{(x, \alpha) \in K \times \mathbb{R}: f(x) \le \alpha\}$ is a convex set

(iii) Revenue is the amount of money realized during a specific period of transaction, including discounts and deductions from returned goods/products .See [3] for more detail on revenue optimization.

CONVEX OPTIMIZATION MODEL WITH INEQUALITY CONSTRAINT

The model is posed as

Min$f(x)$

Subject to $x \in K$

Where, $K = p \cap \{x \in \mathbb{R}^n: h_i\,(x) \le\ 0, i = 1, \dots, k\}$. (p)

$p \subset \mathbb{R}^n$ is open and $h_i,: \mathbb{R}^n \to \mathbb{R}, i = 1, \dots, m,$ where f and K are convex.

We recall some strong implications of convexity such as continuity and differentiability.

CONTINUITY AS AN IMPLICATION OF CONVEXITY

In this section, we discuss continuity as implication of convexity. The result here, according to Sundaram, [4] shows that a convex function must be continuous everywhere on the interior of its domain.

Theorem 2.1: Let $f: K \to \mathbb{R}$ be a convex function. Then if K is open, f is continuous on K. If K is not open it is continuous on the interior of K. The proof is according to [4] and [5]

Proof

We prove that if K is open and f is convex on K, then f must be continuous on K. since int(K) is always open for any set K, and since the convexity of f on K also implies its convexity on int(K), this result will also prove that even if K is not open, f must be continuous on the interior of K. so suppose K is open and $x \in K$. Let $x_k \to x, x_k \in K$ for all K. Since K is open, there is $r > 0$, such that $B(x,r) \subset K$, pick λ such that $0 < \lambda < r$. Let $A \subset B(x,r)$ be defined by:

A= $\{y: \|y - x\| = \lambda\}$. Pick k so large that for all $k \geq K$, we have $||x_k - x|| \leq \lambda$. Since $x_k \to x_1$,it implies that such a k, exists.

Then, for all $k \geq K$, there is $y_k \in$ A such that $x_k = \lambda_k x + (1 - \lambda_k)y_k$ for some $\lambda_k \in (0,1)$.

Since, $x_k \to x_1$ and $\|y_k - x\| = \lambda > 0 \; \forall \; K$. It is the case that $\lambda_k \to 1$.

Therefore by the convexity of f,

$$f(x_k) = f(\lambda_k x + (1 - \lambda_k)y_k \leq \lambda_k f(x) + (1 - \lambda_k)f(y_k).$$

Taking limits, we have

$$\lim_{k\to\infty} \text{Sup } f(x_k) \leq f(x) \qquad (2.1)$$

Secondly, it is also that for all $k \geq K$, there is $z_k \in$ A and $\theta_k \in (0,1)$ such that $x = \theta_k x_k + (1 - \theta_k)z_k$. Exploiting convexity of f, we have

$$f(x) = f(\theta_k x_k + (1 - \theta_k)z_k \leq \theta_k f(x_k) + (1 - \theta_k)f(z_k).$$

Since θ_k must go to 1 as $k \to \infty$

Taking the limits, we have

$$f(x) \leq \lim_{k\to\infty} \inf f(x_k) \qquad (2.2)$$

Since we have established that

$\lim_{k\to\infty} \text{Sup } f(x_k) \leq f(x)$ and $f(x) \leq \lim_{k\to\infty} \inf f(x_k)$, we conclude that,
$\lim_{k\to\infty} f(x_k) = f(x)$

Remark: The continuity of f could fail at the boundary points of K.

Example 1.1

Define $f: [0,1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} \sqrt{x}, & \text{if } 0 < x < 1 \\ -1, & \text{if } x = 0,1 \end{cases}$$

Here f is concave on $[0,1]$ but discontinuous at the boundary points 0 and 1.

DIFFERENTIABILITY AS AN IMPLICATION OF CONVEXITY:

As with continuity, the assumption of convexity also caries strong implications for differentiability of the function involved. Consider the following definition of differentiability of f at x^*.

Definition 2.1: Let K be a nonempty set in $\mathbb{R}^n$ and let $f: K \to \mathbb{R}$. Then f is said to be differentiable at $x^* \in \text{int}(K)$ if there exist a vector $\nabla f(x^*)$, called the gradient vector and a function $\alpha: \mathbb{R}^n \to \mathbb{R}$ such that

$$f(x) = f(x^*) + \nabla f(x^*)(x - x^*) + \|x - x^*\| \alpha\,(x^*, x - x^*),$$

for each $x \in K$, $\lim_{x \to x^*} \alpha(x^*; x - x^*) = 0$.

Definition 2.2: Let K be a nonempty set in $\mathbb{R}^n$, and let $f: K \to \mathbb{R}$. Then f is said to be twice differentiable at $x^* \in \text{int}\,(K)$ if there exist a vector $\nabla f(x^*)$, and $n \times n$ symmetric matrix, $H(x^*)$, called the Hessian matrix, and a function $\alpha: \mathbb{R}^n \to \mathbb{R}$ such that,

$$f(x) = f(x^*) + \nabla f(x^*)^t (x - x^*) + \frac{1}{2}(x - x^*)^t\, H(x^*)(x - x^*) + \, ||x - x^*||^2 \alpha(x^*; x - x^*)$$

$\lim_{x \to x^*} \alpha(x^*; x - x^*) = 0$, for each $x \in$K, The function f is said to be twice differentiable on the open set $K^1 \subset K$ if it is twice differentiable at each point in K^1, where $H(x^*)$ is given by $\partial^2\, f(x^*)/dx, \partial x,$

CHARACTERIZATIONS OF CONVEX FUNCTIONS IN TERMS OF ITS DERIVATIVES

In this section, we characterize convex functions in terms of its first and second derivatives.

The theorem below gives a complete characterization of convexity of an everywhere differentiable function f using its first derivative.

Theorem 3.1: Let K be an open and convex set on $\mathbb{R}^n$, and let $f: K \to \mathbb{R}$ be differentiable on K. Then f is convex on K if and only if

$$f(y) \geq f(x) + \nabla f(x)^t (y - x) \forall x, y \in K.$$

Proof,

Suppose f is convex on K, we show that $f(y) \geq f(x) + \nabla f(x)(y - x), \forall x, y \in K$ holds.

by definition of convex function, we have

$$f(\lambda y + (1 - \lambda)x) \leq \lambda f(y) + (1 - \lambda)\, f(x), \forall x, y \in K \quad (3.1)$$

$$\Rightarrow f\big(x + \lambda(y - x)\big) \leq f(x) + \lambda\big(\,f(y) - f(x)\big)$$

$$\Rightarrow f\big(x + \lambda(y - x)\big) - f(x) \leq \lambda\big(\,f(y) - f(x)\big)$$

$$\Rightarrow \frac{f(x+\lambda(y-x))-f(x)}{\lambda} \leq \lambda(f(y)-f(x))$$

$$\lim_{\lambda\to 0^+} \frac{f(x+\lambda(y-x))-f(x)}{\lambda} = \nabla f(x)^t(y-x),$$

Therefore $\nabla f(x)^t(y-x) \leq f(y)-f(x)$

$\Rightarrow f(y) \geq f(x) + \nabla f(x)(y-x)$ ∎

Conversely,

Suppose $f(y) \geq f(x) + \nabla f(x)^t(y-x)\ ,\forall x,y \in K$ (3.2)

We show that f is convex.

Let $r,s\ \in K \Rightarrow \lambda s + (1-\lambda)r \in K, \forall\lambda \in (0,1)$

Set $w = \lambda r + (1-\lambda)s \in K$

$\Rightarrow w - s = -\lambda \frac{(w-r)}{1-\lambda}$ (3.3)

Now consider the pairs (w,r) and (w,s)

From, eq(3.2), we have

$f(r) \geq f(w) + \nabla f(w)^t(r-w)$ (3.4)

Again

$\Rightarrow f(s) \geq f(w) + \nabla f(w)^t(s-w)$ (3.5)

Form eq (3.3)

$$\Rightarrow f(s) \geq f(w) + \nabla f(w)^t\left(-\lambda\frac{(r-w)}{1-\lambda}\right)$$

$\Rightarrow f(s) \geq f(w) + \left(-\lambda\frac{(r-w)}{1-\lambda}\right)\nabla f(w)$ (3.6)

$\lambda f(r) \geq \lambda f(w) + \lambda\nabla f(w)^t(r-w)$ (3.7)

$(1-\lambda)f(s) \geq (1-\lambda)f(w) - \lambda\nabla f(w)^t(r-w)$ (3.8)

Adding (3.7) and (3.8) we have

$\lambda f(r) + (1-\lambda)f(s \geq f(w)$

$\Rightarrow f(w) \leq \lambda f(r) + (1-\lambda)f(s)$ (3.9)

But, $w = \lambda r + (1-\lambda)s$

Substituting w in (3.9), we have,

$f(\lambda r + (1-\lambda)s) \leq \lambda f(r) + (1-\lambda) f(s), \forall r, s \in K,$ ■

Hence, f is convex.

We now give the second characterization in terms of its derivatives.

Theorem 3.2: Let K be a nonempty open convex set in $\mathbb{R}^n$ and let $f: K \to \mathbb{R}$ be a C^2 function on K. Then f is convex if and only if the Hessian matrix is positive semi-definite at each point in K.

Proof:

Suppose that f is convex and $x^* \in K$. We show that

$x^t H(x^*) x \geq 0,$

for each $x \in \mathbb{R}^n$.

Since$K$ is open, then for any given $x \in \mathbb{R}^n, x^* + \lambda x \in K$ for λ sufficently small.
By convexity and differentiability of f, we have

$$f(x^* + \lambda x) \geq f(x^*) + \lambda \nabla f(x^*)^t x \quad (3.10)$$

and

$$f(x^* + \lambda x) = f(x^*) + \lambda \nabla f(x^*)^t x + \frac{1}{2}\lambda^2 x^t H(x^*) x + \lambda^2 ||x||^2 \alpha(x^*; \lambda x) \quad (3.11)$$

Substituting (3.11) in (3.10), we have

$$\frac{1}{2}\lambda^2 x^t H(x^*) x + \lambda^2 ||x||^2 \alpha(x^*; \lambda x) \geq 0 \quad (3.12)$$

Dividing (3.12) by λ^2 and let $\lambda \to 0$. We have that,

$x^t H(x^*) x \geq 0.$

Conversely,

Suppose that the Hessian matrix is positive semi-definite at each point in K.

Consider x and x^* in K. Then, by the Tailors theorem, we have

$$f(x) = f(x^*) + \nabla f(x^*)^t (x - x^*) + \frac{1}{2}(x - x^*)^t H(x^0)(x - x^*) \quad (3.12)$$

Where $x^0 = \lambda x^* + (1-\lambda)x$ for some, $\lambda \in (0,1)$. Note that $x^0 \in K$, and hence, by assumption, $H(x^0)$ is positive semi-definite. Thus,

$$(x-x^*)^t H(x)(x-x^*) \geq 0,$$

from (3.12), we conclude that,

$$f(x) \geq f(x^*) + \nabla f(x^*)^t (x-x^*)$$

Since the above inequality is true for each $x, x^* \in K$, f is convex.

Example 1.2: Consider a function $f: \mathbb{R}^2 \to \mathbb{R}$, defined by

$$f(x_1,x_2) = 2x_1 + 6x_2 - 2x_1^2 - 3x_2^2 + 4x_1x_2.$$

We verify the convexity or convenient form.

$$f(x_1,x_2) = (2,6)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{1}{2}(x_1,x_2)\begin{pmatrix} -4 & 4 \\ 4 & -6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

In order to check whether the Hessian matrix H is positive semi-definite or negative semi-definite or neither, we compute the eigen-value by solving the following system.

$$\det\left(H - \lambda \underline{1}\right) = 0 \Rightarrow \begin{vmatrix} -4-\lambda & 4 \\ 4 & -6-\lambda \end{vmatrix} = (-4-\lambda)(-6-\lambda) - 16 = 0$$

$$\Rightarrow \lambda^2 + 10\lambda + 8 = 0.$$

The solutions of this equation are

$$\lambda_1 = -5 + \sqrt{17, \lambda_2} = -5 - \sqrt{17}.$$

since, $\lambda_1\lambda_2$ are negative, then H is negative semi-definite and hence f is concave.

ROLE OF CONVEXITY IN OPTIMIZATION PROBLEMS

The role of convexity in optimization theory is very vital in the sense that it gives validity for global optima. In a convex minimization problem, any critical point of f is a global minimizer of f and if f is strictly convex and admits a solution, then the solution must be unique.

Theorem 4.1: Let $f: K \subset \mathbb{R}^n \to \mathbb{R}$ be a convex function with continuous first partial derivatives on some open set containing the convex set K, then any critical point of f is a global minimizer of f.

Proof

Suppose that $x^* \in K$ is critical point of f, then

$$\nabla f(x^*) = 0.$$

Let $y \in K$, so that

$f(y) = f(x^*) + \nabla f(x^*)(y - x^*)$, (by Tailor theorem) (4.1)

But f is convex on K

$\Rightarrow f(x^*) + \nabla f(x^*)(y - x^*) \leq f(y), \forall x^*, y \in K$

$\Rightarrow f(x^*) \leq f(y), \forall y \in K$

Hence, x^* is a global minimizer of f.

The theorem below shows that if f is a strictly convex and admits a solution, then the solution must be unique.

Theorem 4:2 Let $f: \mathbb{R}^n \to \mathbb{R}$ be a function. If f is strictly convex on K, then

(1) f has at most one minimizer, that is if the minimizer exists, then it is unique.

(2) any local minimizer of a convex function f is also a global minimizer.

Proof

Let x_1 and x_2 be two different minimizers of f and let $\lambda \in (0,1)$. Because of the strict convexity of f and the fact that

$f(x_1) = f(x_2) = \min_{x \in K} f(x)$,

We have

$f(x_1) \leq f(\lambda x_1 + (1 - \lambda)x_2) < \lambda f(x_1) + (1 - \lambda)f(x_2) = f(x_1)$,

This is contradiction and therefore,

$x_1 = x_2$.

Proof of (2)

Suppose that x^* is a local minimize of f in K. Then there is a positive number $r > 0$ such that

$f(x^*) \leq f(x), \forall x \in K \cap B(x^*, r)$.

Given any $x \in K$, we show that

$f(x^*) \leq f(x)$.

Chose $\lambda \in (0,1)$ so small that

$x^* + \lambda(x - x^*) = \lambda x + (1 - \lambda)x^* \in K \cap B(x^*, r)$

Then, by convexity of f, we have

$f(x^*) \leq f\big(x^* + \lambda(x - x^*)\big) = f(\lambda x + (1-\lambda)x^*) \leq \lambda f(x) + (1-\lambda)f(x^*)$.

Subtracting, $f(x^*)$ from both sides of the proceeding, inequality,

We have

$0 \leq \lambda\big(f(x) - f(x^*)\big)$.

$\Rightarrow 0 \leq f(x) - f(x^*)$

$\Rightarrow f(x^*) \leq f(x), \forall x \in K$.

This implies that x is a global minimizer of f.

CONVEX AND INEQUALITY-CONSTRAINTS

The main result here is the Theorem of Kuhn-Tucker, (KKT). In this result the first order conditions of the are both necessary and sufficient to identify optima of convex inequality-constrained optimization problems. This is possible if a mild regularity condition is met.

Theorem 5.1 (The Theorem of Kuhn-Tucker): Let f be a $C^{'}$ function mapping p into$\mathbb{R}$, where $P \subset \mathbb{R}^n$ is open and convex. Let $h_i: K \to \mathbb{R}$ also be convex/concave $C^{'}$ function, for, $i = 1, \dots, m$.

Suppose there exist some $x^* \in P$ such that the Slater's condition

$h_i(x^*) > 0, i = 1, \dots,$,m

Then, x^* maximizes f over

$K = \{x \in P: h_i(x) \geq 0, i = 1, \dots, m\}$

If and only if there exists $\lambda^* \in \mathbb{R}^k$ such that the Kuhn-Tucker first order conditions hold:

$[KKT - 1]\ \nabla f(x^*) + \sum_{i=1}^{m} \lambda^* \nabla h_i(x^*) = 0$

$[KKT - 2]\lambda^* \geq 0, \lambda^*\ h_i(x^*) = 0$

Remark: Slater's condition is used only in the proof that, $[KKT - 1]$ and $[KKT - 2]$ are necessary at an optimum. It is not necessary in proof of sufficiency.

APPLICATION

A bock Industry is producing two different types of block 9 inches and 6 inches blocks

From the industry, we discovered the following about their revenue. The sells of the product are in naira

(i) They sell in a defined bundle for each product
(ii) The price for product A (9 inches) x_1=150
(iii) The price for product B (6 inches) $x_2 = 140$

(iv) The demand for $x_1 \geq x_2$, and $x_1 + 2x_2 \leq 100$

The revenue function is given by:
Max $f(x_1,x_2) := 150(x_1^2 + x_1) + 140(x_2^2 + x_2)$

The problem is modeled as follows:

Max$f(x_1,x_2) := 150(x_1^2 + x_1) + 140(x_2^2 + x_2)$,
S.t $x_1 + 2x_2 \leq 100$
$x_1 \geq x_2$

The revenue problem is a convex problem, f, is strictly convex and since it is has inequality constraint, we apply the KKT conditions in the problem above,

L((x,λ_1,λ_2):=$150(x_1^2 + x_1) + 140(x_2^2 + x_2) + \lambda_1 (100 - x_1 - 2x_2) + \lambda_2 (-x_2 - x_1)$

$\frac{\partial L(x,\lambda_1,\lambda_2)}{\partial x_1} = 300x_1 + 150 - \lambda_1 - \lambda_2$

$\frac{\partial L(x,\lambda_1,\lambda_2)}{\partial x_2} := 280x_2 + 140 - 2\lambda_1 - \lambda_2$

$\frac{\partial L(x,\lambda_1,\lambda_2):}{\partial \lambda_1} = \lambda_1 (100 - x_1 - 2x_2)$

$\frac{\partial L(x,\lambda_1,\lambda_2):}{\partial \lambda_2} = \lambda_2 (x_1 - x_2)$

$x_1 + 2x_2 \leq 100$

$-x_1 + x_2 \leq 0$,λ_1,$\lambda_2 \geq 0$

Since, we have two complementary slackness, we have four cases

Case1.
$\lambda_1 = 0, \lambda_2 = 0$

$300x_1 + 150 = 0, \quad x_1 = -\frac{1}{2}$

$280x_2 + 140 = 0, \quad x_2 = -\frac{1}{2}$,

$(x_1, x_2) = (-\frac{1}{2}, -\frac{1}{2})$

is not feasible

Case ii

$\lambda_1 = 0, x_1 - x_2 = 0, x_1 = 50, x_2 = 50$. Not feasible

Case iii

$\lambda_2 = 0,\ 100 - x_1 - 2x_2 = 0,$

$x_1 = \frac{100}{3}, x_2 = 25, \lambda_1 = 515.$

The revenue will be optimized at

$(x_1^*, x_2^*) = (\frac{100}{3}, 25)$

Hence, (x_1^*, x_2^*) is a global solution

SUMMARY/CONCLUSION

So far, the paper has discussed some implications of convexity such as continuity and differentiability and the result have showed that a convex function is continuous and differentiable everywhere on the interior of its domain. Characterizations of convexity in terms of first and second derivatives have been discussed. The role of convexity to optimization problems is very vital in the sense that, if the optimization problem is convex, then any critical point of f is a global minimizer of f and if f is strictly convex and admits a solution, then the solution must be unique. Karush-Kuhn Tucker condition was applied to a small scale revenue problem.

REFERENCES

1. Efor, T. E: Optimality Conditions for equality constrained optimization problems. *International Journal of Mathematics and Statistics invention, 4(4), 2016, pp 28-33.*

2. Robert Phillips: Teaching pricing and revenue optimization, *htt//ite.pubs.inform.org/vol4 NoI/Phillips/, 2004.*

3. Edward and et al: Performance and Constraints of Small Scale Enterprises in the Accra Metropolitan Area of Ghana, *European Journal of Business and Management, Vol.5, No.4, 2013.*

4. Rangarajan, K. Sundaram, A first course in optimization theory, Cambridge, United States; University Press, (1998).

5. Hands, D. W .Introductory Mathmatical Economics, Second Edition, Oxford Univ. Press, 2004