

## A note on convex programming in practical problems

Idowu Ademola Osinuga and Dawud Adebayo Agunbiade

Department of Mathematical Sciences, Faculty of Science, Olabisi Onabanjo University,  
 Ago-Iwoye, Nigeria

**Keywords:** convex programming, optimal design, location theory.

**Abstract.** In recent years, convex programming has become a sophisticated tool of central importance in engineering, finance, operations research, statistics etc. The goal of this paper is to emphasize modeling and present several convex programming problem formulations especially in optimal design and location theory. We build simple models to address this question, investigate their properties and apply a variant of the Weierstrass theorem to prove the existence of a solution. Our results extend and improve some other comparable results of the author [2,8,20,21,22,23,24,25].

### 1. Introduction

The general unconstrained programming problem can be stated as the problem of finding the vector  $x^*$ , where the minimum occurs, i.e.,

$$\text{Minimize } f(x) = f(x)^* \quad (1)$$

The function  $f$  is called the objective function. The elements  $x$  are often called the decision variables. Convex programming (CP) is a subfield of mathematical programming, or simply mathematical optimization, which studies the problem of minimizing convex functions. The usual and most intuitive form of describing a convex programming is:

$$\begin{aligned} &\text{Minimize } f(x) \\ &\text{subject to} \\ &g_i(x) \leq 0, i = 1, 2, \dots, n \\ &h_j(x) = 0, j = 1, \dots, k \end{aligned} \quad (2)$$

where  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is a convex function to be minimized over the variable  $x$ , and each  $f_i$  and  $h_j$  are real-valued functions defined on  $\mathbb{R}^n$ . The constraints  $f_i(x) \leq 0, i = 1, 2, \dots, n$  are referred to as inequality constraints;  $h_j(x) = 0, j = 1, \dots, k$  are referred to as equality constraints.

If the objective and constraint functions are differentiable we refer to the above equation (2) as differentiable programming.

The optimization problem (2) is called a linear programming problem (LPP) if the objective and constraint functions are linear i.e.,

$$f_i(\lambda x + \gamma y) = \lambda f_i(x) + \gamma f_i(y), \forall x, y \in \mathbb{R}^n \text{ and } \lambda, \gamma \in \mathbb{R}. \quad (3)$$

$$\text{e.g minimize } c^T x, \text{ subject to } Ax \leq b, x \geq 0 \quad (4)$$

Here  $c, x \in \mathbb{R}^n$  and  $A$  is an  $m \times n$  matrix and  $b \in \mathbb{R}^m$  and the constraint  $x \geq 0$  is usually referred to as non-negativity restriction. Consider another example given by

$$\text{minimize } \left( \frac{1}{2} x^T Q x + p^T x \right), \text{ subject to } Ax \leq b, x \geq 0 \quad (5)$$

where  $Q$  is an  $n \times n$  matrix and  $b \in \mathbb{R}^m$ . Then (2) is said to be a quadratic programming problem under linear constraints. This class of problems is very important for practical purposes as well as linear programming problems.

Convex programming has applications in a wide range of disciplines, such as automatic control systems, estimation and signal processing, communications and networks, electronic circuit design, data analysis and modeling, statistics and finance [4,5,12,13,27]. With recent improvements in computing and optimization theory, convex programming is nearly as straight forward as linear programming. Due to its ability to solve large, practical engineering problems it has been useful in solving some conic programming problems.

The basic idea and focus of this combined research is that many analyses and designs arising in science and engineering can be cast, or recast, in the form of a convex programming problem, i.e., minimizing a convex objective of some decision variable subject to some convex constraints on the variable. Although such problems can appear difficult, they may have hundreds of variables and nonlinear, nondifferentiable objective functions, they can be solved (numerically) very efficiently by recently developed interior point methods that exploit convexity and the particular problem structure. Thus, the original problem is efficiently solved.

In view of the fact that not all problems arising in science and engineering are convex hence amenable to rapid numerical solution and in cases where the problem is not convex, convex approximations can yield suboptimal solutions that are very useful. In practice. Generally, then, it is useful skill to recognize convexity in engineering and science problems and know how to exploit it.

A large body of literature is devoted to the algorithms development for the solution of the highlighted problems [2,7,8,11,15,19,23] and it is our wish to apply a simple and more general results to resolve(solve) the problem.

Our paper is organized as follows: we specify the problem and some of its basic properties (section 2); we present a general result in finite dimensional setting to resolve the problem; show connections between them and describe how we can apply such result to tackle the problems (section 3). Finally in section 4 we briefly state our conclusions.

## 2. Preliminaries and Problem Setup

In this section, we set up the problems and discuss some of its properties.

The simplest convex programming problem is an unconstrained problem of the form

$$\text{Minimize } \|Ax-b\| \quad (6)$$

where  $A \in \mathbb{R}^{m \times n}$  and  $b \in \mathbb{R}^m$  are problem data,  $x \in \mathbb{R}^n$  is the variable, and  $\|\cdot\|$  is a norm on  $\mathbb{R}^m$ .

A solution of (6) is sometimes called an approximate solution of  $Ax \approx b$ , in the norm  $\|\cdot\|$ . The vector

$$r = Ax - b \quad (7)$$

is called the residual for the problem, its components are sometimes called the individual residuals associated with  $x$  [5].

An  $n$  extension of (6) is the weighted convex programming problem

$$\text{Minimize } \|w(Ax-b)\| \quad (8)$$

where the problem data  $w \in \mathbb{R}^{m \times m}$  is called the weighting matrix.

This kind of problem arises in many disciplines as discussed in the earlier section and some of the many names that have been used depending on areas of discipline include least square norm or least square problems in statistics and engineering, Weber problem in location research and norm approximation in optimization and operations research.

In recent times, research has been carried out on the minimum norm problem which is also convex programming problem and resolved using several techniques, see ([5,12,14,20,21]) to mention a few. Such problems have been found useful in

Approximation theory, statistical estimation problem [5], signal and image reconstruction as well as in other engineering applications [12].

In [12], the author showed that minimum norm problem can be recast as fixed point problem and showed  $x = Tx$ . He further proved the existence and uniqueness of the minimum solution of operator equation, if T is non expansive. The research carried out by [14] was to find the minimum norm solution of linear programs by a Newton-type method which was shown to be globally convergent.

A recent stride reported in [20] was directed at an estimation problem using simple random sampling technique. The idea was used in formulating the estimation problem as an equivalent minimum norm problem in the Hilbert space and resolved by an appropriate application of the classical projection theorem. In [21], a critical survey was carried out on various applications of minimum norm with emphasis on finance.

In this paper, we show that convex programming problem can be recast as a minimum norm problem in practical problems and resolved using a more general approach.

### 2.1 Optimal design

The least squares method describes a frequently used approach to solving overdetermined or inexact systems of equations in an approximate sense. Instead of solving the equations exactly we seek only to minimize the sum of squares of the residuals.

The least squares criterion has important statistical interpretations. If appropriate probabilistic assumptions about underlying error distributions are made, least squares produces what is known as the maximum-likelihood estimate parameters. Even if the probabilistic assumptions are not satisfied, research in this areas has shown that least squares produces useful results [1,2,5,6,7,8,9,10,11,14,15,19,21,22,23,24,25,26]

A very common source of least squares is curve fitting. Let  $x$  be the independent variable and let  $y(x)$  denote an unknown function of  $x$  that we want to approximate. Assume there are  $m$  observations, i.e., values of  $y$  measured at specified values of  $x$  :

$$y_i = y(x_i), i = 1, \dots, m \tag{9}$$

The idea is to model  $y(x)$  by a linear combination of  $n$  basis functions:

$$y(x) \approx \beta_1 \phi_1(x) + \dots + \beta_n \phi_n(x) \tag{10}$$

The design matrix A is a rectangular matrix of order  $m \times n$  with elements  $t_{ij} = \phi_j(x_i)$

The design matrix usually has more rows than columns. In matrix-vector notation, the model is

$$y \approx A\beta \tag{11}$$

Some common models include the following:

straight line:  $y(x) \approx \beta_1 x + \beta_2$  (12)

polynomials:

$$y(x) \approx \beta_0 x^{n-1} + \dots + \beta_{n-1} x + \beta_n \tag{13}$$

log-linear:

$$y(x) \approx Ke^{\lambda x}, \log y \approx \beta_1 x + \beta_2, \text{ with } \beta_1 = \lambda, \beta_2 = \log K, \text{ etc} \tag{14}$$

The residuals in this case are the differences between the observations and the model:

$$r_i = y_i - \sum_{j=1}^n \beta_j \phi_j(x_i, \alpha), i = 1, \dots, m \tag{15}$$

or, in matrix-vector notation

$$r = y - A(\alpha)\beta \quad (16)$$

We want to find the  $\alpha$  's and  $\beta$  's that make the residuals as small as possible. In least squares method we wish to minimize the sum of the squares of the residuals

$$\|r\|^2 = \sum_{i=1}^m r_i^2 \quad (17)$$

If some observations are more important or more accurate than others, then we might associate different weight,  $w_i$ , with different observations and

$$\|r\|^2 = \sum \omega r^2 \quad (18)$$

Other uses of this type of problem in statistical application occur mainly to collection of data, for example, when we design an experiment or a survey so as to minimize experimental errors. Here the residual is essentially an error in the fit of the model, say

$$Y(x) = \beta_1(x) + \beta_2 \quad (19)$$

The fitted line has predicted values at points on the line and hence the residuals are vertical deviations from points to the line. As a result, the least squares procedure produces a line that minimizes the sum of squares of vertical deviation from the points to the line.

In an experimental context in the physical sciences and engineering almost all measured quantities have an error because a perfect apparatus does not exist. Least squares fitting techniques provides some guidelines for determining the values of those errors. Recent publications directed at applications of convex programming in statistical and engineering designs can be found in [1,2,18,19].

## 2.2 Epidemic Models

The least squares fitting techniques has been employed as one of the approaches of estimating epidemiological parameters for daily case notification time series for pandemic diseases. It is used in trajectory matching of epidemic model to epidemic curve data. We use an epidemic model of SEIR-type that classifies individuals as susceptible(S), exposed (E), Infectious (I), recovered (R) and dead(D).The transmission process can be modeled using the following system of nonlinear differential equations:

$$\begin{aligned} S'(t) &= -\beta S(t)I(t) / N(t) \\ E'(t) &= \beta S(t)I(t) / N(t) - kE(t) \\ I'(t) &= kE(t) - (\gamma + \delta)I(t) \\ R'(t) &= \gamma I(t) \\ D'(t) &= \delta I(t) \\ C'(t) &= kE(t) \end{aligned}$$

where the 'prime' denotes time derivatives, infectious individuals either recover or die from the pandemic at the mean rates of  $\gamma$  and  $\delta$  respectively,  $\beta$  is the transmission rate,  $1/k$  is the mean latent period and  $C(t)$  is the cumulative number of case notification (Chowell et.al. 2007).

Specifically, we fit the cumulative number of cases given by equation  $C(t)$  to the cumulative number of case notifications. In the above case the residual is given by

$$\min_t \|C(t) - C'(t)\|^2 = \min \|r_i\|^2 \quad (20)$$

This type of problem possesses existence and uniqueness properties for Hilbert, Banach reflexive, rotund spaces and duals of smooth Banach spaces. The nice properties of least squares as well as their physical interpretations lead to many important applications.

### 2.3 Location Theory

There are numerous publications where facility location problems have been discussed [4,5,8,9,10,13,14,15,17,19,20,21,22,23].The problem has given rise to extraordinary number of generalizations, extensions and modifications. It would literally require volumes to do them justice; space only permits only a brief and somewhat arbitrarily selected summary. However, comprehensive reviews of this type of problem can be found in [16, 26].

This problem is sometimes called Weber problem and it arises in many fields such as signal processing, networking, communications and called many names such as the Fermat –Torricelli problem, the Steiner problem, the Steiner-Weber problem, the median center problem, the minisum problem, and the spartial problem [10].

We are to find the “minisum” point  $(x_f, y_f) \in X$  which minimizes the sum of weighted Euclidean distances from itself to  $n$  fixed points with coordinates  $(x_i, y_i) \in P, i = 1, \dots, n$ . The weights which are associated with the fixed points are denoted by  $w_i$ . One simple scenario for the problem is that of locating a distribution centre and that the weights  $w_i$  are the cost per unit distance of shipping the requirements to customers located at the fixed points  $(x_i, y_i)$ ;  $(x_f, y_f)$  is then the distribution centre that minimizes the transportation costs.

The problem can be stated as:

$$\min \left\{ W(x, y) = \sum_{i=1}^n w_i d_i(x, y) \right\} = \min w_i \|X - P\| \tag{21}$$

Where  $d_i(x, y) = \sqrt{(x - x_i)^2 + (y - y_i)^2}$  is the Euclidean distance between  $(x, y) \in X$  and  $(x_i, y_i) \in P$ . The equation can be rewritten as

$$\min \left\{ C(x, y) = \sum_{i=1}^n w_i d_i^2(x, y) \right\} = \min w_i \|X - P\|^2 \tag{22}$$

indicating the minimization of the sum of squared Euclidean distances that looks similar to (17) and (20).It can depict a wide variety of distance measures; the Euclidean distance is the special case  $p = 2$  and rectilinear distance is  $p = 1$  ; looking at the generalization of Euclidean distance ( $l_p$  distance),which is given by

$$l_p = \sqrt[p]{|x - x_i|^p + |y - y_i|^p} \tag{23}$$

The Weber problem has two very important properties. First, the objective function is a convex function which ensures that any local optimum is also a global optimum. Second, the optimal location for the new facility must lie within the convex hull of the existing facility locations. Hence all the problems discussed are convex programming and as a result finding a global optimum of a norm-based objective function is often tractable.

### 3. The Main Results

We present the results applicable to the earlier discussed problems and show the connections existing between them. In the sequel we state the following definitions and proposition that will be needed for further development.

Definition 3.1. Let  $X$  be a metric space and  $M \subseteq X$  be a non-empty convex set. A Function  $f: X \rightarrow \mathbb{R}$  is convex if  $\forall \alpha \in [0,1]$  and  $\forall x, y \in M$

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(\alpha) + (1 - \alpha)f(y) \quad (24)$$

A strict convex function is defined by

Definition 3.2. Let  $X$  be a metric space and  $M \subseteq X$  be a non-empty convex set. A function  $f: X \rightarrow \mathbb{R}$  is convex if  $\forall \alpha \in (0,1)$  and  $\forall x, y \in M, x \neq y$

$$f(\alpha x + (1 - \alpha)y) < \alpha f(\alpha) + (1 - \alpha)f(y) \quad (25)$$

Definition 3.3. A twice differentiable function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be strongly convex if there exists a real number  $\rho > 0$  such that for all  $x, y \in \mathbb{R}^n$ .

**Theorem 3.1** [3,27]. Consider a convex function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ . Then every local minimum of  $f$  is a global minimum.

Proposition 1.1. The norm function  $f(x) = \|x\|$  is a convex function.

Proof: Let  $f(x) := \|x\|$ . Then for any  $y, z$  and  $\lambda \in [0,1]$ , we have

$$\begin{aligned} f(\lambda y + (1 - \lambda)z) &= \|\lambda y + (1 - \lambda)z\| \leq \|\lambda y\| + \|(1 - \lambda)z\| \\ &= \lambda \|y\| + (1 - \lambda)\|z\| = \lambda f(y) + (1 - \lambda)f(z) \end{aligned} \quad (25)$$

which completes the proof.

The above proposition implies that since every norm is a sublinear function then every norm is a convex function. As a result, finding a global optimum of a norm-based function is often tractable as pointed out in the earlier section.

**Theorem 3.2.** Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is a twice differentiable strongly convex function over a compact set  $X \subset \mathbb{R}^n$  and assume that the  $domf$  is a compact set. Then  $f$  attains a maximum and a minimum on  $X$ .

Proof. Since  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is a strongly convex and hence a convex function it is continuous.

To show compactness of  $domf$ , assume that  $m = \inf_{x \in \mathbb{R}^n} f(x)$ . Then naturally  $m = \inf_{x \in domf} f(x)$

In a similar manner, we can say that there exists a sequence  $x_n \in \mathbb{R}^n$  such that  $f(x_n) \rightarrow m$ . But since  $domf$  is compact we have that there exists a subsequence  $x_{n_i}$  of the stated sequence such that  $x_{n_i} \rightarrow x_0$ , where  $x_0 \in domf$ . Then by lower semicontinuity of  $f$  we have  $f(x_0) \leq m$ . This shows that  $f$  attains the minimum  $m = f(x_0)$ . Similarly, we have by upper semicontinuity that  $f$  attains the maximum  $n=f(x_i)$

Remark. We have used the fact that every real-valued convex functions on  $\mathbb{R}^n$  is Lipschitzian and Lipschitzian function is necessarily lower semicontinuous function. The uniqueness of the minimum norm problem is tractable since strongly convex function are strictly convex.

Example 1.1: This example illustrates the use of the technique developed that identifies the subset having the conditions satisfying existence of solution.

$$\text{Minimize } f(x) = \|Ax - b\|_2^2 = \text{minimize } \sum_{i=1}^n (a_i^T x - b_i)^2, x \in M$$

where  $M \subset \mathbb{R}^n$ .

Since  $f$  is a norm function (also a convex function) is continuous on its domain. By theorem 3.2, a solution exists as long as the set  $M$  is compact. To show that  $M$  is compact, we have  $M \subset \mathbb{R}^n$  a subset of  $X$  and hence  $M$  is compact. By the theorem 3.2 the solution exists and indeed a global solution by theorem 3.1.

#### 4. Conclusion

In this paper, applications of convex programming in practical problems are herein considered and resolved using a more general approach in finite dimensional settings. The work demonstrated the importance of convex programming in the presence of faults, measurement errors and statistical uncertainties. By combining the basic and desirable properties of normed spaces and convexity, we have shown the resolution of practical problems using abstract results.

#### 5. Acknowledgements

Special thanks to the management of Olabisi Onabanjo University, Ago-Iwoye, Nigeria.. The authors thank the management and faculty members of Lahore University of Management Sciences, Lahore for their hospitality and support.

#### References

- [1] Atashgar, A.B., Seifi, A.: Optimal design of multi-response experiments using semi-definite programming, *Optim.Eng.*, 10, 75-90 (2009).
- [2] Atieg, A., Watson, G.A.: Use of Anziam J., 45 4, C187-C200 (2004).l p norms in fitting curves and surfaces to data,
- [3] Bector, C.R., Chandra, S., Dutta, J.: Principles of optimization theory, Narosa publishing house, New Delhi, (2005)
- [4] Boyd, S., Vanderberghe, L.: CRCD programme: Convex optimization for engineering analysis and design, Proceedings of the American control conference, Seetle, Washington, (1996).
- [5] Boyd, S., Vanderberghe, L.: Convex optimization, Cambridge University press, Cambridge, (<http://www.stanford.edu/~boyd/cvxbook>), (2004).
- [6] Butt, S.E., Cavalier, T.M.: An efficient algorithms for facility location in the presence of forbidden regions, *Eur.J.Oper.Res.*, 90, 56-70 (1996)
- [7] Chares, R., Glineur, F.: An interior-point method for the single facility location problem with mixed norms using a conic formulation, *Ecore discussion paper*, Universite Catholique de Louvain, (2007).
- [8] Chowell, G., Nishiura, H and Bettencourt, L.M.A.: Comparative estimation of the reproduction number for pandemic influenza from daily case notification data, *J.R.Soc.Interface*, 4, 155-166 (2007).
- [9] Dasarathy, R., White, L.J.: A maxmin location problem, *Oper.Res.*, 28 6, 1385-1401 (1980).
- [10] Drezner, Z., Klamroth, K., Schobel, A and Wesolowsky, G.O.: The Weber Problem, In: Z.Drezner, H.W.Hamacher (Eds.): *Location analysis: Applications and theory*, 1-36, Springer Verlag, (2002).
- [11] Drezner, Z., Scott, C. and Song, J-S.: The central warehouse location problem revisited, *IMA J.Mgt. Math.*, 14, 321-326 (2003).
- [12] Ferreira, P.S.G.: The existence and uniqueness of the minimum norm solution to certain linear and nonlinear problems, *Signal Processing*, 55, 137-139 (1996).

- 
- [13]Hindi, H.: A tutorial on convex optimization, Palo Alto research centre (PARC), Palo Alto, California [online] available at <http://www2.parc.com/spl/members/hhindi/reports/CvxOptTutPaper.pdf>, accessed 5 October, 2010.
- [14]Kanzow, C., Qi, H. and Qi, L.: On the minimum norm solution of linear programs, Applied Math.Report, AMR00 14, 1-11 (2000).
- [15]Kocak, H.: Convex programming approach to the shopping mall (AVM) site selection problem and Sakarya, Eur.J.Social Sc. 13 2, 219-228 (2010).
- [16]Kreyszig, E.: Introductory functional analysis with applications(student edition), John Wiley, Singapore, (2005).
- [17]Kuhn, H.W.: A note on Fermat's problem, Mathematical Programming, 4, 98-107 (1973).
- [18]Love, F.R., Morris, J.G and Wesolowsky, G.O.: Facilities location:models and methods, Elsevier, New York, (1988).
- [19]Moarref, M., Sayyaadi, H.: Facility location optimization via multi-agent robotic systems, IEEE Trans. Robotics Automation,(2008).
- [20]Oniyide, O.R., Osinuga, I.A.: On the existence of best sample in simple random sampling, J.Math.Assoc.Nigeria, 33 2b, 290-294 (2006)
- [21]Osinuga, I.A., Oniyide, O.R.: An overview of norm approximation applications, Far East J.Math. Sci., 43 2, 189-202 (2010).Osinuga, I.A., Bamigbola, O.M.: On the minimum norm solution to Weber problem, Proceedings of the 25<sup>th</sup> SAMSA (Southern Africa Mathematical Sciences Association) Conference, Windhoek, Namibia, 2, 54-60 (2007).
- [23]Schandl, B., Klamroth, K. and Wiecek, M.M.: Norm-based approximation in bicriteria programming, Computational Optim.Appl., 20, 23-42 (2001).
- [24]Schandl, B., Klamroth, K. and Wiecek, M.M.: Norm-based approximation in convex multicriteria programming, In: B. Freischmann, R. Lasch, U. Derigs, W. Domschke and U. Rieder (Eds.): Operations Research Proceedings 2000, 8-13, Springer Verlag, (2001).
- [25]Schandl, B., Klamroth, K. and Wiecek, M.M.: Norm-based approximation in multicriteria programming, Computers Math.Appl., 44, 925-942 (2002).
- [26]Srebro, R.: Iterative refinement of the minimum norm solution of the bioelectric inverse problem, IEEE Trans. Biomedical Eng., 43 5, 547-552 (1996).
- [27]Sundaram, R.K.: A first course in optimization theory, Cambridge University press, New York, (1996).
- [28]Wesolowsky, G.O.: The Weber problem: History and Perspectives, Location Science, 1, 5-23 (1992).