A Modified Super Convergent Line Search Algorithm for Solving Quadratically Constrained Quadratic Optimization Problems

Type of Article

Received: XX December 20XX
Accepted: XX December 20XX
Online Ready: XX December 20XX

Abstract

In this paper, a modified super convergence line search algorithm is introduced to address the use of the optimal support points of the segmented design space and the bias to evaluate the optimal step length in a quadratically constrained quadratic optimization problem. The optimum number of segment and the minimum number of iterations are considered and the method modified the algorithm by linearising the quadratic constraint through the partial derivative of the Jacobian function to attained optimal step-length and convergence. The new algorithm was applied to two quadratically constrained problems and the results indicated that the modified algorithm satisfied the convergent criteria in six and one iteration and achieved the optimal solutions in both problem set respectively. This shows the effectiveness of the modified algorithm in solving quadratically constrained quadratic optimization problems.

Keywords: Quadratically constrained quadratic optimization problem; segmentation; bias; step length.
2010 Mathematics Subject Classification: 9090; 62L05; 62L20

1 Introduction

A variety of optimal design methods have been shown in literature as methods of optimization of Quadratic Constrained Problems (QCP). Such as; The active set method, The Gradient method, The Barrier Method, The Steepest

*Corresponding author: E-mail: okonkwo.chukwunenye@funai.edu.ng

Ascent Method, The Newton's Method, The Quasi Newton's Method and The Conjugate Direction Method [see[19]]. The set back of the rate of convergence of these methods; Newton's Method, Quasi Newton's Method, Conjugate Direction Method and Steepest Ascent Method was shown in [16], using the Rosen Brook Function to evaluate the rate of convergence and the results showed that none of the methods converged in five iterations. The Ordinary Line Search Method (OLSM) which is said to have originated from Cauchy in 1894, is most frequently used technique in searching for optimum of unconstrained function. Also, the Lagrange Method,the Active Set Method and the Simplex Method are notable for solving constrained problems. [20], [21] and [23], showed that both constrained function and unconstrained function can as well be treated using the same analytical procedure of the Ordinary Least Square (OLS) method. The method can also be used to resolve equality and inequality constrained problems as introduced by [35]. A general approach for solving Quadratic programming problem with quadratic constraints was proposed by [11] in which
he considered the process of minimizing the Lagrange function through the cutting plane algorithm. The approximate dual form converges when the dual and the primal problem has equal objective function. [6] applied two algorithms for solving quadratic problem with negative eigenvalue, the problem with separable quadratic constraint and concave objective function. The initial rectangular domain was divided using the branch and bound algorithm into increasing small rectangle as the sizes tend to zero.In [7], [8], and [9], the use of branch and bound algorithm featured prominently in their propositions. [11] and [12] advocate the decomposition technique and showed that the iteration were between the primal and dual formulation of the main problem which can fail to converge except for special condition on the problem structure, especially with the non convex problems. In some cases, using the decomposition method to solve mixed integer problem was achieved by first converting the problem form and isolating the integer first from the objective function and making the dual problem to be linear integer variables, then, the sequence is solved by a serial integer linear original problem and its continuous convex quadratic primal problems. [13] proposed a method which converges to the global optimum, the process reduces the problem to the equivalence problem type by linearising the Lagrange function to achieve a function of the relaxed dual problem. However, this problem was not to be solved without the branch and bound method. The super convergent line search algorithm method of [1] searches the optimal feasible region in segments. The segments are the feasible region within which the constraints are satisfied. The estimated path of convergence

is achieved by normalizing the search direction and the optimal attained by addition or subtraction of the incremental or reduction factor for minimization and maximization respectively. The search scheme combined the segments using a convex combination to minimize the variance and information matrix component without incorporating the possible bias component and interactions. In the segmented case of the algorithm of [1], the method was shown to converge for quadratic problem with linear constraints in a number of iterations but could not attain convergence for a quadratically constrained quadratic problem due to the inability of the search scheme to evaluate the step length. Although an assumed step length was proposed to solve the problem but it was not iteratively efficient because the conditions of assumption were not clearly stated. In this present study, interest is to seek a method that will circumvent these rigorous means to achieve the optimal of a quadratic constrained quadratic function. This work will present a modified iteration method to resolve the quadratically constrain quadratic optimization function that is capable of evaluating the optimal starting point, the step length and the optimal direction of search that iteratively evaluates the optimal design to advance the super convergent search algorithm of the linear constrained problem to the iterative method of solving quadratic constrained quadratic optimization method by incorporating the bias and the precision matrices as the complete information matrix.

2 Methodology

2.1 The Quadratically Constrained Quadratic Problem

A quadratic constrained quadratic optimization problem is defined as

$$\begin{aligned} \underset{x}{min/max} f(.) &= C^T x + x^T C_B x \\ subject\ to: C^T x + x^T C_Q x &\leq b \\ x &\geq 0 \end{aligned} \tag{2.1}$$

where, $x \in R^n$, $C_B \in R^{n\times n}$, $C_Q \in R^{m\times n}$, $C \in R^n$, $b \in R^m$ and $x^T C_B x$ is the quadratic part of the objective function. C_B and C_Q are positive semi definite square matrices of real numbers. C_B is the coefficient of the quadratic part of the objective function. x is a subset of real numbers n. $C^T x$ is the liner part of the constraint equation and objective function. $x^T C_Q x$ is the Jacobian of the constraint.

2.1.1 The Algorithm

The algorithm can be summarized as

i) Select boundary optimal points from the feasible region of the segmented design space to form design and bias matrices and calculate the weighted means

ii) Evaluate the Average information matrix

iii) Find the optimal direction of search

iv) Calculate the value of the step length

v) Using the line equation evaluate the optimal point

vi) Check for convergence, if yes , stop.

vii) If No, replace the support point of maximum variance with the optimal value of the iteration and repeat the iteration until convergence (vi) is achieved.

2.1.2 Optimum Number of Segments and support points

To evaluate the support points, a graph of the constraint equation in (2.1) would give a picture of the design space and the feasible region especially in an inequality problem. The feasible area can be segmented into non overlapping segment, S_k. The reason for segmentation of the feasible region is to ensure a total search of the feasible area and the reduction of the mean square error of the search scheme. The required optimum number of segments is as given in [1], [16], and [22] is transformed in terms of the number of parameters of the design as

$$\frac{2p}{p^2+p+1} \leq S_k \leq \frac{p^2+p+2}{2p} \tag{2.2}$$

where, p is the number of parameters in the design.
We define S_p as the optimum number of support point per segment that is appropriate for a search; such that,

$$(n+1) \leq S_p \leq \frac{1}{2}n^2+n+1 \tag{2.3}$$

where, $n = p - 1$ and p are as defined in (2.2). It was showed that segmentation of the space X into S subspace such that, in the S_k subspace the function is adequately represented by a first order function as a rapid way to reduce the information matrix and achieve global optimum of the response surface.

2.1.3 Design Measure

A design measure refers to the measures or points that satisfies the constraint equations and lies within the feasible region. The points are made up of optimal support points, x_{ijk}, from all the segments, We elect points that satisfy the constraint equation of the design within the S_k segment. It is often preferred to use boundary points. In [15], the boundary points were shown to converge faster and more efficiently since it covers the whole design space at the same time. Care should be taken not to use the saddle point $(0, 0)$ when using weighted mean as a means of averaging the design measures, (Although, the saddle point does not have any effect when using other methods such as the Arithmetic mean,the Geometric mean and the harmonic mean. The measures of all the optimum design points in the design space are represented by

$$\xi_{N(N\times p)} = \begin{pmatrix} \xi_{\mathbf{1}} \\ \xi_{\mathbf{2}} \\ \vdots \\ \xi_{\mathbf{N}} \end{pmatrix}$$

where $\xi_{N(N\times p)}$ represents the measures of all the optimum design points in the design space and ξ are column vectors

2.1.4 Design Matrix

In each of the segments, the boundary points form the design measure denoted by x_{ijk} . for the $k-th$ segment corresponding to $\xi_{\mathbf{i}}$. For a $k-th$ segment case, the matrix of the first segment is given as

$$X_1 = Xij1 = \begin{pmatrix} x_{111} & x_{121} & \cdots x_{1p1} \\ x_{211} & x_{221} & \cdots x_{2p1} \\ \vdots & \vdots & \vdots \\ x_{n11} & x_{n21} & \cdots x_{np1} \end{pmatrix}. \tag{2.4}$$

the second segment is

$$X_2 = Xij2 = \begin{pmatrix} x_{112} & x_{122} & \cdots x_{1p2} \\ x_{212} & x_{222} & \cdots x_{2p2} \\ \vdots & \vdots & \vdots \\ x_{n12} & x_{n22} & \cdots x_{np2} \end{pmatrix}. \tag{2.5}$$

while the $k-th$ segment will be given as

$$Xk = X_{ijk} = \begin{pmatrix} x_{11k} & x_{12k} & \cdots x_{1ps} \\ x_{21k} & x_{22k} & \cdots x_{2ps} \\ \vdots & \vdots & \vdots \\ x_{n1k} & x_{n2k} & \cdots x_{nps} \end{pmatrix} \tag{2.6}$$

where $x_{i(j \times n)} = (x_1, x_2, \cdots, x_n)$. $S = 1, 2, \cdots, k$. $X_{n \times p \times k}$ is the design matrix of the segment, such that, $i = 1, 2, ..., n.j = 1, 2, ..., p.$ and $S = 1, 2, ..., k.$ and $X_1, X_2, \cdots X_k$ are design matrices for 1, 2 to k segments such that n is the row, p is the column and k is the segment.

2.1.5 Optimal Starting Point

This is a point where the search begins often evaluated as an average of all the support points.To evaluate the mean of the support points in all the segments, use can be made of the arithmetic mean, the harmonic mean or the weighted mean. The weighted mean is most preferred because it assigns weights to the design measures X_{ij} of the design matrix. The weighted mean is given as

$$W_i = \frac{a_i^{-1}}{\sum_{i=1}^{N} a_i^{-1}} \tag{2.7}$$

where, $a_i = \mathbf{x_i^T x_i}$, and $\sum_{i=1}^{N} W_i \mathbf{x_{ij}^T}$ is the optimal starting point shown as the weighted mean of $\bar{X}_j$. The optimal starting point of a design is vital in forming the design, and it affects the rate of convergence of the search. The method of selecting the best starting points is usually of great concern, because, it is the base on which the search scheme is built. [21] confirmed that once an optimal starting point $\bar{X}_j$ is achieved, the search scheme is left with evaluating the incremental value for maximization or the decremental value for minimization case respectively. The starting point of the sequence is known to affect performance and should be given serious consideration for an optimal design of experiment. The weighted mean is effective in forming optimal starting point. The reasons for performing an experiment have been to generate the distribution of information throughout the region of interest[see [30], and [31]]. This is so, because within the feasible region lies the optimal point. It is worthy of note without loss of generality, that all feasible points are local optimal points. The boundary points are effective for forming the designs for segmented and non-segmented designs.

2.1.6 The Coefficient of Bias and the Bias Vector and Bias Matrix

The coefficient of bias, C_B, is defined as the coefficient of the quadratic component and the interaction of the objective function in (2.1). An expanded expression of the objective function for p variables is given as

$$\underset{x}{\text{Min/Max}}\ f(.) = a_{11}x_1^2+\cdots+a_{pp}x_p^2+b_{12}x_1x_2+\cdots+b_{(p-1)(p)}x_{p-1}x_p+c_1x_1+\cdots+c_px_p \quad (2.8)$$

where, $a_{11},\cdots,a_{pp}$ are the coefficients of the quadratic terms, $b_{12}\cdots,b_{(p-1)p}$ are the coefficients of the interaction terms, and $c_1,\cdots c_p$ are the coefficients of the linear terms (variables). The expression (2.8) can be reduced for two variables case as,

$$\underset{x}{\text{Min/Max}}\ f(.) = ax_1^2 + bx_2^2 + cx_1x_2 + \beta x_1 + \theta x_2 + \gamma x_0. \quad (2.9)$$

The coefficient of bias of the function is

$$C_B = (a, b, c)$$

where,
a is the coefficient of x_1^2, b is the coefficient of x_2^2, c is the coefficient of x_1x_2 the interaction term. The bias matrix of each of the partitions of the design is given as X_{iB} for the $i-$th partition. In each segment, the bias matrix is obtained by using the powers and interactions of the entries of the objective function such that

$$X_{iB} = [x_{i1}^2, x_{i2}^2, x_{i1}x_{i2}] \quad (2.10)$$

where,
x_{i1}^2; is the first squared factor of the equation; x_{i2}^2; is the second squared factor of the equation and $x_{i1}x_{i2}$ is the (interaction) of the first factor and the second factor of the equation. The minimax property of the response function can be express by considering a response function

$$f(.) = g_1(x) + g_2(x) \quad (2.11)$$

Assuming that the regression function is given as

$$y(x) = g(x) + \varepsilon \quad (2.12)$$

and ε is normally distributed with mean 0; then,

$$E(g(x)) = y(x) \neq f(.)$$

The regression function (2.12) is said to be bias in (2.11), and shows that g_1 and g_2 are n component vectors of the gradient and the biasing effect respectively. The least square estimate of the regression equation in (2.12) is given as

$$E(\hat{g}(x)) = (X^TX)^{-1}X^TY, \tag{2.13}$$

If $Y = X_BC_B$, then (2.13) is given as

$$E(\hat{g}_2(x)) = (X^TX)^{-1}X^TX_BC_B.$$

Hence,

$$E(f(.)) = \hat{g}(x) + (X^TX)^{-1}X^TX_BC_B$$

It should be noted that

$$E(f(.)) = \Delta f$$

where,
Δf is the gradient of the function, and
$\hat{g}(x)$ is the precision matrix given as $(X^TX)^{-1}$; and the true estimate of (2.11) is

$$E(f(.)) = (X^TX)^{-1} + (X^TX)^{-1}X^TX_{iB}C_B. \tag{2.14}$$

where,
C_B and X_{iB} are as defined in (2.9) and (2.10) respectively.

$$E(f(.)) \geq \hat{b}$$

where,
$\hat{b}$ is the bias in each segment. If the design is segmented into k segments, then $\hat{b}$ will be represented as $\hat{b}_k$, such that

$$\hat{b}_k = (X_k^TX_k)^{-1}X_k^TX_{iB}C_B.$$

where, $k = 1, 2, \cdots, S$ provided in (2.4) to (2.6).

2.1.7 The Mean Square Error

In the presence of the bias, the response function for the kth segment in (2.12) can be restructured (see, [16], [1], and [23]), as

$$f(.) = X_kC_k + X_{Bk}g_2$$

where, g_2 is the vector of biasing effect, X_k is the design matrix for the $k-th$ segment, X_{Bk} is the coefficient matrix of bias effects for the $k-th$ segment. $Min(M(dt)) = min_{hk}(M(\sum h_{tk}C_{tk}))$, $Min(M(dt)) = min_{htk}(M(\sum h_{tk}^2M(C_{tk}))$, $M(C_{tk})$ is the mean square error of C_{tk} the diagonal element of the $M(C_{tk})$. It follow from the relation shown in (2.14) that the mean square error for (2.13) is given as

$$M(C_tk) = M_{(\varepsilon_kNk)}^{-1} + \left[M_{(\varepsilon_kNk)}^{-1} + X_k^TX_{Bk}g_2\right]\left[g_2^TX_{Bk}^TX_k + M_{(\varepsilon_kNk)}^{-1}\right]^T \tag{2.15}$$

where,
$g_2^T X_{Bk}^T X_k + M^{-1}_{(\varepsilon_k Nk)} = b_k^T$, $M^{-1}_{(\varepsilon_k Nk)} + X_k^T X_{Bk} g_2 = b_k,\ M(C_t k) = MSE,\ M^{-1}_{(\varepsilon_k Nk)}$ is the precision matrix; such that,

$$M(C_t k) = M^{-1}_{(\varepsilon_k Nk)} + b_k b_k^T, \tag{2.16}$$

$$\varepsilon_k = \begin{pmatrix} x_1 & x_2 & \cdots & x_{Nk} \\ w_1 & x_2 & \cdots & w_{Nk} \end{pmatrix},$$

and

$$N = \sum_{k=1}^{S} N_k.$$

W_i is the $i-$th weight defined in (2.7), y_k is the regression defined in (2.12), such that $g(x) = C_k^T x_k$; and x_k are the $N - th$ support point in ε_{Nk}.

2.1.8 The Information Matrix

The Fisher information is a measure of the amount of information about parameters provided by experimental data. [37] showed that in Fisher 1912, it is a well-established characteristic of an experimental design used to asses and optimize the design for maximization of expected accuracy of parameter estimates. [38] the Fisher information is calculated for each pair of parameters and it is in this notation denoted as information matrix, M . In some cases when the information matrix is not of full rank, the precision matrix M^{-1} is usually difficult to obtain. This problem is usually overcome through matrix operation and linear algebra (see [36]). Given a convex design space $S(x)$, the set of all non-singular $n \times n$ matrices $S^{n \times n}$ such that $X \in S(x)$ and $S(x)$ is convex if $x \in S(x)$ is a convex combination of the boundary set $S(b)$, then

$$S_b = \{x : x^T A x = r^2\}$$

such that $x = Hx_1 + (1 - H)x_2$.
Let $M_x = \sum_{x \in S(x)} x x^T \delta_e^2$, ie,

$$M_x = \sum_{x_1, x_2 \in S(b)} ((Hx_1 + (1 - H)x_2)(Hx_1 + (1 - H)x_2))^T \delta_e^2$$

If $x_1 x_2^T = x_2 x_1^T = 0$ and $x_1 x_1^T \neq x_2 x_2^T \neq 0$,

$$M_x \leq H^T M_1 H + (1 - H)^T M_2 (1 - H)$$

where,
M_x is the information matrix, S_y is the boundary set, S_x is the convex set of x_i^s and S_n is the function space. [26] showed the conditions and proof for concave case showing that the sign of the inequality will change when the design is segmented into S segments as shown in (2.2) to obtain the average information matrix M_A of the whole design space by the convex combination of the information of the respective segments

$$M_A = H_1^T M_1 H_1 + H_2^T M_2 H_2 + \cdots + (1 - H_S)^T M_s H_s \tag{2.17}$$

The diagonal elements in (2.17) of M_A are the variances and the off diagonal are the covariances.

2.1.9 Convex Matrix

The convex combination for two segment design can be shown by considering an optimization problem given as

$$\begin{aligned} Optimize\ f(x_1, x_2, \cdots, x_n) \\ Subject\ to : g(x) \end{aligned}$$

The constraint equation $g(x)$, is graphed and segmented to evaluate M_{ck1}, M_{ck2} , $\cdots$, M_{cks} for k_{th} segments as shown in (2.16), the convex combination matrix H_1, H_2, $\cdots$, $1 - H_k$ is as defined in (2.17)
such that for a two variable case,

$$d = H_1 b_1 + (1 - H_1) b_2$$

expressed as

$$\begin{aligned} d_0 &= h_0 b_{10} + (1 - h_0) b_{20} \\ d_1 &= h_1 b_{11} + (1 - h_1) b_{21} \\ d_2 &= h_2 b_{12} + (1 - h_2) b_{22} \end{aligned}$$

The variance of the relation in (2.4), to (2.6) can be evaluated as

$$var(d_i) = h_i^2 var(b_{1i}) + (I - h_i)^2 var b_{2i}$$

Taking a partial derivative ∂_i with respect to h_i and solving for h_i, for $i = 0, 1, 2$, would yield

$$\begin{aligned} h_0 &= \frac{v_{20}}{v_{10} + v_{20}} \\ h_1 &= \frac{v_{21}}{v_{11} + v_{21}} \\ h_2 &= \frac{v_{22}}{v_{11} + v_{22}} \end{aligned}$$

$H_{n\times n} = diag(h_0, h_1, h_2)$ and H is a symmetric matrix
where,

$$H^T H + (1-H)^T(1-H) = 1$$

The normalized convex matrices H^* and $(I-H)^*$ are achieved by dividing each diagonal entry of the symmetric matrix H by the square root of the sum of square of the corresponding entries of H and $(1-H)$ respectively, Using the same rules, the convex matrix can be extended to k segments.Thus,

$$H^* = diag\left\{\frac{h_0}{\sqrt{h_0^2+(1-h_0)^2}}, \frac{h_1}{\sqrt{h_1^2+(1-h_1)^2}}, \cdots, \frac{h_k}{\sqrt{h_k^2+(1-h_k)^2}}\right\}$$

.

2.1.10 The Direction Vector and The Direction of Search

In the response function defined in (2.12), there exists a first-order model,

$$f(.) = Xa + \varepsilon$$

Such that if $x^T x \geq 0$ of degree, $m \geq 2$ (see for instance,[16]) then

$$Z(.) = M_{\xi N} d + u(.)$$

and if $M_{\xi N} = wX^T X w^T$ and $d = wa$ then

$$Z(.) = wX^T X w^T + u(.)$$

By least square estimate

$$\hat{a} = (X^T X)^{-1} X^T f(.)$$

$$\hat{d} = w(X^T X)^{-1} X^T f(.)$$

$$\hat{d} = M_{\xi N}^{-1} \mathbf{z}(.)$$

where, w is a convex combination matrix, such that $w^T w = I$ and The direction vector $\mathbf{z}(.)$ is achieved by substituting the corresponding entries of the mean square error into the objective function of (2.1)
$\mathbf{z}(.) = \mathbf{f}(\mathbf{m_1}, \mathbf{m_2}, \cdots, \mathbf{m_n}) = (\mathbf{z_1}, \mathbf{z_2}, \cdots, \mathbf{z_n})^{\mathbf{T}}$

$$var(\hat{d}) = w(X^T X)^{-1} w^T$$

$\hat{d}$ is an $(n \times 1)$ directional column vector. To evaluate the optimal directional vector, we normalize the directional column vector $\hat{d}$ to obtain

$$\hat{d}^* = \begin{pmatrix} \frac{d_1}{\sum_i^n d_i^2} \\ \frac{d_2}{\sum_i^n d_i^2} \\ \vdots \\ \frac{d_n}{\sum_i^n d_i^2} \end{pmatrix}$$

2.1.11 The Step Length

The Step length represents the length of search in the design space given as;

$$\rho_{n\times 1} = \left\{ \frac{C_T x_j^* - b_i}{C_T \hat{d}^*} \right\}$$

where, ρ is step length, C_T is the value of the partial differential coefficient of the Jacobian of the quadratic constraint, $\mathbf{x_j^*}$ is the weighted optimal starting point, $\hat{\mathbf{d}}^*$ is the optimal direction of search, and $\mathbf{b_i}$ is constants of the constraint equation. In the case of a minimization problem, $\rho_{\mathbf{min}}$ is used to evaluate the optimal decreased (although these are usually dependent on the sign) value while for maximization problem, ρ_{max} is used to evaluate increased optimal value. The constraints coefficient are transformed by linearising the Jacobian of the constraint equation, $C^T x + x^T C_Q x$ in (1). Let C_T be defined as the partial deferential coefficient of the constraint equation given as,

$$C_T = \left[\frac{\partial (Cx)}{\partial x} + \frac{\partial (x^T C_Q x)}{\partial x} \right] = 0$$

such that,
C_T linearises the quadratic constraint see ([39])

$$\rho_{min/max} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_n \end{pmatrix} \tag{2.18}$$

The step length helps in upward or down adjustment of the search along the estimated optimal direction of search to achieve the global optimum.

2.1.12 Optimal Point

The optimal move is achieved by substituting the values of the optimal starting point, the step length and the direction of search in the line equation to evaluate x^* the optimal point. The difference for maximization problem and minimization problem are discussed as $\mathbf{x}^*$ is the value that minimizes or maximizes the $f(.)$ and satisfies the constraint equations

$$\mathbf{x}^* = (\mathbf{x_j^*} - \rho_{\mathbf{min}}\hat{\mathbf{d}}) \tag{2.19}$$

for minimization problem and

$$\mathbf{x}^* = (\mathbf{x_j^*} + \rho_{\mathbf{max}}\hat{\mathbf{d}}) \tag{2.20}$$

for maximization problem. where, $\rho_{max}\hat{d}$ is increment value that maximizes the optimal stating point $\mathbf{x_j^*}$ while $\rho_{min}\hat{d}$ is decrement value that minimizes the optimal point starting point.

2.1.13 Variance Replacement

This is a procedure used to eliminate the design support point $\begin{pmatrix} x_{1j} \\ x_{2j} \end{pmatrix}$ with the maximum variance and replace it with the algorithm improved point $\begin{pmatrix} x^*_{1j+i} \\ x^*_{2j+i} \end{pmatrix}$ where, $i = 1, 2, ...$

The variance is evaluated as

$$Var(x) = \mathbf{x_i M_A}^{-1}\mathbf{x_i^T} \tag{2.21}$$

where, $\mathbf{x_i}$ are the row vectors of the support points of all the segments and ${M_A}^{-1}$ is the precision matrix of the design respectively. [15] demonstrated that by improving an existing experimental design the optimizer of the response function is approached by adding points of minimal variance to an initial design leading to the minimum of the response function for a minimization problem and otherwise adding for maximization problem. The minimum variance adjustment technique is introduced as a means of reducing variability and attaining convergence

2.1.14 Convergent Criteria

Let $\mathbf{x^*_{j+i}} = \begin{pmatrix} x^*_{1j+i} \\ x^*_{2j+i} \end{pmatrix}$ such that $\mathbf{x^*_{j+i}}$ be the iteration estimated optimal value at the i^{th} iteration given in either (2.19) or (2.20). If $\epsilon \in [0, 0.01)$ then the sequence is said to have converged if

a) $\mid f(x_j^*) - f(x_{j+1}^*) \mid \leq \epsilon$

b) If $\begin{pmatrix} x_{1j+i}^* \\ x_{2j+i}^* \end{pmatrix}$ satisfies all the constraint equation in (2.1).

c) The sequence is said to have converged, if sequence satisfies all the convergent criteria stated in number (a) and (b) above then $\begin{pmatrix} x_{1j+i}^* \\ x_{2j+i}^* \end{pmatrix}$ is a global optimum. Otherwise, do a variance replacement and repeat the iteration.

3 Application

Two numerical examples are considered, the first problem is taken from [29] as F5, also available in [4] and [28]. The second problem is taken from [1] as standard quadratically constrained optimization problems.

3.1 Example 1

Let us consider the problem taken from [29] as F5, available in [4] and [28] given as;

$$\begin{aligned} min\ f(x) = (x_1 - 1)^2 + (x_2 - 1)^2 \\ subject\ to: x_1 + x_2 \leq 2 \\ x_2^2 - x_1 \leq 0 \end{aligned}$$

The feasible area is partitioned into two segments of $X1$ and $X2$ as the first and second segments are given as

$$X_1 = \begin{pmatrix} 1.00 & 1.0000000 \\ 0.25 & 0.5000000 \\ 0.50 & 0.7071068 \\ 6.25 & 2.5000000 \end{pmatrix} .\text{and} \quad X_2 = \begin{pmatrix} 1.0000 & 1.000000 \\ 1.2081 & 0.790800 \\ 3.2500 & 1.802776 \\ 0.2500 & -0.500000 \end{pmatrix}.$$

The bias matrices for the two segments are evaluated as

$$X_{1B} = \begin{pmatrix} 1.0000 & 1.0000 \\ 0.0620 & 0.2500 \\ 0.2500 & 0.5000 \\ 39.0620 & 6.2500 \end{pmatrix} .\text{and} \quad X_{2B} = \begin{pmatrix} 1.0000 & 1.0000 \\ 1.460 & 0.625 \\ 10.562 & 3.250 \\ 0.062 & 0.250 \end{pmatrix}.$$

The coefficient of bias of the function is

$$\mathbf{C_B} = \begin{pmatrix} 1.0000 & 1.0000 \end{pmatrix}.$$

The partial deferential coefficient of the constraint equation C_T is evaluated as

$$C_T = \begin{pmatrix} 1.0000 & 1.0000 \\ -1.0000 & 2.0000 \end{pmatrix}.$$

The right hand side constant of the constraint equation is

$$\mathbf{b} = \begin{pmatrix} 2.0000 \\ 0.0000 \end{pmatrix}.$$

The information matrices of the first and the second segments are obtained as

$$X_1^T X_1 = \begin{pmatrix} 40.37500 & 17.10355 \\ 17.10355 & 8.000000 \end{pmatrix} \text{ and } \quad X_2^T X_2 = \begin{pmatrix} 13.084506 & 5.689386 \\ 5.689386 & 5.125365 \end{pmatrix}.$$

The bias vectors are calculated as

$$\mathbf{b_1} = \begin{pmatrix} 9.905 \\ -6.680 \end{pmatrix} \text{.and} \quad \mathbf{b_2} = \begin{pmatrix} 3.310 \\ 1.084 \end{pmatrix}.$$

and Bias for segment the two segments are

$$b_1 b_1^T = \begin{pmatrix} 98.10902 & -66.1654 \\ -66.1654 & 44.6224 \end{pmatrix} \text{.and} \quad b_2 b_2^T = \begin{pmatrix} 10.95610 & 3.588040 \\ 3.588040 & 1.175056 \end{pmatrix}.$$

The mean square errors are evaluated as

$$M_1 = \begin{pmatrix} 98.37 & -66.73 \\ -66.73 & 45.95 \end{pmatrix} \text{.and} \quad M_2 = \begin{pmatrix} 11.10 & 3.42 \\ 3.42 & 1.55 \end{pmatrix}.$$

The convex matrices are evaluated as

$$H_1 = diag.\begin{pmatrix} 0.8986024, & 0.9673684 \end{pmatrix}.$$

and

$$H_2 = diag.\begin{pmatrix} 0.1013976, & 0.03263158 \end{pmatrix}.$$

and the normalized convex matrices

$$H_1^* = diag.\begin{pmatrix} 0.9936938, & 0.9994316 \end{pmatrix}.$$

and

$$H_2^* = diag.\begin{pmatrix} 0.1121277, & 0.03371314 \end{pmatrix}.$$

Average information matrix is calculated be

$$M_A = \begin{pmatrix} 97.27279 & -66.25857 \\ -66.25857 & 45.89954 \end{pmatrix}.$$

The direction vector and the normalized direction of search are calculated as

$$\mathbf{d} = \begin{pmatrix} 14303.11 \\ 20787.88 \end{pmatrix}.$$

and

$$\mathbf{d}^* = \begin{pmatrix} 0.5668365 \\ 0.8238303 \end{pmatrix}.$$

The weighted X are

$$\bar{x}_j = \begin{pmatrix} 0.4535172 \\ 0.1621163 \end{pmatrix}.$$

The Step lengths ρ and minimum step length ρ^* for minimization problem is evaluated from respectively as

$$\mathbf{p} = \begin{pmatrix} -0.9954696 \\ -0.1196167 \end{pmatrix}.$$

Optimal Step length

$$\mathbf{p}^* = \mathbf{-0.9954696}$$

Optimal value is

$$\mathbf{x}^* = \begin{pmatrix} 1.0177857 \\ 0.9822143 \end{pmatrix}.$$

3.1.1 Convergence criteria

To evaluate the Convergent Criteria, the optimal starting point $\bar{x}_j$ and the optimal value $\mathbf{x}^*$ are used to evaluate the convergent criteria,$f(\bar{x}_j)$ as f_0 and $f(\mathbf{x}^*)$ as f_1. being function of objective function. To check if the solutions satisfies the constraint equations,
$\mathbf{x}^*$ is substituted in to the constraint equations.
Convergent Criteria 1
let $Q = |(f(\bar{x}_j) - f(\mathbf{x}^*))| < \epsilon$
where ϵ is a small value $0 < \epsilon < 0.01$, then,

$$Q = |(f_0((0.4535172, 0.1621163) - f_1(1.0177857, 0.9822143)| < 0.01$$

Q is not TRUE and $\mathbf{x}^*$ does not satisfies the objective function convergent condition at $\epsilon = 0.01$ and $\mathbf{x}^*$ = (x_1= 1.0177857, x_2 =0.9822143).
Convergent Criteria 2
To test with the constraints equations, $\mathbf{x}^*$ is substituted to check if the optimal

values $\mathbf{x}^*$ satisfies all the constraint equations. The results obtained showed that the values of $\mathbf{x}^*$ = (x_1= 1.0177857, x_2 = 0.9822143) satisfies the constraint equations. $\mathbf{x}^*$ is not considered to be a global solution because first
convergence criteria is not satisfied. Hence, a variance replacement is done to evaluate the support point with maximum Variance and replace it with $\mathbf{x}^*$. and Repeat the Iteration. The Maximum Variance is evaluated as Vm =59.95698 corresponding to the variance of the fourth support point V4. The fourth support point is replaced with (1.0177857,0.9822143) and the iteration is repeated. The iteration was repeated five more times before convergence was attained in the sixth iteration. the maximum variance were evaluated at V4,V7,V1,V4 and V1 corresponding to support points X4,X7,X1,X4 and X1 which was replaced with the algorithm evaluated optimal points X^* as (1.0678,0.9322),(1.0436,0.9564),(1.0539,0.9461), (1.0622,0.9378) and(1.0666,0.9334) evaluated at each iteration respectively. see Table 1 below;

Table 1: Result of Six Iterations

IT	p^*	d^*	$\bar{x_j}$	$\mathbf{x}^*$	Q	Vm	Rule 1	Rule 2
1	-0.9955	0.5668,0.8238	0.4535,0.1621	1.0178,0.9822	2.6403	-	False	True
2	-.09468	0.6323,0.7747	0.4691,0.1987	1.0678,0.9322	0.1087	59.9570	False	True
3	-0.9257	0.6103,0.7922	0.4787,0.2230	1.0436,0.9564	0.0538	4.5780	False	True
4	-.09245	0.5200,0.7846	0.4807,0.2208	1.0539,0.9461	0.0224	1.3441	False	True
5	-0.9235	0.6278,0.7784	0.4824,0.2189	1.0622,0.9378	0.0187	1.3277	False	True
6	-0.9232	0.6320,0.7750	0.4832,0.2179	1.0666,0.9334	0.0099	1.3093	True	True

3.2 Example 2

consider the optimization problem from [1] considered as a General Constrained Optimization Problem. Although, the problem set assumed the optimal step length, we shall run the iteration with the algorithmic evaluated step length and direction of search using the modified coefficients of the constraint equation.

$$
\begin{aligned}
min\ f(x) = x_3 - \frac{1}{2}{x_2}^2 \\
subject\ to : x_1^2 + x_2 + x_3 \leq 7 \\
x_1^2 - x_2 + x_3 \leq 5 \\
x_3 \geq 0
\end{aligned}
$$

Let the area be partitioned into two and the boundary points that satisfies the constraint equation given as;

$$X_1 = \begin{pmatrix} 2.44949 & 1 & 0.00 \\ -2.44949 & 1 & 0.00 \\ 2.00000 & 2 & 0.00 \\ -2.00000 & 0 & 0.00 \\ 2.00000 & 0 & 0.01 \\ -2.00000 & 0 & 0.01 \end{pmatrix}.$$

for Segment two

$$X_2 = \begin{pmatrix} 2.236068 & 1 & 0.01 \\ 2.236068 & 1 & 0.01 \\ 1.732051 & 3 & 0.01 \\ 1.732051 & 2 & 0.01 \\ 2.236068 & 0 & 0.00 \\ 2.236068 & 0 & 0.00 \end{pmatrix}.$$

The bias matrix for the first segment

$$X_1B = \begin{pmatrix} 4 \\ 4 \\ 6 \\ 3 \\ 6 \end{pmatrix}.$$

and

$$X_2B = \begin{pmatrix} 5 \\ 3 \\ 6 \\ 4 \\ 4 \end{pmatrix}.$$

The coefficient of Bias

$$C_B = \left(-\frac{1}{2}\right).$$

The differential of the Jacobian matrix

$$C_T = \begin{pmatrix} 2 & 1 & 1 \\ 2 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

The Right hand side constant

$$b = \begin{pmatrix} 7 \\ 5 \\ 0 \end{pmatrix}.$$

The Bias Vector and bias are calculated as

$$b_1 = \begin{pmatrix} 0.00 \\ -1.40 \\ -200.0 \end{pmatrix}.$$

The Bias Vector and bias are evaluated as

$$b_2 = \begin{pmatrix} -1.1180 \\ 0.4365 \\ -43.6492 \end{pmatrix}.$$

and

$$b_1 b_1^T = \begin{pmatrix} 0 & 0.00 & 0 \\ 0 & 1.96 & 280 \\ 0 & 280.00 & 40000 \end{pmatrix}.$$

$$b_2 b_2^T = \begin{pmatrix} 1.249924 & -0.4880070 & 48.79981 \\ -0.488007 & 0.1905322 & -19.05288 \\ 48.799806 & -19.0528758 & 1905.25266 \end{pmatrix}.$$

The mean Square for first segment

$$M_1 = \begin{pmatrix} 0.0357 & 0.00 & 0 \\ 0.0000 & 2.06 & 280 \\ 0.0000 & 280.00 & 45000 \end{pmatrix}.$$

the mean square error is calculated as

$$M_2 = \begin{pmatrix} 1.3499 & -0.4376 & 21.3990 \\ -0.4376 & 1.2159 & -182.8634 \\ 21.3990 & -182.8634 & 34413.3193 \end{pmatrix}.$$

The Convex matrices and the normalized convex combination is given as

$$H_1 = diag.\left(0.02576501, \quad 0.6288348, \quad 0.5666556\right).$$

and

$$H_2 = diag.\left(0.974235, \quad 0.3711652, \quad 0.433444\right).$$

and the normalized convex matrices are

$$H_1^* = diag.\left(0.02643716, \quad 0.8611778, \quad 0.7943446\right).$$

and

$$H_2^* = diag.\left(0.9996505, \quad 0.5083039, \quad 0.6074674\right).$$

Average information matrix

$$M_A = \begin{pmatrix} 1.348981 & -0.222356 & 12.99465 \\ -0.222356 & 1.841907 & 135.07593 \\ 12.994652 & 135.075925 & 41093.33834 \end{pmatrix}.$$

The direction vector

$$F_z = \begin{pmatrix} 12.08478 \\ 135.05120 \\ 41008.90785 \end{pmatrix}.$$

The direction of search and the optimal direction of search are calculated as
The direction of search is

$$d = \begin{pmatrix} -0.64516711 \\ 0.05851019 \\ 0.99795709 \end{pmatrix}.$$

and

$$d^* = \begin{pmatrix} -0.54225635 \\ 0.04917721 \\ 0.83877271 \end{pmatrix}.$$

The optimal weighted mean as optimal starting point is obtained as

$$\bar{x}_j = \begin{pmatrix} 1.038584 \\ 0.822400 \\ 0.005446 \end{pmatrix}.$$

Step length p and minimum step length p^* for minimization is evaluated

$$\mathbf{p} = \begin{pmatrix} 20.83297 \\ 12.6808 \\ 0.00649282 \end{pmatrix}.$$

optimal step length for minimization is

$$p^* = 0.00649282$$

Optimal value is calculated as

$$\mathbf{x}^* = \begin{pmatrix} 1.0421049 \\ 0.8220807 \\ 0.000000 \end{pmatrix}.$$

3.2.1 Convergence Criteria

To evaluate the Convergent Criteria, the optimal starting point $\bar{x}_j$ and the optimal value $\mathbf{x}^*$ are used to evaluate the convergent criteria,$f(\bar{x}_j)$ as f_0 and $f(\mathbf{x}^*)$ as f_1. being function of objective function. To check if the solutions satisfies the constraint
equations,$\mathbf{x}^*$ is substituted in to the constraint equations.
Convergent criteria 1
let $Q = |(f(\bar{x}_j) - f(\mathbf{x}^*))| < \epsilon$
where ϵ is a small value $\epsilon \in [0, 0.01)$, then,

$$Q = |(f_0((1.038584, 0.822400, 0.005446) - f_1(1.0421049, 0.8220807, 0.0000000)| < 0.01$$

Q is TRUE and $\mathbf{x}^*$ satisfies the objective function convergent condition of $\epsilon \in [0, 0.01)$ and $\mathbf{x}^*$ = (x_1= 1.0421049, x_2 = 0.8220807 and x_3=0.000000).
Convergent Criteria 2
To test with the constraints equations, $\mathbf{x}^*$ is substituted to check if the optimal values $\mathbf{x}^*$ satisfies all the constraint
equations. The results obtained also showed that the values of $\mathbf{x}^*$ = (x_1= 1.0421049, x_2 = 0.8220807 and x_3=0.000000) satisfies the constraint equations in the first iteration. Hence, $\mathbf{x}^*$ is considered a global solution. There will be no need for another iteration.

4 Conclusion

A modified super convergent line search algorithm for solving quadratically constrained quadratic optimization problem has been developed. The two analytical problems showed the validity of the method in solving quadratically constrained problems. The approach is simple, straight forward and can be implemented for solving various quadratically constrained quadratic optimization problems in the fewest number of iterations.

References

[1] Onukogu, I. B. and Chigbu, P .E. (2002). Super Convergent Line Series in Optimal Design of Experiments and Mathematical Programming. AP Express Publishers, Nsukka.

[2] I. Etukudo, Optimal designs technique for locating the optimum of a second order response function, American Journal of Operations Research, 7(05), (2017).

[3] H. W.Kuhn, Nonlinear programming: a historical view, Traces and emergence of nonlinear programming, 393-414, (2013).

[4] J. Bracken and G. P. McCormick, Selected Applications of Nonlinear Programming. Wiley, Hoboken,NJ,USA.(1968).

[5] X. Wu, Optimal designs for segmented polynomial regression models and web-based implementation of optimal design software. State University of New York at Stony Brook,(2007).

[6] P. M. Pardalos and S. A. Vavasis, Quadratic programming with one negative eigenvalue is NP-hard, Journal of Global optimization, 1, 15-22, (1991).

[7] J. E. Falk and R. M. Soland, An algorithm for separable nonconvex programming problems, Management science, 15(9), 550-569 (1969).

[8] R. M. Soland, An algorithm for separable nonconvex programming problems II: Nonconvex constraints, Management Science, 17(11), 759–773 (1971).

[9] G. R. Reeves, Nonconvex all-quadratic programming, Washington University, U.S.A.(1973),

[10] G, R. Reeves, Global minimization in nonconvex all-quadratic programming, Management Science, 22, 76-86, (1975)

[11] D. P. Baron, Quadratic programming with quadratic constraints, Naval Research Logistics Quarterly, 19(2),253–260, (1972).

[12] R. Lazimy, Improved algorithm for mixed-integer quadratic programs and a computational study, Mathematical Programming, 32, 100-113, (1985).

[13] C.A. Floudas and V. Visweswaran, Quadratic optimization. Handbook of global optimization, 217–269(1995).

[14] M. P. Iwundu and S. U. Ekpo, Rules for Identifying the Initial Design Points for Use in the Quick Convergent Inflow Algorithm, International Journal of Statistics and Probability, 5, (2016).

[15] M. P. Iwundu and J. E. Hezekiah, Algorithmic Approach to Solving Linear Programming Problems on Segmented Regions, Asian Journal of Mathematics & Statistics, 7(2),(2014)

[16] I. B. Onukogu, Foundations of optimal exploration of response surfaces. Ephrata, Nsukka,(1997).

[17] W. Zhou, Computing optimal designs for regression models via convex programming, Department of Mathematics and Statistics, University of Victoria, (2015).

[18] V. Boyds, Convex optimization. Cambridge University Press, UK, (2004).

[19] H. A. Taha, Operations research: an introduction. Pearson Education, India, (2013).

[20] R. Fletcher, Practical methods of optimization. John Wiley & Sons, New York, U.S.A, (2000).

[21] I. B. Onukogu, Global Convergent Sequence for Solving Optimization Problems, Journal of Statistics Applications & Probability, 2(3), 243-250(2013).

[22] A. Pázman, Foundations of optimum experimental design, (1986).

[23] T. A. Ugbe and S. S. Akpan, On Optimal Solution of Boundary and Interior Support Points for Solving Quadratic Programming Problems Through Super Convergent Line Series, Journal of Computations & Modelling, 4(2), 79–92(2014).

[24] R. Fletcher, A nonlinear programming problem in statistics (educational testing), SIAM Journal on Scientific and Statistical Computing, 2(3), 257–267(1981).

[25] C. R. Rao and K. Mitra, Generalized inverse of a matrix and its applications. Proceedings of the sixth Berkeley symposium on mathematical statistics and probability, 601-620 (1972).

[26] R. M. Freund , Quadratic functions, optimization, and quadratic forms, Massachusetts Institute of Technology (2004).

[27] J. Bracken and G. P. McCormick, Selected applications of nonlinear programming. Wiley New York, (1968)

[28] W. Hock and K. Schittkowski, Usage of the Fortran Subroutines. Test examples for nonlinear programming codes, 12 -16(1981).

[29] P. A. Simionescu, A collection of bivariate nonlinear optimisation test problems with graphical representations, International Journal of Mathematical Modelling and Numerical Optimisation, 10(4), 365-398(2020).

[30] G. E. P. Box and N. R. Draper, Robust designs, Biometrika, 62(2), 347-352 (1975).

[31] A. Atkinson, Optimum experimental designs. Clarendon Press, New York, (1992).

[32] J. T. Betts, An accelerated multiplier method for nonlinear programming, Journal of Optimization Theory and Applications, 21(2), 137-174(1977).

[33] K. Kreutz-Delgado, The complex gradient operator and the CR-calculus, arXiv preprint arXiv:0906.4835(2009).

[34] A. C. Miguel, Lecture Notes On Numerical Optimization, EECS, University of California, Merced.(2020)

[35] H. W. Kuhn and A. W. Turker. Nonlinear Programming. In J.Neyman, edition.proc. of the Second Berkeley symposium on maths. stat and prob., california press,Berkeley California.(1951)

[36] C.R Rao and S. K. Mitra. Generalized Inverse of a Matrix and its Applications. s Sixth Berkeley Symposium on Maths. Stat. and Prob.(1972).

[37] S. Geisser. Break Through in Statistics, Foundation and Basic Theory. Springer Series in Statistics, New York. (1992)

[38] C. Kreutz and J. Timmer. Optimal experiment, Design, Fisher Information. Encyclopedia of system Biology, Springer, New York (2013)

[39] S. Poonam and K, Ashok. Optimal Design for Generalized Pareto Model.International Journal of Mathematical, Engineering and Management Sciences 4(5) 1264-1276 (2019)
