Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.12785/jsap/020307

Global Convergent Sequence for Solving Optimization Problems

Ike Basil Onukogu[*]

Department of Mathematics/Statistics, University of Uyo, Nigeria

Received: 22 Aug. 2012, Revised: 24 Feb. 2013, Accepted: 13 Mar. 2013
Published online: 1 Nov. 2013

Abstract: The paper presents an effective, gradient-based procedure for locating the optimum for either constrained or unconstrained response surfaces. The procedure is widely applicable covering linear and curvilinear constraints as well as constraints with redundancy. Each move of the sequence is made by pooling information from all the support points in the feasible region, making the search a global one and convergence to the global optimum almost certain.

Keywords: Optimization Problems, Line search algorithm, constrained and unconstrained, Global Convergence, Least squares–Minimum Norm.

1 Introduction

The problem considered here is one that occurs very frequently in many scientific enquiries; i.e. to locate the optimizer (specifically the minimizer) of a multivariable objective function, defined in a continuous, finite dimensional space;see, e.g. Gass (1964, page 257) for a bibliography of applications covering several areas of practical optimization problems, Dantniz (1964, page 63), Box and Wilson (1951).

The Ordinary Line Search Methods OLSM which originated from Cauchy (1894), albeit modified severally over the decades, are perhaps, the most frequently employed techniques in the search for the optimum of unconstrained functions UCF. On the other hand, the Active set, the Lagrangian, the simplex methods, etc are the favoured techniques for constrained functions CF. Interestingly, Fletcher (1981) by portraying these CF techniques as line search algorithms has shown that both the UCF and CF problems could be treated from the same analytical framework, all of which make-up the OLSM. Generally, each line search sequence is built-up from four (4) composite parts:

(a) The starting point, which is an n–component vector $\underline{\bar{x}}$, (b) An n-component direction vector $\underline{d}$, (c) Computation of the step–length ρ, and (d) A movement to the iterate, $\underline{x}\,(=\underline{\bar{x}} \pm \rho \underline{d})$.

This set of sequential activities applies as well to the new technique herein introduced, albeit with some very significant differences.

First, exploration in the OLSM is made in a sub-region $S_p \subset \tilde{X}_f$, and the starting point, $\underline{\bar{x}}$, is at the center of S_p; where, $\tilde{X}_f$ is the entire feasible region, also known as the experimental space; see, e.g. Wilde and Beightler (1967, section 7.03, page 281). But in the new series, which may be called Super Convergent Line Series SCLS, the whole space $\tilde{X}_f$ is partitioned into non-overlapping segments S, and each segment is explored separately. The starting point $\underline{\bar{x}}$, i.e. the center of the space $\tilde{X}_f$, is then obtained as the arithmetic average of the segment means. Good convergence has also been reported by Onukogu and Chigbu (2002, page 106) using the harmonic mean of the segments. However, because $\tilde{X}_f$ often has a regular geometric form the use of $\underline{\bar{x}}$ is generally satisfactory.

Second, the direction of search, under the OLSM is either along the n–component gradient vector $\underline{g}$ or along a deflected gradient $A\underline{g}$; with A as the $n \times n$ deflection matrix; see, e.g. Wilde and Beightler (1967, page 293). The gradient direction vector, $\underline{g}$ (where $A = I$ is an identity matrix), doesn't take into account the bias due to interactions and higher order effects that may be present in the response function hence the performance is generally poorer than the $A\underline{g}$ direction. In the SCLS however, the direction vector is a weighted average from all the segments; the weights being proportional

[*] Corresponding author e-mail: ikebasilonukogu@yahoo.com

to the mean square errors from the segments. It is shown in section three (3) that the SCLS direction vector has a least squares-minimum norm LSMN property.

Third, the basic idea for obtaining the step-length ρ is essentially the same for both OLSM and SCLS and this again depends on whether or not the optimizer $\underline{x}^*$, is expected to be an interior point or a boundary point. When $\underline{x}^*$, is an interior point, ρ is obtained (i.e. after obtaining $\underline{\bar{x}}$, and $\underline{d}$) by expressing the objective function as a function of ρ only and solving the derivative equation that follows. On the contrary,when $\underline{x}^*$ is a boundary point (again after getting $\underline{\bar{x}}$, and $\underline{d}$) the step-length is computed for each constraint, giving $\{\rho_k\}, k = 1,2,\ldots,K$ and $\rho = \min_k \{\rho_k\}_1^K$. Detailed computations are illustrated in section six (6) of this report.

Fourth, the SCLS is globally convergent, whereas the OLSM may fail to converge or may converge to a local optimum. After 200 iterations the gradient procedure failed to reach the minimum of the Rosenbrock function which is example 1, section 6 of this paper. The gradient procedure obtained $x_1 = -0.605, x_2 = 0.375), f(x_1,x_2) = 2.578$; see, Wilde and Beightler table 7-1, page 312.

In the OLSM, difficulties could arise if some of the constraints are redundant or curvilinear whereas no such difficulties occur with the SCLS; see Onukogu and Chigbu (2002, page 127). Redundancy occurs when the set of constraints are not linearly independent and this gives rise to the problem of inversion of matrices that are not of full rank; see, e.g. Fletcher starting from page 14. The curvilinear constraints often give rise to feasible regions that have irregular geometry and an optimizer $\underline{x}^*$, that may now be an interior point. In such a situation, sequential methods like the Simplex, Active Set, etc are likely to fail to reach the optimizer. Indeed, in the SCLS, the major role of the constraints is to ensure that the designs in the segments are feasible, and that the iterate is also feasible.

Exploring new ways to tackle old problems on optimization have continued to generate fascinating research interest: interior point algorithm by Karmarker (1984) shows a refreshing light on solution of constrained optimization, Umoren (1996) has obtained solutions to linearly constrained optimization problems using experimental design techniques. Taking all these efforts along with that of Concentric Balls technique by Onukogu (2012), one can see an emergence of a unified class of algorithms for solving optimization problems regardless of whether they occur in experimental designs, in Operations Research or in the diversity of practical problems.

2 The Sequential Steps

The problem is to find $\underline{x}^*$ the minimizer of an n–variate, m–degree polynomial objective function;

$$f(\underline{x};m) = \underline{a}'\underline{x} = \left(\underline{a}_1'\underline{x}_1 + \underline{a}_2'\underline{x}_2\right) + e; \quad m \geq 1; \tag{2.1}$$

where, $\underline{a}' = (\underline{a}_1', \underline{a}_2')$, $\underline{x} = \begin{pmatrix} \underline{x}_1 \\ \underline{x}_2 \end{pmatrix}$, are respectively, p–component row and column vectors of known parameters and input variables.

The $\underline{a}_1$, $\underline{a}_2$ are $(n+1)$ and $(p-n-1)$ vectors of first and higher effects, $\underline{x}_1$, $\underline{x}_2$ are their respective vector of input variables, e represents random observation error. For constrained optimization COP, the feasible region,

$$\tilde{X}_f \equiv \tilde{X}_c = \left\{x_1, x_2, \ldots, x_n;\ \underline{c}_k'\underline{x}_k \leq, =, \geq b_k, k = 1,2,\ldots,K\right\};$$

whereas,

for unconstrained optimization UCOP,

$$\tilde{X}_f \equiv \tilde{X}_u = \{x_1, x_2, \ldots, x_n;\ \textstyle\prod_{i=1}^n (t_{1i} - t_{0i});\ t_{0i} < x_i < t_{1i}\};$$

is a product of intervals.

In $\tilde{X}_c$, $\underline{c}_k'\underline{x}_k$ is a polynomial of degree $m_k \geq 1$, $\underline{c}_k$ is a p_k–component vector of known constants, $\underline{x}_k$ its corresponding vector of input variables, b_k, t_{0i}, t_{1i} are scalars. The $\{x_i\}$ are considered quantitative, non-stochastic variables; hence, $\tilde{X}_f$ is a continuous, finite dimensional, compact, metric space; see e.g. Pazman (1986 starting from chapter two, page 15). The sequential steps are therefore as follows:

(i) Partition $\tilde{X}_f$ into $S \geq 2$, non-overlapping segments $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_S$

(ii) Represent $f(\underline{x};m)$ in the sth segment by a first-order regression function $\underline{y}_s = X_s\underline{q}_s + \underline{e}_s$, $s = 1,2,\ldots,S$.

Thus,

$\underline{q}_s = \underline{a}_1 + X_{Bs}\underline{a}_2; \quad \underline{x} \in \tilde{X}_s$

$\underline{y}_s$, $\underline{e}_s$ are respectively, N_s–component vectors of unknown responses and random observation errors,

X_s is $N_s \times (n+1)$ design matrix; N_s is the number of support points in the sth segment, $N_s \geq n+1$

$\underline{q}_s$ is $(n+1)$–component vector of unknown parameters

X_{BS} is $N_s \times (p-n-1)$ bias matrix corresponding to $\underline{a}_2$, the bias vector in (2.1)

(iii)From (ii) above, compute $\overline{\underline{x}}_s$, the arithmetic mean vector in each segment, the grand mean, $\overline{\underline{x}} = S^{-1}\sum_{s=1}^{S}\overline{\underline{x}}_s$, the normalized information matrix, $M_s = N_s^{-1}X_s'X_s$ and the bias matrix,

$$B_S = M_S^{-1} + M_S^{-1}X_S'X_{BS}\underline{a}_2\underline{a}_2'X_{BS}'X_SM_S^{-1},\ s = 1,2,\ldots,S$$

(iv)From (3.1) through (3.5) compute the direction vector $\underline{d}$ and the step-length ρ from (3.6a) or (3.6b).

(v)Move to $\underline{x} = \overline{\underline{x}} - \rho\underline{d}$, and at the jth iterate, move to $\underline{x}_j = \overline{\underline{x}}_{j-1} - \rho_{j-1}\underline{d}_{j-1}$

(vi)Is $\| f\left(\underline{x}_{j-1};m\right) - f\left(\underline{x}_j;m\right) \| \leq \delta$?; $\delta \geq 0$

Yes: Stop, and set $\underline{x}_j = \underline{x}^*$

No: Let $\underline{x}_j$ fall into the tth segment and define for this segment the information matrix,
$M_{jt} = M_t + \underline{x}_j\underline{x}_j'$; $\quad M_t = X_t'X_t$

$$X_{jt} = \begin{pmatrix} X_t \\ \cdots \\ \underline{x}'_j \end{pmatrix},$$

and the mean square matrix,

$$B_t = M_t^{-1} + M_t^{-1}X_{jt}'X_{Bt}\underline{a}_2\underline{a}_2'X_{Bt}'X_{jt}M_t^{-1}.$$

Return to step (iii) above, taking notice of the fact that $\overline{\underline{x}}_s, M_s$ and B_s remain unchanged for all $s \neq t$.

3 The direction vector and step-Length

By pooling the segmented models in section 2, step (ii) we obtain a single model,

$$\underline{y} = \underline{X}\underline{q} + \underline{e}; \tag{3.1}$$

$\underline{y} = \left(\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_S\right)'$ is an N–component column vector of unknown responses, $\underline{e} = (\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_S)'$ is an N–component vector of random observation error. $\underline{X} = diag\{X_1, X_2, \ldots, X_S\}$ is an $N \times S(n+1)$design matrix $\underline{q} = \left(\underline{q}_1, \underline{q}_2, \ldots, \underline{q}_S\right)'$ is an $S(n+1)$–component column vector of unknown parameters. We define the $(n+1)$–component direction vector $\underline{d} = (\underline{d}_0, \underline{d}_1, \ldots, \underline{d}_n)'$ as a vector of convex combinations,

$$d_i = \sum_{s=1}^{S} h_{is}q_{is}; \quad h_{is} \geq 0, \quad \sum_{s=1}^{S} h_{is} = 1, \quad i = 0,1,2,\ldots,n \tag{3.2}$$

with mean square error, $\overline{m}(d_i) = \sum_{s=1}^{S} h_{is}^2\overline{m}_{ii(s)}$; $\overline{m}_{ii(s)}$ is the ith diagonal element of the matrix B_s. We minimize $\overline{m}(d_i)$ from the partial derivative equations
$\{\partial\overline{m}(d_i)/\partial h_{is}\}_{s=1}^{S-1} = 0$.
Hence, the matrix equation,

$$\left(D_i + \overline{m}_{ii(S)}J\right)\underline{h}_i = \overline{m}_{ii(S)}\underline{1}; \quad i = 0,1,\ldots,n. \tag{3.3}$$

In (3.3), $D_i = diag\left\{\bar{m}_{ii(1)}, \bar{m}_{ii(2)}, \ldots, \bar{m}_{ii(S-1)}\right\}$ is an $(S-1)\times(S-1)$ diagonal matrix, $\underline{h}_i = \left(h_{i1}, h_{i2}, \ldots, h_{i(S-1)}\right)'$, $\underline{1} = (1,1,\ldots,1)'$ is an $(S-1)$–component vector of unit elements. Therefore, $\underline{h}_i = \left(D_i + \bar{m}_{ii(S)}J\right)^{-1}\bar{m}_{ii(S)}\underline{1}$; $J = \underline{1}\underline{1}'$ is an $(S-1)\times(S-1)$ matrix with unit elements,

$\left(D_i + \bar{m}_{ii(S)}J\right)^{-1} = D_i^{-1} - D_i^{-1}\bar{m}_{ii(S)}\underline{1}\left(1 + \underline{1}'D_i^{-1}\underline{1}\right)^{-1}\underline{1}'D_i^{-1}$. And define,

$\underline{\underline{h}}_i = \begin{pmatrix} \underline{h}_i \\ h_{iS} \end{pmatrix}$; $h_{iS} = 1 - \sum_{s=1}^{S-1} h_{is}$.

By re-arranging the vectors $\left\{\underline{\underline{h}}_i\right\}_{i=0}^{n}$ we obtain the matrices of coefficients of convex combinations,

$H = (H_1, H_2, \ldots, H_S)$ of dimension $(n+1)\times S(n+1)$; $H_s = \text{diag}\left\{h_{0s}, h_{1s}, \ldots, h_{ns}\right\}$, $s = 1,2,\ldots,S$, which is now normalised such that $HH' = I$, I is an identity matrix. In vector notations, (3.2) is written,

$\underline{d} = H\underline{q}$,

which from above normalization of the H matrix implies $\underline{q} = H'\underline{d}$.

Then, from (3.1) we get the model

$$\underline{z} = M_H\underline{d} + \underline{\upsilon}; \qquad 3.4$$

$\underline{z} = H\underline{X}'\underline{y}$, $M_H = H\underline{X}'\underline{X}H'$, $\underline{\upsilon} = H\underline{X}'\underline{e}$ of dimensions $(n+1)$, $(n+1)\times(n+1)$ and $(n+1)$ respectively.

To obtain the least squares estimate of $\underline{d}$ from (3.4); first, we write M_H in expanded form:

$$M_H = (m_{ki}) = \left(\begin{array}{c|cccc} m_{00} & m_{01} & \cdots & \cdots & m_{0n} \\ m_{10} & m_{11} & \cdots & \cdots & m_{1n} \\ \vdots & & & \vdots & \\ m_{n0} & m_{n1} & \cdots & \cdots & m_{nn} \end{array}\right)$$ and obtain from (2.1),

$z_0 = f(m_{01}, m_{02}, \ldots, m_{0n}; m)$, $z_1 = f(m_{11}, m_{12}, \ldots, m_{1n}; m)$, $\ldots, z_n = f(m_{n1}, m_{n2}, \ldots, m_{nn}; m)$; which are obtained by replacing $\{x_i\}$ in $f(x_1, x_2, \ldots, x_n; m)$ with $\{m_{ki}\}$ in the kth row of M_H, $k = 0,1,\ldots,n$, $i = 1,2,\ldots,n$.

Hence the least squares estimate

$$\hat{\underline{d}} = M_H^{-1}\underline{z} = \begin{pmatrix} \hat{d}_0 \\ \ldots \\ \hat{d}_1 \\ \vdots \\ \hat{d}_n \end{pmatrix} \rightarrow \underline{\underline{d}} = \begin{pmatrix} \hat{d}_1 \\ \hat{d}_1 \\ \vdots \\ \hat{d}_n \end{pmatrix} \qquad 3.5$$

and normalize to $\underline{\underline{d}}'\underline{\underline{d}} = 1$.

Notice, the optimal properties of the direction vector:

(a) No more than $(n+1)$ independent responses $\{z_i\}$ are needed to obtain the estimate, $\hat{\underline{d}}$, as may be expected from Brooks and Mickey (1961).

(b) The estimator $\hat{\underline{d}}$, has a least squares–minimum norm property being solution to the Lagrangian function,

$L(\underline{d}, \underline{\lambda}) = \min(\underline{d}'\bar{B}\underline{d})$, subject to $\underline{z} = M_H\underline{d}$;

$\bar{B} = H\text{diag}\{B_1, B_2, \ldots, B_S\}H'$ is an $(n+1)\times(n+1)$ matrix.

$\underline{\lambda}$ is a vector of Lagrangian multiplier.

This derivation of the direction vector $\underline{d}$, applies to constrained, as well as, unconstrained problems; however, for the step-length ρ, the computational procedure depends on whether the optimizer $\underline{x}^*$ is an interior or a boundary point.

4 Computation of ρ

When $\tilde{X}_c$ is a convex set, as is true, when the constraints are first-order linear, $\underline{x}^*$ is a boundary point and the step-length ρ is obtained from the equation:

$$\rho = \min_k \left\{ \underline{c}'_k \left(\overline{\underline{x}} - \rho_k \underline{\underline{d}} \right) - b_k \right\} = \min_k \left\{ \left(\underline{c}'_k \overline{\underline{x}} - b_k \right) / \underline{c}'_k \underline{\underline{d}} \right\}, \quad k = 1, 2, \ldots, K. \qquad 4.1$$

Now, for constrained problems with non-linear constraints as well as unconstrained problems, $\underline{x}^*$ is expected to be an interior point, ρ is obtained by solving for the roots of the $m-1$ degree polynomial from the derivative equation of the objective function in (2.1);

$$d f((\overline{\underline{x}} - \rho \underline{\underline{d}}); m) / d\rho = 0 \qquad 4.2$$

5 Global Convergence

From Section 2, step (vi) we notice that $M_{tj} = M_t + \underline{x}_j \underline{x}'_j$, and from (3.4),

$M_{Hj} = M_{Hj-1} + \left(H_{jt} \underline{x}_j \right) \left(H_{jt} \underline{x}_j \right)'$;

H_{jt} is a matrix of convex combinations at the jth step when $\underline{x}_j$ falls into the tth segment.

$\Rightarrow \left\{ M_{Hj} \right\}_{j=1}^{\infty}$ is a non-decreasing sequence.

Consequently,

$\det \left(M_{Hj} M_{Hj-1}^{-1} \right) = 1 + \omega_j$;

$\omega_j = \left(H_j \underline{x}_j \right)' M_{Hj-1}^{-1} \left(H_j \underline{x}_j \right)$ is the jth step variance incremental factor; hence

$\det \left(M_{Hj} M_{Hj-1}^{-1} \right) < \det \left(M_{Hj-1} M_{Hj-2}^{-1} \right)$

$\Rightarrow \left\{ \omega_j \right\}_{j=1}^{\infty}$ is an non-increasing sequence.

Therefore, Kolmogorov's criterion for absolute convergence is satisfied; see e.g. Feller (1966, page 259). Now, since exploration at each step is global, encompassing the entire feasible region $\tilde{X}_f$ (see section 2) the convergence to $\underline{x}^*$ is global as well.

6 Numerical Examples

Two examples are covered in this section. The first is the unconstrained Rosenbrock test function taken from Onukogu (1997). The second is a tri-variate constrained quadratic function from Fletcher (1981, page 101).

6.1 Example 1. (Rosenbrock function)

The problem is defined in Wilde and Beightler (1967, page 301) as
$\min f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$.
And to reflect the biasing vector, we re-write $f(x_1, x_2)$ thus;

$f(x_1, x_2) = 1 - 2x_1 + (x_1^2 + 100x_2^2 - 200x_1^2 x_2 + 100x_1^4)$,

and let $\tilde{X} = \{x_1, x_2; -2.2 < x_1 < -0.2,\ 0 < x_2 < 2\}$.

Hence, the bias vector, $\underline{c}_B = (1, 100, -200, 100)'$

For two segments $(S = 2)$, using $x_1$to create segments, we define,

$$\tilde{X}_1 = \{x_1, x_2; -1.2 < x_1 < -0.2,\ 0 < x_2 < 2\}$$

$$\tilde{X}_2 = \{x_1, x_2; -2.2 < x_1 < -1.2, 0 < x_2 < 2\}$$

The design matrices for the two segments are, respectively, for $y_s = a_{0s} + a_{1s}x_1 + a_{2s}x_2 + e_s;\ s = 1,2,$

$$X_{11} = \begin{pmatrix} 1 & -1.2 & 0.0 \\ 1 & -0.2 & 0.0 \\ 1 & -1.2 & 2.0 \\ 1 & -0.2 & 2.0 \end{pmatrix} \text{ and } X_{12} = \begin{pmatrix} 1 & -1.2 & 0.0 \\ 1 & -2.2 & 0.0 \\ 1 & -1.2 & 2.0 \\ 1 & -2.2 & 2.0 \end{pmatrix},$$

from which we derive the bias matrices,

$$X_{B1} = \begin{pmatrix} 1.44 & 0.0 & 0.0 & 2.0736 \\ 0.04 & 0.0 & 0.0 & 0.0016 \\ 1.44 & 4.0 & 2.88 & 2.0736 \\ 0.04 & 4.0 & 0.08 & 0.0016 \end{pmatrix},\ X_{B2} = \begin{pmatrix} 1.44 & 0.0 & 0.0 & 2.0736 \\ 4.84 & 0.0 & 0.0 & 23.4256 \\ 1.44 & 4.0 & 9.68 & 2.0736 \\ 4.84 & 4.0 & 9.68 & 23.4256 \end{pmatrix}$$

The mean square error matrices are obtained by

$$\bar{M}_1 = \left(X'_{11}X_{11}\right)^{-1} + \underline{b}_1\underline{b}'_1$$

and

$$\bar{M}_2 = \left(X'_{12}X_{12}\right)^{-1} + \underline{b}_2\underline{b}'_2$$

for segments 1 and 2 respectively.
Where,
$\underline{b}_1 = (X'_{11}X_{11})^{-1}X'_{11}X_{B1}\underline{c}_B$, $\underline{b}_2 = (X'_{12}X_{12})^{-1}X'_{12}X_{B2}\underline{c}_B$. Let, $\underline{q}_1$ and $\underline{q}_2$ be the gradient vectors given in step (ii), section 2, the direction vector,

$$\underline{d} = \begin{pmatrix} d_0 \\ d_1 \\ d_2 \end{pmatrix} = H\underline{q}_1 + (I - H)\underline{q}_2;$$

$$H = \begin{pmatrix} h_{10} & 0 & 0 \\ 0 & h_{11} & 0 \\ 0 & 0 & h_{12} \end{pmatrix} = \begin{pmatrix} 0.9992 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.8201 \end{pmatrix}$$

and

$$HH' + (I - H)(I - H)' = I.$$

with the information matrices for the segments as

$$M_1 = HX'_{11}X_{11}H'$$

and

$$M_2 = (I - H)X'_{12}X_{12}(I - H)',$$

the pooled information matrices from the segments is

$$M_H = M_1 + M_2 = \begin{pmatrix} 4.000 & sym & \\ -2.7978 & 2.9600 & \\ 1.7840 & -1.1482 & 4.6204 \end{pmatrix}.$$

Hence, the response vector,

$$\underline{z} = (z_0,\ z_1,\ z_2)',$$

$$z_0 = f(-2.7978,\ 1.7840) = 3.667.0359$$

$$z_1 = f(2.9600, -1.1482)\ = 9825.0284$$

$$z_1 = f(-1.1482, 4.6204)\ = 1094.9594$$

Therefore, the direction vector

$$\underline{d} = \begin{pmatrix} \hat{d}_0 \\ \cdots \\ \hat{d}_1 \\ \hat{d}_2 \end{pmatrix} = M_H^{-1}\underline{z}$$

$$\rightarrow \underline{\underline{d}} = \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2 \end{pmatrix} = \begin{pmatrix} 0.9997 \\ -0.0245 \end{pmatrix};$$

The step-length ρ is obtained from the derivative function

$$df\left(\underline{\bar{x}} - \rho\underline{\underline{d}}\right)/d\rho = 0 \rightarrow \rho = 2.174$$

A move is now made to

$$\underline{x}_1 = \begin{pmatrix} -1.2 \\ 1.0 \end{pmatrix} + 2.174\begin{pmatrix} 0.9997 \\ -0.0245 \end{pmatrix} = \begin{pmatrix} 0.97 \\ 0.95 \end{pmatrix}$$

Since, the optimizer

$$\underline{x}^* = \begin{pmatrix} 1.0 \\ 1.0 \end{pmatrix}$$

a second move is considered unnecessary.

6.2 Constrained Tri-Variate Objective Quadratic

The example is taken from Fletcher (1981, page 101) and requires the minimum of

$$f(x_1, x_2, x_3) = 2x_1 - x_2 - x_1x_2 - x_2x_3 + x_1^2 + x_2^2 + x_3^2$$

s.t.

$$\tilde{X}_f = \{x_1, x_2, x_3; 3x_1 - x_2 + x_3 \geq 0, 2x_1 - x_2 - x_3 \leq 0\}.$$

With two segments $(S = 2)$ defined by

$\tilde{X}_1 = \left\{x_1, x_2, x_3; x_1 \leq \frac{1}{2}, x_2, x_3\right\}$ and $\tilde{X}_2 = \left\{x_1, x_2, x_3; x_1 \geq \frac{1}{2}, x_2, x_3\right\}$

We develop the design matrices

$$X_{11} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & -\frac{1}{3} & 0 & 1 \\ 1 & \frac{1}{2} & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & -\frac{1}{2} & \frac{1}{2} \\ 1 & \frac{1}{3} & 0 & 1 \\ 1 & -\frac{1}{2} & -1 & \frac{1}{2} \end{pmatrix},\ X_{12} = \begin{pmatrix} 1 & \frac{1}{2} & 1 & 0 \\ 1 & \frac{1}{2} & 0 & 1 \\ 1 & \frac{2}{3} & \frac{1}{3} & -1 \\ 1 & \frac{1}{2} & 4 & 3 \end{pmatrix},\ \underline{c}_B = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Just as in (6.1) the other matrices; i.e. $X_{B1}, X_{B2}, \overline{B}_{1t}, \overline{B}_{2t}, H, M_H$ are similarly generated. The direction vector,
$\underline{\underline{d}} = (0.8136, -0.2419, 0.5287)'$; $\underline{\underline{d}}'\underline{\underline{d}} = 1$, $\rho = 0.1999$, $\underline{\overline{x}} = (0.1806, 0.4444, 0.5417)$,
$\underline{x}_1^* = \underline{\overline{x}} - \rho \underline{\underline{d}} = (0.0180, 0.4927, 0.4360)'$
$f(\underline{x}_1^*) = -0.2472$.

Since, the minimizer $\underline{x}^* = (0, \frac{1}{2}, \frac{1}{2})$ and $f(\underline{x}^*) = -\frac{1}{4}$ a second iterate is considered unnecessary.

7 Summary and Conclusions

The paper introduces a method, named SCLS for solving either constrained or unconstrained objective function that frequently occur in different areas of scientific research. The method is shown to be more effective than other line search techniques. The SCLS can be identified by the following properties:

(i)Exploration at each move is global, covering the entire feasible region through segmentation of $\tilde{X}_f$;
(ii)The direction vector is along the least-square, minimum-norm direction;
(iii)No more than maximum of $(n+1)$ independent responses are needed at each iteration point; and
(iv)Convergence is global.

References

[1] Brooks, S. H. and Mickey, M. R., Optimum Estimation of Gradient in Steepest Ascent Experiments; *Biom*, **17**, 48–56 (1961).
[2] Box, G. E. and Wilson, K. B., On Experimental Attainment of Optimum Conditions, *Jour. of the Royal Statistical Society Series B*, **13**, 1–45 (1951).
[3] Cauchy, A., Methode Générale pour la Résolution des Systémes D'équations Simultanées; *Compt Rend Acad Sci Paris*, **25**, 536–538 (1894).
[4] Cauchy, A., *Oeuvres Completes*; Gauthier-Villars, Paris, 399–406 (1901).
[5] Dantniz, G. B., *Linear Programming and Extensions*, Princeton University Press, N. J., (1963).
[6] Dantniz, G. B. and Wolfe, P., The Decomposition Algorithms for Linear Progams, *Econometrica*, **29**, (1961).
[7] Feller, W., *An Introduction to Probability Theory and its Applications* 3rd ed.; J. Wiley and Sons Inc, (1966).
[8] Flecher, R., *Practical Methods of Optimization, Constrained Optimization;* Printice Hall, **2**, (1981).