MAXIMUM ENTROPY APPROACH IN LINEAR PROGRAMME

DR. RAMESH KAR
Associate professor,
Department of Mathematics,
Bidhan Chandra College, Rishra,
Rishra, Hooghly-712248, West Bengal, India.

ABSTRACT:
The Paper deals with a modified maximum-entropy method of solution of linear programming problem transforming it by an equivalent surrogate program.

KEYWORDS: Shannon Entropy, Optimization Technique, Surrogate, Lagrangian parameter.

INTRODUCTION:

Let us consider a general optimization problem of the following form:

Problem – I

Minimize $\{g(x) = g(x_1, x_2, \ldots, x_n)\}$ (1.1)
$\qquad \{x_i, i = 1, 2, \ldots n\}$

Subject to the constraints $d_j(x) \le 0, j = 1, 2, \ldots, m$ (1.2)

Where x denotes vector of real, continuous-valued variables $x_i (i = 1, 2, \ldots, n)$ and the constraint functions $d_j (j = 1, 2, \ldots, m)$ represent constraint vector **d**. Problems of type – I with non-linear functions **g** have far reaching applications and there are plenty of techniques of solving this type of problems [11, 14]. The numerical solutions of this type of problem are of great importance for engineering design, synthesis and analysis [4, 18]. Numerical solution of this type of problems on the basis of Jaynes maximum-entropy principle [9] is first due to Templeman and Li [17]. The method was modified by Das, Mazumder and De [2] on the basis of kullback minimum cross-entropy principle [10]. Later on Das and Chakrabarti [3] has formulated a modified maximum-entropy principle, which removes some flaws and demerits of the paper of Templeman and Li [17] and that of Das, Mazumder and De [2]. This method enables a powerful algorithm of tackling numerical optimization problems, which are very often of great need in engineering science.

The object of the paper is to show how this modified maximum-entropy algorithm technique can be successfully applied to the numerical solution of the linear and non-linear programming problems. We have attempted to establish our claims with four numerical examples pointing out the advantageous points of the technique.

EQUIVALENT SURROGATE PROGRAMMING: MODIFIED MAXIMUM-ENTROPY APPROACH

For clear understanding of the mathematical background of the technique we describe it briefly as follows: The **Problem I** have an equivalent surrogate form [17].

Problem – II

Minimize $\{ g(x) = (x_1, x_2, \ldots, x_n) \}$

$\{ x_i, i = 1, 2, \ldots, n \}$

Subject to the single constraint $\sum\limits_{j=1}^{m} \mu_j \, d_i(x) = 0$ (2.1)

Where μ_j (j = 1, 2, …, m) are non-negative multipliers known as surrogate multipliers. The surrogate multipliers μ may be assumed to be normalized without loss of generality, i.e.

$$\sum\limits_{j=1}^{m} \mu_j = 1 \qquad\qquad (2.2)$$

The solution process is set in a probabilistic context by considering the **Problem I** as posed initially and estimating what level of certainty (probability) should be assigned to the event that each constraint is active at the problem solution. Denoting these probability by μ_j (j = 1, 2, …, m) it is known that at least one of the constraints must be active so that (2.2) must hold. The desired solution x^* of the **Problem I** will here be sought indirectly through a sequence of solutions of the **Problem II**. This approach assumes therefore, that **Problems I** and **II** are equivalent at the solution point specially that a set of multipliers μ^* exists and can be found such that x^* which solves **Problem II** with μ^* also solves **Problem I**.

The Lagrangian of problem II has the form

$$L_{II}(x, \mu, a) = g(x) + \beta \sum\limits_{j=1}^{m} \mu_j \, d_j(x) \qquad\qquad (2.3)$$

Where β is the Lagrange multiplier associated with the constraint (2.1) in problem II. An essential condition to be satisfied is that (x^*, μ^*, a^*) should be a saddle point of the lagrangian L_{II} of the Problem II

$$L_{II}(x, \mu^*, \beta^*) \geq L_{II}(x^*, \mu^*, a^*) \geq L_{II}(x^*, \mu, \beta) \qquad\qquad (2.4)$$

The saddle point condition (2.4) may be satisfied, however by iterative means using problem II itself with alternative iterations in x – space and μ space. A typical scheme is as follows:

An initial set of multipliers μ^0 is chosen and problem II is solved to yield corresponding values of x^0. The multipliers are then updated to μ^1 and problem II is solved again to give x^1. The process is repeated until the sequence $(\mu^0, x^0), (\mu^1, x^1) \ldots (\mu^k, x^k)$, converges to the solution of problem II and hence also of problem I, at (x^*, μ^*). The interation process should satisfy the inequality [17]

$$\mu_j^{k+1} \begin{cases} > \mu_j^k & \text{if} & d_j(x^k) > 0 \\ = \mu_j^k & \text{if} & d_j(x^k) = 0 \\ < \mu_j^k & \text{if} & d_j(x^k) < 0 \end{cases} \qquad\qquad (2.5)$$

The new multipliers μ^{k+1} must satisfy the normality conditions and the surrogate constraints of problem II. Thus

$$\sum_{j=1}^{m} \mu_j^{k+1} = 1 \tag{2.6}$$

$$\sum_{j=1}^{m} \mu_j^{k+1} \, d_j(x^{k+1}) = 0 \tag{2.7}$$

For the conditions (2.7) to be satisfied we require values for $d_j(x^{k+1})$, j = 1, 2, …, m which are not known as yet. Therefore, the best current estimates for $d_j(x^{k+1})$ are the values $d_j(x^k)$ which are available and must be used in their place. Consequently (2.7) is modified to

$$\sum_{j=1}^{m} \mu_j^{k+1} \, d_j(x^k) = \in \tag{2.8}$$

Where $\in$ represents the unknown error introduced by using $d_j(x^k)$ in place of $d_j(x^{k+1})$, j = 1, 2, …, m. Furthermore, we would expect $\in$ to approach zero as the sequence of iterations x^k, x^{k+1}, … approaches to x^*.

Now let us write $\quad y_j^{k+1} = (\mu_j^{k+1} - \mu_j^k) \, d_j(x^k)$, $j = 1, 2, \dots, m.$ (2.9)

In view of (2.1) and (2.8) we have then

$$\sum_{j=1}^{m} y_j^{k+1} = \sum_{j=1}^{m} (\mu_j^{k+1} - \mu_j^k) \, d_j(x^k) = \varepsilon$$

Which tends to zero as $k \to \infty$

Since $\quad \displaystyle\sum_{j=1}^{m} \mu_j^k \, d_j(x^k) = 0 \quad$ we thus have then

$$\sum_{j=1}^{m} y_j^{k+1} = \varepsilon \tag{2.10}$$

$$\sum_{j=1}^{m} \frac{y_j^{k+1}}{d_j(x^k)} = 0 \tag{2.11}$$

and $\quad \mu_j^{k+1} \geq 0, \quad j = 1, 2, \dots, m.$ (2.12)

To determine y_j^{k+1} we pose the following problem :

Problem – III

Maximize $S = -\displaystyle\sum_{j=1}^{m} y_j^{k+1} \ln y_j^{k+1}$ (2.13)

Subject to the constraints $\quad \displaystyle\sum_{j=1}^{m} y_j^{k+1} = \in$ (2.14)

$$\sum_{j=1}^{m} \frac{y_j^{k+1}}{d_j(x^k)} = 0 \tag{2.15}$$

and $\quad \mu_j^{k+1} \geq 0, \quad j = 1, 2, \dots, m.$ (2.16)

The solution of problem III gives :

$$In\ y_j^{\,k+1} = (\alpha - 1) + \frac{\eta}{d_j(x^k)}\ ,\quad j = 1, 2, \ldots, m.$$

Or $y_j^{\,k+1} = Be^{\eta/d_j(x^k)}$ (2.17)

Where η is determined from the equation

$$\sum_{j=1}^{m} \frac{e^{\eta/d_j(x^k)}}{d_j(x^k)} = 0 \tag{2.18}$$

Let the value of η be η* then from (2.17), we have

$$y_j^{\,k+1} = Be^{\eta^*/d_j(x^k)}$$

So, $\displaystyle\sum_{j=1}^{m} y_j^{\,k+1} = \sum_{j=1}^{m} Be^{\eta^*/d_j(x^k)} = \varepsilon$

Implying $B \displaystyle\sum_{j=1}^{m} Be^{\eta^*/d_j(x^k)} = \varepsilon$ (2.19)

   Since we expect ∈ gradually approaches to zero as the iteration proceeds (i.e ∈ → 0 as $K \to \infty$) therefore, equation (2.19) states that B is a non-negative quantity which infact gradually approaches to zero as the iteration increases.

By virtue of (2.9) the solution of the problem III gives the following assignment of $\mu_j^{\,k+1}$:

$$\mu_j^{\,k+1} = \mu_j^{\,k} + \frac{Ae^{\eta/d_j(x^k)}}{d_j(x^k)},\quad j = 1, 2, \ldots, m \tag{2.20}$$

$\mu_j^{\,k+1} \geq 0$ we have

$$B \leq \frac{-\mu_i^{\,k}\, d_j(x^k)}{e^{\eta/d_j(x^k)}}\ ,\quad j = 1, 2, \ldots, m \tag{2.21}$$

   In view of this it should be possible to prescribe procedures for selecting sequence {B} of positive numbers that converges to zero. The above method satisfies the conditions (2.5) absolutely confirming the convergence of μ [3].

APPLICATION TO PROGRAMS:

   Here we are going to deliberate an interesting application of the above modified maximum-entropy method in the solution of linear and non-linear programming problems. Bregman-Romanovsky [1] and Erlander [5] studied entropy constrained linear programs. Bregman-Romanovsky [1] add an entropy term to the objective function with the purpose of introducing an element of smoothing into the solutions of linear programs. Erlander [5] adds an entropy constraint to the classical linear programs in order to achieve a desired level of accessibility in the solutions.

   In the present section we are going to deal with the technique of section 2 in the solution of linear and non-linear programming problems. This technique has been shown to be very successful in generating an iterative procedure in the numerical solution of linear and non-linear programs. The different steps in the iterative procedure are as follows:

   a) The first step is to reduce the original program into a surrogate and reduce the number of constraints into a single constraint
   b) The next step is to add an entropy term to the new objective function so that the new objective function becomes a non-linear function (in case of linear program) of its arguments in consistent with the formalism of the section 2.
   c) Next is to write down the lagrangian for the surrogate problem introducing a Lagrangian parameter β (say).

d) Maximization of the Lagrangian with respect to the new surrogate variables y (say)leads to the set of equations which determine the lagrangian parameter β.

e) We start with the computation scheme by setting surrogate multipliers μ equal so that their total sum be unity. This is consistent with Laplaces principle of insufficient knowledge-a particular case of Jaynens'maximum-entropy principle [9].

f) By starting with the initial values $\mu_j{}^o$, we upgrade their values step by step and find the corresponding values of the surrogate variables y_i and hence original variables x_i ($i = 1, 2, …, n$). The whole iterative process can be arranged in a tabular form for a better representation of the solution. The procedure has been explained more explicitly by four numerical examples presented below.

Numerical Example 1:

A firm can produce three types of cloth, say A, B and C. Three kinds of wool are required for it, say, red wool, green wool and blue wool. One unit length of type A cloth needs 2 yards of red wool and 3 yards of blue wool; one unit length of type B cloth needs 3 yards of red wool, 2 yards of green wool and 2 yards of blue wool and one unit of type C cloth needs 5 yards of green wool and 4 yards of blue wool. The firm has only a stock of 8 yards of red wool, 10 yards of green wool and 15 yards of blue wool. It is assumed that the income obtained from the one-unit length of type A cloth is Rs. 3.00, of type B cloth is Rs. 5.00 and of type C cloth is Rs. 4.00.

The L.P. model is written as:

Maximize $Z = 3x_1 + 5x_2 + 4x_3$

Subject to constraints:

$d_1 \equiv 2x_1 + 3x_2 - 8 \leq 0$

$d_2 \equiv 2x_2 + 5x_3 - 10 \leq 0$ (1.a)

$d_3 \equiv 3x_1 + 2x_2 + 4x_3 - 15 \leq 0$

and $x_1, x_2, x_3 \geq 0$

The surrogate form of (1.a) is given by

Maximize $\overline{Z} = 3x_1 + 5x_2 + 4x_3 - x_1\ln x_1 - x_2\ln x_2 - x_3\ln x_3$

Subject to

$\mu_1(2x_1 + 3x_2 - 8) + \mu_2(2x_2 + 5x_3 - 10) + \mu_3(3x_1 + 2x_2 + 4x_3 - 15) = 0$

Or $a_1x_1 + a_2x_2 + a_3x_3 - a_4 = 0$ (1.b)

Where $a_1 = (2\mu_1 + 3\mu_3)$

$a_2 = (3\mu_1 + 2\mu_2 + 2\mu_3)$

$a_3 = (5\mu_2 + 4\mu_3)$

$a_4 = (8\mu_1 + 10\mu_2 + 15\mu_3)$ (1.c)

Now the lagrangian for the above surrogate problem is

$L(X, \beta) = (3x_1 + 5x_2 + 4x_3 - x_1\ln x_1 - x_2\ln x_2 - x_3\ln x_3) + \beta(a_1x_1 + a_2x_2 + a_3x_3 - a_4)$ (1.d)

Now, the maximization requires

$$\frac{\partial L}{\partial x_1} = 0 \Rightarrow x_1 = e^{(2+a_1\beta)}$$ (1.e)

$$\frac{\partial L}{\partial x_2} = 0 \Rightarrow x_2 = e^{(4+a_2\beta)}$$ (1.f)

$$\frac{\partial L}{\partial x_3} = 0 \Rightarrow x_3 = e^{(3+a_3\beta)}$$ (1.g)

Putting x_1, x_2, x_3, in (1.b) we have

$$a_1 e^{(2+a_1\beta)} + a_2 e^{(4+a_2\beta)} + a_3 e^{(3+a_3\beta)} - a_4 = 0 \tag{1.h}$$

Which gives the value of β for known a_1, a_2, a_3, a_4 (determined from (1.c) for given μ_1, μ_2, μ_3). We start the process by setting $\mu_1 = \dfrac{1}{3}$, $\mu_2 = \dfrac{1}{4}$ $\mu_3 = \dfrac{5}{12}$. Infact Jaynes maximum-entropy principle' generates this starting set for μ_1, μ_2, μ_3 because in the absence of any others extra information about the problem the least baised assumption that we can make is that all the constraints are equally weighted. Table 1 give the iterative result and is as exact as the three decimal place accuracy.

TABLE – 1

K	B	μ_1	μ_2	μ_3	β	x_1	x_2	x_3	d_1	d_2	d_3	η
0	–	$\dfrac{1}{3}$	$\dfrac{1}{4}$	$\dfrac{5}{12}$	-1.20254	.99552	3.428018	.54583	3.94832	-.69327	-3.22987	1.44748
1	.007	.33525	.33289	.33198	-1.20277	.99433	3.32516	.57821	3.94521	-.69324	-3.23495	1.44502
2	.003	.33555	.33287	.33161	-1.20289	.99345	3.32855	.56121	3.94532	-.69321	-3.23582	1.44500
3	.0001	.33578	.33284	.33169	-1.20282	.99341	3.28501	.56123	3.94535	-.69320	-3.23640	1.44480
4	00005	.33571	.33281	.33165	-1.20280	.99340	3.28500	.56120	3.94510	-.66180	-3.23610	1.44441

Numerical Example 2:

Let us consider the following optimization problem, which is of significant importance in engineering design [11]. A firm manufactures two types of products of x_1 and x_2 units. The profit function is given by
$$f(x_1, x_2) = x_1 + 3x_2$$

The problem here to maximize the linear function
Maximize $f(x_1, x_2) = x_1 + 3x_2$ (2.a)

Subject to constraints :

$$3x_1 + 6x_2 \leq 8$$
$$5x_1 + 2x_2 \leq 10, \qquad x_1, x_2 \geq 0 \tag{2.b}$$

Now first rewrite the L.P.P by introducing a transformation of variables defined by

$$y_1 = \frac{x_1}{2} \qquad \text{and} \qquad y_2 = \frac{x_2}{5}$$

We use this transformation in order to make one of the constraints of (2.b) in its normality from which infact is consistent with the method developed in this paper. Then the equation (2.a) and (2.b) transformed to

Maximize $\quad f(y_1, y_2) = 2y_1 + 15y_2$
Subject to $\qquad d_1 \equiv 3y_1 + 15y_2 - 4 \leq 0$ (2.c)
$\qquad\qquad d_2 \equiv y_1 + y_2 - 1 \leq 0$
$\qquad\qquad y_1, \quad y_2 \geq 0$

The surrogate form of (2.c) is given by

Maximize $\quad \overline{f}(y_1, y_2) = 2y_1 + 15y_2 - y_1\ln y_1 - y_2\ln y_2$
Subject to $\quad \mu_1(3y_1 + 15y_2 - 4) + \mu_2(y_1 + y_2 - 1) = 0$
Or $\qquad a_1 y_1 + a_2 y_2 - a_3 = 0$ (2.d)
Where
$\qquad\qquad a_1 = (3\mu_1 + \mu_2)$
$\qquad\qquad a_2 = (15\mu_1 + \mu_2)$ (2.e)

$$a_3 = (4\lambda_1 + \mu_2)$$

Following the same argument as in Numerical Example 1, Table 2 give the iterative results and is as exact as the three decimal place accuracy.

TABLE – 2

K	B	μ_1	μ_2	β	y_1	y_2	d_1	d_2	η
0	–	$\dfrac{3}{5}$	$\dfrac{2}{5}$	-1.84325	.506149E-01	.39786	.74281	-.54762	-0.876045E-13
1	0.007	0.6234	0.45632	-1.78965	.528821E-01	.39432	.70768	-.54642	-0.875084E-01
2	0.0004	0.6282	0.45611	-1.78698	.530307E-01	.39441	.70542	-.54252	-0.885041E-01
3	0.0001	0.6280	0.45601	-1.78611	.530365E-01	.39440	.70501	-.54250	-0.885011E-01

IMPLICATION:

In this paper we have analyzed the applicability of the modified maximum-entropy principle in the numerical solution of linear and non-linear programming problems. Infact, it may be justified to conclude the paper by saying that this metamorphosis to the surrogate mode has diminished the number of hindrances into a single one and, furthermore, this has lessened the manifold numerical complications of the problem undoubtedly to a large measure.

REFERENCES:

[1] BREGMAN, I.M. & ROMANOVSKY, JV. (1975): Raverstka I optimizatsija V Zadatjach Raspredelenija', University of Leningrao; USSR.

[2] DAS. N.C., MAZUMDER, S.K. & DE, KAJAL (1988): 'Constrained non-linear programming', Minimum cross-entropy algorithm, **Eng. Optimization** (in press)

[3] DAS. N.C. AND CHAKRABARTI. C.G. (1988): 'Non-linear constrained optimization: Information-Theoretic algorithm', **SAMS** (In press)

[4] DEB. K. (1995): **'Optimization for engineering design' Algorithms and examples,** Prentice Hall of India, New Delhi.

[5] ERLANDER, S. (1979): **'Entropy in Linear programs', mathematical programming,** pp. 137-150.

[6] GLOVER, F. (1986): 'Surrogate constraints', **Operations Research**, Vol. 16, pp. 741-749.

[7] GOULD, F.J. (1969): 'Extensions of Lagrange multipliers in non-linear programing', **SIAM Journal of Applied Mathematics**, Vol. 13, No. 6, pp. 1280-1297.

[8] GREENBERG, H.J. & PIERSKALLA, W.P. (1970): 'Surrogate Mathematical programming', **Operations Research**, Vol. 18, 924-939.

[9] JAYNES, E.T. (1957): 'Information theory and statistical mechancis', **The Physical Review**, Vol. 106, pp. 620-630.

[10] KULLBACK, S. (1959): 'Information theory and Statistical', **John Wiley**, New York.

[11] KAMBO, N.S. (1991): 'Mathematical Programming Techniques', Affiliated East-West Press.

[12] ORTEGA, J.M. AND REHEINBOLDT, W.C. (1970): 'Iterative solution of non-linear equations in several variables', Academic Press, New York.

[13] SHANNON, C.E. (1949): 'A mathematical theory of communication', **Bell System Technical Jr.** 27 (379-423).

[14] SWARUP KANTI, GUPTA, P.K., AND MAN MAHAN (1980): **'Operations Research'**, Sultan Chand & Sons, New Delhi.

[15] SINGIRESU, S.RAO (1988): Engineering Optimization. Theory and Practical.

[16] TRIBUS, M. (1969): Rational description and designs', Pergamon Press, New York.

[17] TEMPLEMAN, A.B. & XINGSI. LI (1987): 'A maximum-entropy approach to constrained non-linear programming', **Engg. Optimization**. Vol. 112, No. 2, pp. 191-205.

[18] TEMPLEMAN, A.B. (1993): **'Operations Research – An introduction', Macmillan Pub. Co. Inc.,** New York.

[19] J.N. KAPUR, Maximum Entropy Model in Science and Engineering, Wiley Eastern, New Delhi, 1990.