Why Unified Statistics Theory by MCMC towards Estimation of Stationary Joint Probabilities of Stochastic Matrix?

Dr. Usama H. Abou El-Enien

Abstract— **Unified statistics theory by MCMC is considered. A new proposed algorithm is presented to obtain surely empirical analysis conclusions in order to turn to surely theoretical analysis results about stationary joint probabilities of stochastic matrix of any genetic or clonal selection algorithm of any general too large dimensional deterministic and probabilistic (the grouping data, both continuous and discrete) linear or nonlinear programming problem in order to introduce a complete framework and to estimate any stationary joint probabilities of stochastic matrix by the proposed algorithm towards the fourth resounding success of unified statistics theory by MCMC.**

Index Terms—**Clonal selection algorithms, Genetic algorithms, General linear programming Problem, General nonlinear programming problem, Grouping data, joint probabilities, Stochastic matrix, Unified statistics theory by MCMC**

I. INTRODUCTION

Genetic and clonal selection algorithms within the framework of Markov chains and MCMC have been studied by El-Nady et al. [1], [2], [3], and [4].

Unified statistics theory by MCMC has been proposed by Abou El-Enien [5].

Unique chromosomes by simple random sampling without replacement theorem for any objective real valued function of n - variables of the form $f(x_1, x_2, \ldots, x_n)$, where $a_i \leq x_i \leq b_i$ for $i = 1, 2, \ldots, n$ within the framework of unified statistics theory by MCMC has been proposed by Abou El-Enien and Khalil [6].

Subset of unique chromosomes by simple random sampling without replacement theorem to solve any too large dimensional deterministic and probabilistic linear or nonlinear programming problems toward two obvious criteria (speed and accuracy) has been proposed by Abou El-Enien [7].

Estimation of stationary transition probabilities of stochastic matrix theorem for any objective real valued function (linear or nonlinear) of n - variables to estimate any stationary transition probabilities of stochastic matrix of any genetic or clonal selection algorithm of any general too large

dimensional deterministic and probabilistic linear or nonlinear programming problem has been proposed by Abou El-Enien [8].

These works do not develop estimation of stationary joint probabilities of stochastic matrix of any genetic or clonal selection algorithm of any general too large dimensional deterministic and probabilistic (the grouping data, both continuous and discrete) linear or nonlinear programming problem to provide a general framework.

The rest of the paper is organized as follows. In Section 2, we give the formulation of the problem. In Section 3, we state the main result. Then in Section 4, the proof of the main result is given in five steps. In Section 5, we propose the algorithm. In Section 6, we give numerical example. In Section 7, we give some concluding remarks.

II. FORMULATION OF THE PROBLEM

In this paper, we consider a problem, namely: Why unified statistics theory by MCMC towards estimation of stationary joint probabilities of stochastic matrix?.

Throughout this paper, we consider any objective real valued function (linear or nonlinear) of n - variables $f(x_1, x_2, \ldots, x_n)$, where $a_i \leq x_i \leq b_i$ for $i = 1, 2, \ldots, n$ are domains of each variable x_i and a_i and b_i are real numbers:

$u \geq 0$ equations:

$q_i(x_1, x_2, \ldots, x_n) = 0$ (linear or nonlinear), $i = 0, \ldots, u, \rightarrow$ (I)

and $m - u \geq 0$ inequalities:

$q_i(x_1, x_2, \ldots, x_n) \leq 0$ (linear or nonlinear), $i = u + 1, \ldots, m . \rightarrow$ (II)

Proposition 2.1. We restrict an arbitrary uncountable set $S = \{a_i \leq x_i \leq b_i \text{ for } i = 1, 2, \ldots, n\}$ to be a subset of n -space $\mathbf{R^n}$ as a sample space, restrict an arbitrary countable set T to be set of all $(x_1, x_2, \ldots, x_n)$ in $S = \{a_i \leq x_i \leq b_i\}$ for which $P(x_1, x_2, \ldots, x_n) > 0$ as a sample space (see [5]).

Proposition 2.2. We divide each interval $a_i \leq x_i \leq b_i$, $i = 1, 2, \ldots, n$ into **k** (**k** is a different optional integer number for each interval $a_i \leq x_i \leq b_i$) subintervals $a_i \leq x_i \leq c_i$, $c_i \leq x_{i+1} \leq d_i$, $\ldots, w_i \leq x_{k-1} \leq y_i$, $y_i \leq x_k \leq b_i$ and c_i, d_i, $\ldots, w_i$ and y_i (the new population) are optional real numbers, list the possible simple

random samples without replacement of n (n as n of S) subintervals from this new population (see [6]).

Proposition 2.3. We restrict any simple random sample without replacement of n subintervals of Proposition 2.2, get unique chromosomes $(= 2^k)$, k bits {0,1}, substitute in (I) and (II), get subset of unique chromosomes (see [7]), generate all possible combinations of states of subset of unique chromosomes.

III. MAIN RESULT

In this section, we shall state the main theorem.

Theorem 3.1. For any objective real valued function (linear or nonlinear) of n - variables $f(x_1, x_2, \quad, x_n)$, where $a_i \leq x_i \leq b_i$ for $i = 1, 2, \quad, n$ are domains of each variable x_i and a_i and b_i are real numbers:

$u \geq 0$ equations:

$q_i (x_1, x_2, \quad, x_n)$ = 0 (linear or nonlinear), $i = 0, \ldots, u$, and $m - u \geq 0$ inequalities:

$q_i (x_1, x_2, \quad, x_n) \leq 0$ (linear or nonlinear), $i = u + 1, \ldots, m$, the following holds:

(1) A real valued function is one that contains all possible simple random samples without replacement of n (n as n of S) subintervals from the new population.

(2) Every simple random sample without replacement of n subintervals has unique chromosomes, subset of unique chromosomes and all possible combinations of states of subset of unique chromosomes.

(3) By applying genetic or clonal selection algorithms. We have the probability of each $(f_{m-1} = s, f_m = t)$

$\forall \ m \geq 1$ (P), where

$$P = \frac{\text{number of times } (f_{m-1} = s, f_m = t) \text{ appeared}}{n - 1}$$

and have stationary joint probabilities of stochastic matrix.

IV. PROOF OF THE MAIN RESULT

In this section, we prove the main result in Theorem 3.1. We start with a useful theorem.

Theorem 4.1. Let (S, β, P) be a probability space and let T denote the set of all x in S for which $P(x) > 0$ Then T is countable (see [9]).

We shall prove Theorem 3.1 in five steps.

Proof of Theorem 3.1. Step 1. For the real valued function, we define a probability space (S, β, P) (see [5]).

Step 2. We prove that T is a countable subset of S , and define n - dimensional random variable defined on a sample space T (see [5]).

Step 3. We divide each interval $a_i \leq x_i \leq b_i$, $i = 1, 2, \quad, n$ into **k** subintervals, define new population, and list the possible simple random samples without replacement of n subintervals from this new population.

Step 4. We restrict any simple random sample without replacement of n subintervals, get unique chromosomes

$(= 2^k)$, k bits {0,1}, substitute in (I) and (II), get subset of unique chromosomes and generate all possible combinations of states of subset of unique chromosomes.

Step 5. We pick one state randomly, apply genetic or clonal selection algorithms on the state for n -iterations, where n is a large number, count for each $(f_{m-1} = s, f_m = t) \ \forall \ m \geq 1$ (see [5])

such that $f_0, f_1, \ldots$ are outcome functions

and s, t are any possible sequence of outcomes the number of times it appeared, calculate the probability of each $(f_{m-1} = s, f_m = t)$ (P), where

$$P = \frac{\text{number of times } (f_{m-1} = s, f_m = t) \text{ appeared}}{n - 1}$$

and get the stationary joint probabilities of stochastic matrix.

On the basis of Steps 1-5, we complete the proof of Theorem 3.1.

V. PROPOSED ALGORITHM

We prepared programs by using MATLAB 7.5. We named the proposed algorithm **Abou El-Enien Estimation of Stationary Joint Probabilities of Stochastic Matrix (Abou El-Enien ESJPSM)** and named the proposed algorithm of **ESTPSM (see [8]) Abou El-Enien ESTPSM**, the basic steps of **Abou El-Enien ESJPSM** algorithm are as follows:

1. Divide each interval $a_i \leq x_i \leq b_i$, $i = 1, 2, \quad, n$ into **k** subintervals, define the new population.

2. List the possible simple random samples without replacement of n subintervals from this new population.

For any restricted simple random sample without replacement of n subintervals, do the following:

a. Input number of bits k .

b. Get unique chromosomes = 2^k .

c. Substitute in (I) and (II).

d. Get subset of unique chromosomes.

e. Generate all possible combinations of states of subset of unique chromosomes and give each state a number.

f. Pick one state randomly.

g. Apply genetic or clonal selection algorithms on the state for n -iterations, where n is a large number.

h. Count for each $(f_{m-1} = s, f_m = t) \ \forall \ m \geq 1$

such that $f_0, f_1, \ldots$ are outcome functions

and s, t are any possible sequence of outcomes the number of times it appeared.

i. Calculate the probability of each $(f_{m-1} = s, f_m = t)$ (P), where

$$P = \frac{\text{number of times } (f_{m-1} = s, f_m = t) \text{ appeared}}{n - 1}.$$

j. Get the stationary joint probabilities of stochastic matrix.

VI. NUMERICAL RESULTS AND DISCUSSION

Consider the following function:

$$f(x) = x \cdot \sin(10 \prod \cdot x) + 1, \ x \in [-1, 2]$$

if $k = 5$ bits, $m = 2$ chromosomes, probability of crossover = 0.6, probability of mutation = 0.9 then **unique chromosomes = {** 00000 = -1.000000, 00001 = - 0.903226, ……, 11101 = 1.806452, 11110 = 1.903226, 11111 = 2.000000 **}** and globally optimum value = 1.225806. All possible combinations of states of unique chromosomes = $2^{k \times m}$ = { 0: (00000, 00000), 1: (00000, 00001), . . . , 1023: (11111, 11111) }

Pick one state randomly (state 1006), apply GAs with bit mutation on the state for n = 100000 and then get the stationary joint probabilities of stochastic matrix**.**
1006,17,1006,17,972,……,63,960,30,977,55.
One of the stationary joint probabilities of stochastic matrix P (1006,17) =0.000080.

DISCUSSION

In this paper, the main result is Abou El-Enien estimation of stationary joint probabilities of stochastic matrix theorem for any objective real valued function (linear or nonlinear) of n - variables. Using this, we propose Abou El-Enien estimation of stationary joint probabilities of stochastic matrix method to estimate any stationary joint probabilities of stochastic matrix of any genetic or clonal selection algorithm of any general too large dimensional deterministic and probabilistic linear or nonlinear programming problem.

REFERENCES

[1] Kh. El-Nady, U. Abou El-Enien, A. Badr, Why are genetic algorithms Markov?. InterStat, InterStat, (2010), no.4.
[2] Kh. El-Nady, U. Abou El-Enien, A. Badr, Why are clonal selection algorithms Markov?. InterStat, InterStat, (2010), no.5.
[3]   Kh. El-Nady, U. Abou El-Enien, A. Badr, Why are clonal selection algorithms MCMC?. International Journal of Open Problems in Computer Science and Mathematics, ICSRS Publication, **4**(2011), no.3, pp. 37-55.
[4] Kh. El-Nady, U. Abou El-Enien, A. Badr, Why are genetic algorithms MCMC2 Markov Chain Monte Carlo. Advances in Modelling, B Signal Processing and Pattern Recognition, **54**(2011), no 1-2.
[5] U. H. Abou El-Enien, A New Unified MCMC Methods Toward Unified Statistics Theory by MCMC, LAP, Germany, 2012**.**
[6] U. Abou El-Enien, M. Khalil, Why unified statistics theory by MCMC towards resounding success ?., Indian Journal of Computer Science: Theory and Practical., GBS Publishers & Distributors (I), **1**(2013), no.1, pp. 13-16.
[7] U. Abou El-Enien, Why unified statistics theory by MCMC towards linear and nonlinear programming problems?, International Journal of Research in Business and Technology, Techmind Research, **4** (2014), no. 3, pp. 531-533.
[8] U. Abou El-Enien, Why unified statistics theory by MCMC towards estimation of stationary transition probabilities of stochastic matrix?, Journal of Computer and Mathematical Sciences, 6(2015), no. 7, pp. 411-415.
[9] T. Apostol, Calculus, vol.II, John Wiley, Inc., 1969.

Author

Dr. Usama H. M. Abou El-Enien is an assistant professor of math and statistics. - Department of Administrative Information Systems- The High Institute for Tourism, Hotels & Computer. His research interests in Markov Chains and MCMC