Journal of Al-Nahrain University Vol. 9(1), June, 2006, pp.108-117 Science

An Evolution Strategy for Likelihood Estimator of ARMA (1, 1) Model

[1]Dr. Bara'a Ali Attea, [1]Basa'd Ali Hussain
[1]Baghdad University, Iraq/ Baghdad

Abstract

This paper presents a multi-membered evolution strategy $(\mu+\lambda)$-ES for optimizing the maximum likelihood function $ln(L(\phi_1,\theta_1))$ of the mixed model $ARMA(1,1)$. The presented evolution strategy is composed of three main steps: local recombination, mutation, and selection. The experimental design is based on simulating the ES algorithm with different values of (ϕ,θ), and sample size n. The results are compared with those of moment method. Depending on MSE value obtained from both methods, one can conclude that $(\mu+\lambda)$-ES can give good estimators (ϕ_1,θ_1) of $ARMA(1,1)$ parameters and more reliable than estimators obtained by moment method.

Key Words: $ARMA(1,1)$, likelihood function, moment, $(\mu+\lambda)$-ES, local recombination

Introduction

A time series is an order sequence of observations in equal interval space; this ordering is through time or other dimensions such as space. Time series occur in variety of fields (e.g., agriculture, business, economics, and engineering). The series is formally represented in a stochastic model known as mixed autoregressive-moving average model, ARAM (p, q) [1]:

$$z_t=\phi z_{t-1}+\dots+\phi_p z_{t-p}+a_t-\theta a_{t-1}-\dots-\theta_q a_{t-q} \quad \text{............(1)}$$

or

$$\phi(B)z_t=\theta(B)a_t \quad \text{....................................(2)}$$

Which employs $(p+q+2)$ unknown parameters: $\mu,\phi_1,\dots,\phi_p,\theta_1,\dots,\theta_q$ and σ_a^2 to be estimated from data series. Sometime time series has a non-stationary

behavior about a fixed mean. This kind of behavior can be represented by denoted by $ARMA(1,1)$, takes the formula [1][2]:

(p,d,q) model for the $d'th$ difference of the data series to be stationary [1][2]:

$$\phi(B)(1-B)^d z_t=\theta(B)a_t \quad \text{.......................(3)}$$

$$\phi(B)w_t=\theta(B)a_t \quad \text{...................................(4)}$$

where

$$w_t=\nabla^d z_t \quad \text{...(5)}$$

Model in (Eq. 4) is powerful for describing stationary and non-stationary time series and is called an autoregressive integrated moving average ARIMA model of order (p,d,q). This model is defined as [1]:

$$w_t=\phi_1 w_{t-1}+\dots+\phi_p w_{t-p}+a_t-\theta a_{t-1}-\dots-\theta a_{t-q} \quad \text{......(6)}$$

With $w_t=\nabla^d z_t$.

Mixed model of first order, known as first order autoregressive-moving average model, denoted by $ARMA(1,1)$, takes the formula [1][2]:

$$z_t=\phi_1 z_{t-1}+a_t-\theta_1 a_{t-1} \quad \text{..............................(7)}$$

or

$$(1-\phi_1 B)z_t=(1-\theta_1 B)a_t \quad \text{...............................(8)}$$

where a_t the error term and has $(i.i.d)N(0,\sigma_a^2)$ distribution. This model is stationary if the root of $(1-\phi_1 B)=0$ lies

outside the unit circle and invertible if the root of $(1-\theta_1 B)=0$ lies outside the unit, so

we can get that $|\phi_1|<1$, and $|\theta_1|<1$. The moment of this stationary system is [2]:

Mean function

$$E(z)=0 \quad \text{..(9)}$$

Variance function is

$$\gamma_0=var(z_t)=\frac{1+\theta_1^2-2\phi_1\theta_1}{1-\phi_1^2}\sigma_a^2 \quad \text{...............(10)}$$

Auto-covariance function

$$\gamma_k=\begin{bmatrix}\dfrac{(\phi_1-\theta_1)(1-\phi_1\theta_1)}{1-\phi_1^2}\sigma_a^2 & k=1\\ \phi_1\gamma_{k-1} & k\geq 2\end{bmatrix} \quad \text{.......(11)}$$

ISSN:1814-5922

Autocorrelation function

$$\rho_i = \begin{cases} 1 & k=0 \\ \dfrac{(\phi_i-\theta_i)(1-\phi_i\theta_i)}{1+\theta_i^2-2\phi_i\theta_i} & k=1 \\ \phi_i\rho_{k-1} & k\geq 2 \end{cases} \quad \ldots\ldots\ldots\ldots(12)$$

So, we can show that the marginal distribution of time series which has ARMA(1,1) model is:

$$z_t \sim N\left(0, \frac{1+\theta_i-2\phi_i\theta_i}{1-\phi_i^2}\sigma_a^2\right), \text{ If } \quad a_t \sim N(0,\sigma_a^2)$$

The analysis of time series has four steps [3]: identification, estimation, diagnostic checks, and forecasting or control. The second step, estimation, means efficient use of the data to make inference about parameters conditional to the adequacy of entertained model.

The model under study is nonlinear as $a_t = \dfrac{1-\phi_i B}{1-\theta_i B} z_t$ is nonlinear. As a result, there is no direct method to estimate the model's parameters. However, one can use indirect method (i.e. iterative method). This

method starts with an initial value and then modifies this value iteratively using some numerical algorithms [3].

Mathematical Part

The likelihood function is one of fundamental importance in estimation theory. This principle says that the data has to tell us about the parameters contained in the likelihood function, all other aspects of the data being irrelevant. In moderate and large samples, the likelihood function will be unimodel and can be adequately approximated over a sufficiently extensive region near the maximum by a quadratic function. Hence, in these cases the log-likelihood function can be described by its maximum and its second derivatives at the maximum. The values of parameters which maximize the likelihood function, or equivalently the log-likelihood function, are called maximum likelihood (ML) estimates. The second derivatives of the log – likelihood provide measurers of "spread" of the likelihood function and can be used to calculate approximate standard errors for the estimates [2].

Now, to study the likelihood function of ARAM(1,1) let as suppose the $N = n + d$ original observations z, from a time series which can be denoted by $z_{-d+1},\ldots,z_0,z_1,z_2,\ldots,z_n$. We assume that this series is generated by an $ARMA(1,1)$ model. From these observations, we can generate a series w of $n = N - d$ differences $w_1, w_2, \ldots, w_n$, where $w_t = \nabla^d z_t$. The stationary mixed ARMA (1, 1) model in Eq. 7 may be written as [2]:

$$a_t = w_t - \phi_i w_{t-1} + \theta_i a_{t-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(13)$$

Where $E(w_t) = 0$. Suppose that $\{a_t\}$ has the normal distribution with zero mean and constant variance equal to , then the likelihood function can get as follows [2]:

$$L = (2\pi\sigma^2)^{-\frac{n}{2}} \left|M^{(1,1)}\right|^{\frac{1}{2}} \exp\left(\frac{-s(\phi_i,\theta_i)}{2\sigma_a^2}\right) \ldots\ldots\ldots(14)$$

Where

$$M^{(1,1)} = var\text{-}cov(\phi_i,\theta_i) = I^{-1}(\phi_i,\theta_i)$$

$$= \frac{1}{I(\phi_i,\theta_i)} adj(I(\phi_i,\theta_i)) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (15)$$

where

$$I(\phi_i,\theta_i) = \frac{n}{\sigma_a^2}\begin{bmatrix} \dfrac{\sigma_a^2}{1-\phi_i^2} & \dfrac{-\sigma_a^2}{1-\phi_i\theta_i} \\ \dfrac{-\sigma_a^2}{1-\phi_i\theta_i} & \dfrac{\sigma_a^2}{1-\theta_i^2} \end{bmatrix}$$

$M^{(1,1)} = I^{-1}(\phi_i,\theta_i)$, then the log-likelihood function is:

$$\ln(L) = -\frac{n}{2}(2\pi\sigma_a) + \frac{1}{2}\ln(|M^{(1,1)}|) - \frac{s(\phi_i,\theta_i)}{2\sigma_a^2} \ldots\ldots\ldots(16)$$

where:

$$s(\phi_i,\theta_i) = \sum_{t=-\infty}^{n} (a_t|\phi_i,\theta_i,w)^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(17)$$

is the sum squares errors, n is the sample size, and $[a_t|\phi_i,\theta_i,w] = E[a_t|\phi_i,\theta_i,w]$

denotes the expectation of a_t conditional on ϕ_i, θ_i and w. Sum squares errors can be found by unconditional calculation of the $[a]'s$ are computed recursively by taking conditional expectations in Eq. 13. A back-calculation provides the values

$[w_{-j}], j = 0,1,2,\ldots$ This back-forecasting needed to start off the forward recursion.

For moderate and large values of n in Eq. 17 is dominated by $s(\phi_i,\theta_i)/2\sigma_a^2$ and thus the contours of the unconditional sum squares function in the space of the parameters (ϕ, θ) are vary nearly contours of likelihood and of log likelihood . It follows, in particular, that the parameter estimates obtained by minimizing the sum of squares in Eq. 17, called least square estimates will usually provide very close approximation to the (maximum likelihood estimator).

Another method is Yule-Walker which is known by moment method, obtained by equating a sample moment such as sample mean $\bar{z}$, sample variance $\hat{\gamma}_0$, and sample ACF $\hat{\rho}_l$ to their theoretical values counterparts and solving the resultant equations. For the presented model, the approximate values for the parameters are obtained by substituting the estimators for ρ_1, and ρ_2 [1]:

$$\hat{\rho}_k = r_k = \frac{\sum_{t=1}^{n-k}(z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^{n}(z_t - \bar{z})^2} \quad \text{.........(18)}$$

$(\mu \overset{+}{,} \lambda)$ Evolution Strategies

H.-P. Schwefel proposed the multi-membered evolution strategies, the so-called $(\mu \overset{+}{,} \lambda)$-ES. In their most general form, these strategies are described in the coming subsections.

Representation and Fitness Evaluation

An individual $\vec{a} = (\vec{x}, \vec{\sigma}) \in I$ in $(\mu \overset{+}{,} \lambda)$-ES can consist of the components [4] [5]:

$\vec{x} \in R^n$: The vector of object variables.

$\vec{\sigma} \in R_+^{n_\sigma}$: A vector of step length or standard deviations $(1 \le n_\sigma \le n)$ of the normal distribution. The strategy parameter $\vec{\sigma}$ (also called the internal model) determines the variances of the n-dimensional normal distribution, which is used for exploring the search space. The user of an evolution strategy, depending on his feeling about the degree of freedom required, can vary the amount of strategy parameters attached to an individual. As a rule of thumb, the global search reliability increases at the cost of computing time when the number of strategy parameters is increased. The setting most commonly used which form the extreme cases are:

- $n_\sigma = 1$, (Uncorrelated mutation with one single standard deviation controlling mutation of all components of $\vec{x}$)
- $n_\sigma = n$: (Standard mutations with individual step sizes $\sigma_1, ..., \sigma_n$ controlling mutation of the corresponding object variables x_i individually)The only part of $\vec{a}$ entering the objective function evaluation is $\vec{x}$, and the fitness of an individual $\phi(\vec{a})$ is identical to its objective function value $f(\vec{x})$, i.e. $\phi(\vec{a}) = f(\vec{x})$.

Mutation Operator

The generalized structure of $(\mu \overset{+}{,} \lambda)$-ES mutation operator consists of the addition of a normally distributed random number to each component of the object variable vector, corresponding to a step in the search space. The variance of the step-size distribution is itself subject to mutation as a strategy variable. Formally speaking, mutation operator $m_{\{\tau_0, \tau\}} : I \rightarrow I$, is defined as follows [4][6]

$$m_{\{\tau_0,\tau\}}(\vec{a}) = m_x(\vec{x}) \circ m_\sigma(\vec{\sigma}) = (\vec{x}', \vec{\sigma}') \quad \text{........(19)}$$

Which proceeds by first mutating the strategy parameters $\vec{\sigma}$:

$m_\sigma : R_+^{n_\sigma} \rightarrow R_+^{n_\sigma}$

$$m_\sigma(\vec{\sigma}) = \vec{\sigma}' = (\sigma_1 \exp(\varepsilon_1 + \varepsilon_0), ..., \sigma_{n_\sigma} \exp(\varepsilon_{n_\sigma} + \varepsilon_0)) \quad \text{......(20)}$$

Where

$\varepsilon_0 \sim N(0, \tau_0^2)$, $\varepsilon_i \sim N(0, \tau^2)$ $\forall i \in \{1, ..., n_\sigma\}$

To prevent standard deviations from becoming practically zero, a minimal value of ε_σ is algorithmically enforced for all σ_i.

Secondly, modifying $\vec{x}$ according to the new set of strategy parameters obtained from mutating $\vec{\sigma}$:

$m_x : R^n \rightarrow R^n$

$$m_x(\vec{x}) = \vec{x}' = (x_1 + z_1, ..., x_n + z_n) \quad \text{.........(21)}$$

Recombination Operators

In $(\mu \overset{+}{,} \lambda)$-ESs, different recombination mechanisms are used in either local form, producing one new individual from two randomly selected parent individuals, or in global form, allowing components to be taken for new individual from potentially all individuals available in the parent population. Furthermore, recombination is performed on strategy parameters as well as the object variables, and the recombination type may be different for object variables, and standard deviations.

Depending on the recombination type rec [4][6]:

$$rec = \begin{cases} 0 & \text{No recombination} \\ 1 & \text{Discrete recombination of pair of parents} \\ 2 & \text{Intermediate recombination of pair of parents} \\ 3 & \text{Discrete recombination of all parents} \\ 4 & \text{Intermediate recombination of all parents in pair} \end{cases}$$

Sometimes, the choice of a useful recombination operator for a particular optimization problem is relatively difficult and requires performing some experiments [2].The rules of recombination operator $r : I^{\mu} \to I$ for creating an individual, $r_{\{rec_x, rec_\sigma\}}(P) = \vec{a}' = (\vec{x}', \vec{\sigma}') \in I$, are given respectively by referring to arbitrary vectors $\vec{b}$ and $\vec{b}'$ where $\vec{b}$ and $\vec{b}'$ denote here the part (i.e., either $\vec{x}$ or $\vec{\sigma}$) of a pre-selected parent individuals and the part of an offspring vector receptively. Each of $\vec{b}$ and $\vec{b}'$ are of length $m \in \{n, n_\sigma\}$;

$(\forall i \in \{1,...,m\})$

$$b_i' = \begin{cases} b_i & \text{if } rec = 0 \\ b_{\chi_1,i} \text{ or } b_{\chi_2,i} & \text{if } rec = 1 \\ (b_{\chi_1,i} + b_{\chi_2,i}) \cdot 0.5 & \text{if } rec = 2 \\ b_{\chi_3,i} & \text{if } rec = 3 \\ (b_{\chi_3,i} + b_{\chi_4,i}) \cdot 0.5 & \text{if } rec = 4 \end{cases} \quad (21)$$

where $\chi_1, \chi_2 \sim U(\{1,...,\mu\})$ for each offspring, and $\chi_3, \chi_4 \sim U(\{1,...,\mu\})$ for each i.

Selection Operator

There are two main classifications for selection according to the survival property of the parents [6]:

Extinctive_ (μ, λ) strategy; where parents live for a single generation only.

$S : I^{\lambda} \to I^{\mu}$

$S(P) = P'$,where $|P| = \lambda$ and $|P'| = \mu$,and

$\forall \vec{a}' \in P' : \exists \vec{a} \in P - P' : f(\vec{x}) \le f(\vec{x}')$

Preservative_ $\mu + \lambda$ strategy; where selection operates on the joined set of parents and offspring, i.e., very fit individuals may survive indefinitely:

$S : I^{\mu+\lambda} \to I^{\mu}$

$S(P) = P'$, where $|P| = \mu + \lambda$, $|P'| = \mu$, and

$\forall \vec{a}' \in P' : \exists \vec{a} \in P - P' : f(\vec{x}) \le f(\vec{x}')$

The ratio μ/λ is known as selection pressure. In the choice of μ and λ, there is no need to ensure that λ is exactly divisible by μ. The association of offspring to parents is made by a random selection of evenly distributed random integers from the range $[1, \mu]$. It is only necessary that λ exceeds μ by a sufficient margin that on average at least one offspring can be better than its parent. Hoffeister and Bäck in [7] have stated that $\mu/\lambda \approx 1/6$ are tuned for a maximum rate of convergence, and as a result tend to reduce their genetic variability, i.e., the number of different alleles (specific parameter setting) in a population, as soon as they are attracted by some local optimum.

Conceptual ES Algorithm

Combining mutation, recombination, and selection as defined in the previous subsections, the conceptual algorithm can then be formulated as [4][5]:

t:= 0; // t is the generation number Initialize

$P(0) := \{\vec{a}_1(0),...,\vec{a}_\mu(0)\} \in I^{\mu}$

Where $I = R^{n+n_{sp}}$ and

$\vec{a}_k = (x_i, \sigma_j \quad \forall i \in \{1,...,n\}, \forall j \in \{1,...,n_\sigma\})$;

Evaluate $P(0) : \{\phi(\vec{a}_1(0)),...,\phi(\vec{a}_\mu(0))\}$

where $\phi(\vec{a}_k(0)) = f(\vec{x}_k(0))$

While ($\iota(P(t)) \ne true$) do

Recombine:

$\vec{a}_k'(t) := r_{\{rec_x, rec_\sigma\}}(P(t)) \quad \forall k \in \{1,...,\lambda\}$;

Mutate:

$\vec{a}''(t) := m_{\{\tau_0,\tau\}}(\vec{a}'_k(t)) \forall k = \{1,\dots,\lambda\};$

Evaluate:

$P''(t) := \{\vec{a}_1''(t),\dots,\vec{a}_\lambda''(t)\} : \{\phi(\vec{a}_1''(t)),\dots,\phi(\vec{a}_\lambda''(t))\}$

Where $\phi(\vec{a}_k''(t)) = f(\vec{x}_k''(t));$

Select:

$P(t+1) := if\ (\mu,\lambda)\_selection$

then

$S(P''(t));$

else

$S(P(t) \cup P''(t));$

$t = t+1;$

End

Experimental Results

In this section, the conceptual algorithm for $(\mu+\lambda)$-ES is adopted for the likelihood estimator of ARMA (1, 1). μ and λ are set to 30 and 200 respectively. The experimental results performed here are based on two different sample size (i.e. $n = 25,50$), ϕ set to $(\pm 0.1, \pm 0.4)$ and θ set to $(\pm 0.2, \pm 0.5)$. The random sample are generated using Box-Muller formula provided by randan statement in Matlab workspace. All results were obtained by running each experiment 5 different runs and each iterates with 150 generations and averaging the resulting data. Moreover, the initial values for strategy parameters $\vec{\sigma}$ are set to 3.0. Further, the results of $(\mu+\lambda)$-ES are compared with those obtained by moment method (with 1000 runs). The comparison made was based on Mean square error (MSE): $MSE = var(\phi) + bias$. Results are given in figure 1-2 (a, b, c, and d) and table 1. The experiments on a set of data give some impressions of the behaviors of both ES and moment methods. As one can see that the MSEs of ES are smaller than those of moment. This indicates that ES is more reliable than moment to give estimator of the parameters of the model under study.

Moreover, one can see that value of MSE decreases as the sample size increase. For ES one can also see that the value of sum square increases when increasing the number of generation and sample size. Also, from figures below, we can see that the mean objective function is more stable than both best and worst objective functions. The behavior of ES when the objective function parameters (ϕ_i, θ_i) take positive values is better than when they are negative. Finally, the behavior of ES, also, reflects the maximum likelihood estimators are very close to least squares estimators.

Conclusion

In this paper, we present a multi-membered $(\mu+\lambda)$-ES for estimating parameters of log-likelihood function of mixed model of first order. It provides effective results for two random samples with different sizes. By comparing (based on MSE) ES with moment method, we found that more reliable results can be obtained by ES.

sample size	ϕ	θ	MSE of moment		MSE of ES	
			ϕ	θ	ϕ	θ
25	.1	.2	.366	.332	.084	.122
	-.1	-.2	.319	.344	.062	.121
	.4	.5	.446	.452	.013	.104
	-.4	-.5	.621	.715	.059	.136
50	.1	.2	.313	.281	.072	.111
	-.1	-.2	.303	.291	.053	.045
	.4	.5	.386	.415	.021	.067
	-.4	-.5	.668	.551	.098	.194

Table (1) MSE compression of moment and $(\mu+\lambda)$-ES

References

[1] W. W. S. Wei, Time Series Analysis: Univariate and Multivariate Methods, Wesely Publishing com. INC, New York, 1990.

[2] G. E. P. Box, and G. M. Jenkins, "Time series analys is: forecasting and control", Holden-Day, San Francisco, 1976.

[3] S. Makridakis, S. C. Wheelwright, Forecasting Method and Application, John Wiley& sons, New York, 1999.

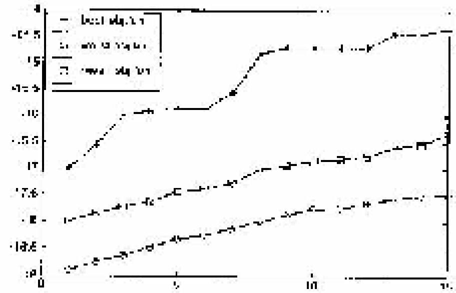
[4] T. Bäck, and H,-P, Schwefel, "An overview of evolutionary algorithms for parameter optimizations", Evolutionary Computation, Cambridge, Vol. 1, No. 1, pp. 1-23, 1993.

[5] T. Bäck, and H,-P, Schwefel, "Evolution strategies I: variants and their computational implementation", Genetic Algorithms in Engineering and Computer Science, pp. 11-26, 1995.

[6] H,-P, Schwefel, Numerical Optimization of Computer Models. New York: John Wiley & Sons. 1981.

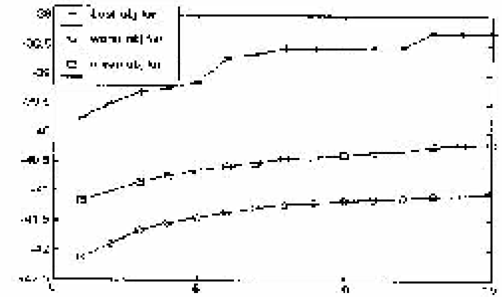
[7] F. Hoffmeister , and T. Back, "Genetic Algorithms and Evolutionary Strategies: Similarities and Differences", Parallel

Problem Solving from Nature PPSN I, Vol. 496, pp.455-469,



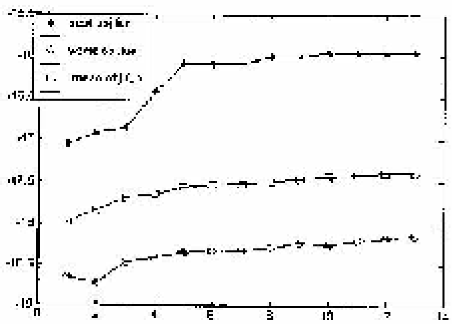
$\phi_1 = .4, \theta_1 = .5, n = 25$

-a-



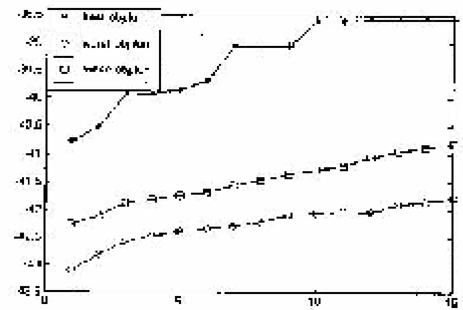
$\phi_1 = .4, \theta_1 = -.5, n = 25$

-b-



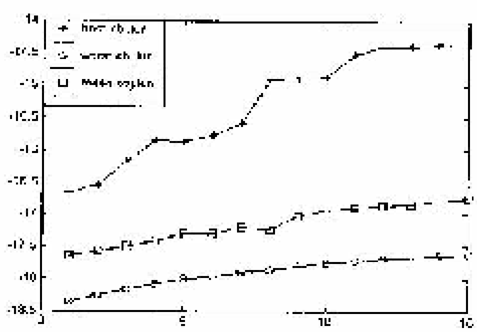
$\phi_1 = -.4, \theta_1 = -.5, n = 25$

-c-



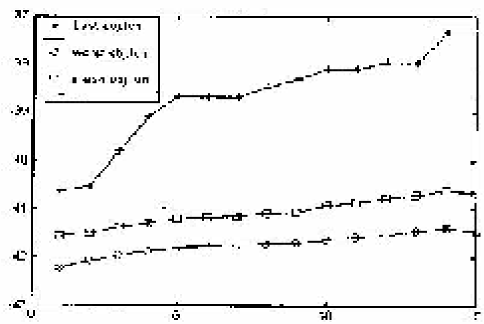
$\phi_1 = .4, \theta_1 = -.5, n = 50$

-e-



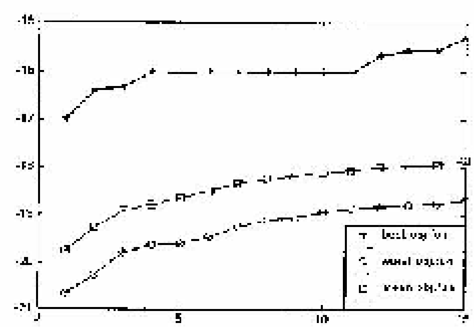
$\phi_1 = .1, \theta_1 = .2, n = 25$

-a1-



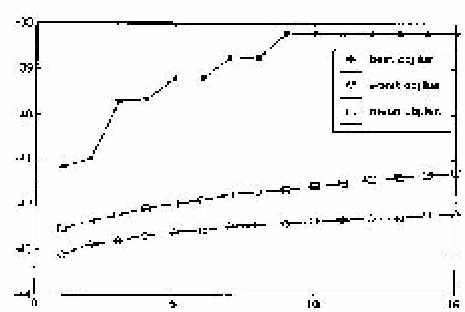
$\phi_1 = .1, \theta_1 = .2, n = 50$

-b1-



$\phi_1 = -.1, \theta_1 = -.2, n = 25$

-c1-



$\phi_1 = -.1, \theta_1 = -.2, n = 50$

-d1-

Figure 1: Results of ES for 150 generations depicted as: blue ("*") (best objective function), red (33 "O") mean function, and green ("□") worst for 5 different runs

Table 1-a: Results where $\phi_1 = .4, \theta_1 = .5, n = 25$

best			Worst			mean		
Φ_1	Θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$
0.37124	0.4424	-17.006	0.4098	0.4792	-18.9125	0.3995	0.4872	-17.9939
0.41398	0.4385	-16.569	0.4268	0.5169	-18.7482	0.3983	0.4732	-17.8597
0.38368	0.3942	-15.975	0.4086	0.4954	-18.6255	0.3963	0.4604	-17.7205
0.39298	0.4035	-15.907	0.4591	0.5019	-18.4643	0.3964	0.4526	-17.6271
0.3923	0.4021	-15.865	0.4257	0.4673	-18.3205	0.3894	0.4378	-17.4485
0.3923	0.4021	-15.865	0.4352	0.4746	-18.243	0.3780	0.4223	-17.3841
0.3953	0.3953	-15.551	0.3970	0.4940	-18.1169	0.3820	0.4218	-17.2618
0.4604	0.4604	-14.808	0.3942	0.4508	-17.9846	0.3805	0.4137	-17.0195
0.4604	0.4605	-14.698	0.3812	0.4544	-17.8656	0.3774	0.4071	-16.9323
0.4604	0.4605	-14.698	0.3578	0.3954	-17.7565	0.3721	0.3969	-16.8499
0.4048	0.4048	-14.69	0.381	0.4311	-17.7383	0.3713	0.3949	-16.8122
0.4048	0.4048	-14.69	0.2863	0.3047	-17.6515	0.3740	0.3935	-16.7575
0.2857	0.2858	-14.441	0.3906	0.3905	-17.5583	0.3724	0.3887	-16.6026
0.2857	0.2858	-14.441	0.3819	0.4182	-17.5067	0.3682	0.3833	-16.5248
0.3560	0.3361	-14.750	0.3635	0.4048	-17.4872	0.3641	0.3774	-16.3303

Table 2-b: Results where $\phi_1 = .4, \theta_1 = .5, n = 50$

best			worst			mean		
Φ_1	Θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$
0.267	0.26721	-14.7346	0.34539	0.357425	-17.2439	0.34037	0.35393	-16.25032
-0.36975	-0.44358	-17.048	-0.4176	-0.50554	-18.652	-0.4092	-0.4823	-17.98214
-0.398	-0.45304	-16.9199	-0.452	-0.4964	-18.7251	-0.412	-0.4752	-17.83414
-0.420575	-0.45332	-16.8524	-0.4407	-0.46074	-18.4747	-0.4169	-0.473	-17.70076
-0.4529	-0.4651	-16.4307	-0.4436	-0.4629	-18.4169	-0.4195	-0.4719	-17.6518
-0.3896	-0.42418	-16.0793	-0.4345	-0.47648	-18.3467	-0.4219	-0.4709	-17.56892
-0.387675	-0.42276	-16.0731	-0.4264	-0.48676	-18.3319	-0.4218	-0.4701	-17.53168

Best			worst			mean		
Φ_1	Θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	Θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	Θ	$\ln(L(\Phi_1, \theta_1))$
0.3886	0.4437	-39.7547	0.29472	0.3917	-42.1224	0.3887	0.4699	-41.1588
0.4078	0.4448	-39.5	0.4119	0.4797	-41.9021	0.3948	0.46356	-41.0214
0.4282	0.4501	-39.3078	0.47332	0.4954	-41.6592	0.3929	0.44828	-40.8489
0.4108	0.4323	-39.2413	0.42412	0.5068	-41.5474	0.4013	0.4503	-40.737
0.4112	0.4332	-39.1194	0.42184	0.5019	-41.4557	0.3991	0.44344	-40.6508
0.4244	0.4344	-38.7183	0.4565	0.5166	-41.3572	0.3943	0.43332	-40.5929
0.4372	0.4371	-38.6559	0.44524	0.4831	-41.2947	0.3939	0.42902	-40.5178
0.4296	0.4296	-38.5656	0.35356	0.4117	-41.2247	0.3906	0.42318	-40.4469
0.4296	0.4296	-38.5656	0.3881	0.3891	-41.2082	0.3844	0.41582	-40.4264
0.4296	0.4296	-38.5656	0.33716	0.3731	-41.1627	0.3817	0.41104	-40.3811
0.4295	0.4296	-38.5328	0.34396	0.3433	-41.1525	0.3828	0.41036	-40.3397
0.4295	0.4296	-38.5328	0.434	0.4528	-41.15	0.3846	0.41146	-40.3220
0.4229	0.4229	-38.2937	0.43682	0.45608	-41.1008	0.3855	0.41054	-40.2408
0.4229	0.4229	-38.2937	0.39586	0.4508	-41.0688	0.3825	0.40754	-40.2235
0.4229	0.42298	-38.2937	0.45002	0.468	-41.041	0.3824	0.40564	-40.2092

Table 3 - c : Results where $\phi_1 = -.4, \theta_1 = -.5, n = 25$

-0.387675	-0.42276	-16.0737	-0.4359	-0.49734	-18.3264	-0.4237	-0.4714	-17.52094
-0.4385	-0.46354	-15.9844	-0.4645	-0.50592	-18.2922	-0.4252	-0.4702	-17.4867
-0.4385	-0.46354	-15.9844	-0.4547	-0.47368	-18.2298	-0.4263	-0.4692	-17.46854
-0.465625	-0.4745	-15.9442	-0.4156	-0.48458	-18.2579	-0.4282	-0.4691	-17.42966
-0.46686	-0.4745	-15.9442	-0.406	-0.44592	-18.2029	-0.4287	-0.4672	-17.40696
-0.466867	-0.4745	-15.9442	-0.4052	-0.48276	-18.1729	-0.4278	-0.4651	-17.39202
-0.46686	-0.4745	-15.9442	-0.4283	-0.48072	-18.1485	-0.428	-0.4627	-17.38182

Table 4- d : Results where $\phi_1 = -.4, \theta_1 = -.5, n = 50$

best			worst			mean		
Φ_1	Θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$
-0.42784	-0.4604	-40.7702	-0.3995	-0.50128	-43.0991	-0.3895	-0.4743	-42.24868
-0.41566	-0.4455	-40.5115	-0.4044	-0.52302	-42.8219	-0.4053	-0.4728	-42.09772
-0.35196	-0.35206	-39.9205	-0.4034	-0.44754	-42.5725	-0.4003	-0.4547	-41.8662
-0.35196	-0.35206	-39.9205	-0.4219	-0.48868	-42.467	-0.3923	-0.4423	-41.80652
-0.3827	-0.38252	-39.8481	-0.4838	-0.5042	-42.3845	-0.3859	-0.4302	-41.74172
-0.48362	-0.48374	-39.6778	-0.4673	-0.50986	-42.3243	-0.3865	-0.4273	-41.6773
-0.29074	-0.2908	-39.0547	-0.4528	-0.4943	-42.2728	-0.3777	-0.4144	-41.5418
-0.29074	-0.2908	-39.0547	-0.3271	-0.38956	-42.1995	-0.3532	-0.387	-41.46392
-0.29074	-0.2908	-39.0547	-0.4363	-0.45666	-42.0763	-0.3447	-0.3725	-41.34148
-0.30238	-0.30242	-38.6066	-0.5189	-0.55746	-42.0413	-0.3336	-0.3608	-41.27334
-0.30238	-0.30242	-38.6066	-0.4614	-0.48202	-42.0136	-0.3216	-0.3461	-41.20228
-0.18588	-0.1859	-38.5763	-0.3941	-0.47054	-42.0321	-0.2972	-0.3211	-41.0559
-0.18588	-0.1859	-38.5763	-0.3344	-0.37104	-41.8953	-0.2843	-0.3044	-40.9533
-0.18588	-0.1859	-38.5763	-0.3122	-0.38236	-41.8452	-0.2574	-0.2603	-40.87933
-0.18588	-0.1859	-38.5763	-0.2218	-0.2614	-41.7886	-0.2693	-0.2852	-40.81728

Table 1 - a1 : Results where $\phi_1 = .1, \theta_1 = .2, n = 25$

best			worst			mean		
Φ_1	Θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$		θ	$\ln(L(\Phi_1, \theta_1))$
0.1122	0.1266	-16.6647	0.1193	0.2387	-18.3586	0.10488	0.1826	-17.64972
0.1418	0.1423	-16.543	0.0422	0.1076	-18.2521	0.10192	0.17	-17.582
0.1472	0.1472	-16.1596	0.1098	0.2197	-18.1611	0.09896	0.1604	-17.50382
0.1666	0.1668	-15.8455	0.1067	0.2133	-18.0917	0.10226	0.1578	-17.42298
0.0912	0.0913	-15.8596	0.0872	0.1649	-18.0011	0.1027	0.1488	-17.29275
0.1112	0.1114	-15.7755	0.1065	0.1867	-17.9608	0.10026	0.1460	-17.29876
0.1154	0.1158	-15.5935	0.0872	0.1477	-17.895	0.10004	0.1419	-17.20818
0.0701	0.0704	-14.9122	0.1216	0.1606	-17.8644	0.1002	0.1308	-17.24982
0.0742	0.0742	-14.9038	0.1174	0.1929	-17.8108	0.0975	0.1321	-17.02084
0.1269	0.1269	-14.8811	0.108	0.1825	-17.754	0.09854	0.1293	-16.94276
0.1271	0.1271	-14.5201	0.0067	0.0813	17.7345	0.09832	0.1276	-16.89696
0.1196	0.1196	-14.3989	0.0565	0.1127	-17.6949	0.08608	0.1132	-16.86142
0.1196	0.1196	-14.3989	0.1014	0.1575	-17.6755	0.08808	0.1147	-16.84546
0.1171	0.1169	-14.3793	0.1167	0.1712	-17.6405	0.09294	0.1179	-16.77806
0.1171	0.1169	-14.3793	0.1148	0.1506	-17.6308	0.08988	0.1129	-16.75816

Table 3 - c1 : Results where $\phi_1 = -.1, \theta_1 = -.2, n = 25$

best			worst			mean		
Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$
0.1278	0.12758	-14.6313	0.10532	0.13581	-17.473	0.09712	0.12129	-16.60974
-0.313	-0.3134	-16.9926	-0.0573	-0.1011	-20.677	-0.1621	-0.1879	-19.76536
-0.3043	-0.3044	-16.3973	-0.1458	-0.1582	-20.269	-0.1923	-0.2085	-19.29336
-0.2877	-0.2878	-16.328	-0.1528	-0.1573	-19.800	-0.191	-0.2011	-18.91352
-0.2432	-0.2431	-16.0091	-0.2483	-0.2436	-19.636	-0.245	-0.2553	-18.7953
-0.2311	-0.2310	-16.0332	-0.3248	-0.3265	-19.602	-0.1917	-0.1995	-18.65694
-0.2155	-0.2155	-16.0086	-0.174	-0.1985	-19.460	-0.1927	-0.1992	-18.49306
-0.1347	-0.1346	-16.0034	-0.2048	-0.2274	-19.24	-0.2128	-0.2185	-18.33906
-0.1347	-0.1346	-16.0034	-0.2054	-0.2049	-19.118	-0.2276	-0.2314	-18.25186
-0.1347	-0.1346	-16.0034	-0.2445	-0.2652	-19.051	-0.2313	-0.2351	-18.17706
-0.1347	-0.1346	-16.0034	-0.2944	-0.2938	-18.936	-0.2333	-0.2364	-18.11332
-0.1347	-0.1346	-16.0034	-0.2996	-0.3003	-18.863	-0.2359	-0.239	-18.0721
-0.1375	-0.1374	-15.6724	-0.2439	-0.2590	-18.825	-0.2404	-0.2436	-18.01396
-0.1531	-0.1531	-15.5651	-0.267	-0.2692	-18.797	-0.2391	-0.2418	-17.9852
-0.1531	-0.1531	-15.5651	-0.2709	-0.2716	-18.772	-0.2396	-0.2422	-17.96144
-0.0643	-0.0643	-15.3063	-0.1631	-0.1839	-18.69	-0.2457	-0.2476	-17.87354

Table 4 - b1 : Results where $\phi_1 = -.1, \theta_1 = .2, n = 50$

Best			W0rst			mean		
Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$
0.10168	0.13096	-40.6356	0.12684	0.2183	-42.2337	0.1026	0.18484	-41.57858
0.07518	0.10484	-40.5377	0.1453	0.23024	-42.0923	0.09014	0.16262	-41.51268
0.14118	0.14084	-39.7993	0.05898	0.10166	-41.9713	0.08656	0.14652	-41.39522
0.07714	0.0769	-39.1025	-0.0813	-0.02224	-41.8945	0.0852	0.13918	-41.30694
0.08886	0.0887	-38.6991	0.10358	0.17986	-41.8137	0.09804	0.14476	-41.22656
0.08886	0.0887	-38.6991	0.15242	0.20796	-41.751	0.09802	0.1401	-41.1897
0.08886	0.0887	-38.6991	0.15114	0.18774	-41.7552	0.09154	0.13178	-41.1744
0.01764	0.0175	-38.4942	0.20804	0.24426	-41.7239	0.08492	0.12146	-41.10458
0.04726	0.0471	-38.3169	0.18782	0.22474	-41.7125	0.08448	0.12038	-41.07578
0.11962	0.1196	-38.1174	-0.0265	-0.0079	-41.6475	0.08656	0.11716	-40.91492
0.11962	0.1196	-38.1174	0.15568	0.17243	-41.5947	0.09564	0.12392	-40.87046
0.129	0.1289	-37.9657	0.15662	0.18932	-41.5472	0.07542	0.09904	-40.78816
0.129	0.1289	-37.9657	0.20164	0.21658	-41.4621	0.0362	0.05594	-40.7188
0.03738	0.0373	-37.3376	-0.1261	-0.09552	-41.3863	0.02354	0.04172	-40.60504
0.03738	0.0373	-37.3376	-0.0452	-0.02942	-41.488	0.02544	0.04564	-40.68356

Table 4 - d1 : Results where $\phi_1 = -.1, \theta_1 = -.2, n = 50$

best			worst			mean		
Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$	Φ_1	θ	$\ln(L(\Phi_1, \theta_1))$
-0.14132	-0.1578	-41.1767	-0.1554	-0.22706	-43.1062	-0.105	-0.1842	-42.5445
-0.1773	-0.17734	-41.0011	-0.1331	-0.2251	-42.9026	-0.1152	-0.1776	-42.38158
-0.14036	-0.13986	-39.6972	-0.1238	-0.21314	-42.8157	-0.1198	-0.1753	-42.258
-0.11482	-0.11466	-39.6593	-0.1362	-0.2012	-42.6854	-0.125	-0.1673	-42.10034
-0.11548	-0.11518	-39.1988	-0.1528	-0.1938	-42.6297	-0.1269	-0.1631	-41.9978
-0.11548	-0.11518	-39.1988	-0.1009	-0.16036	-42.5749	-0.125	-0.1564	-41.8984
-0.14956	-0.14946	-38.7442	-0.1249	-0.1661	-42.4851	-0.1251	-0.1508	-41.77762

-0.14956	-0.14946	-38.7442	-0.1191	-0.13974	-42.4547	-0.125	-0.148	-41.71538
-0.17026	-0.17022	-38.2114	-0.1775	-0.21614	-42.42	-0.1266	-0.1481	-41.66762
-0.17026	-0.17022	-38.2114	-0.1357	-0.19366	-42.3681	-0.1256	-0.1422	-41.58146
-0.17026	-0.17022	-38.2114	-0.1795	-0.1966	-42.3325	-0.1267	-0.1408	-41.51816
-0.17026	-0.17022	-38.2114	-0.2096	-0.22603	-42.2974	-0.1215	-0.133	-41.45592
-0.17026	-0.17022	-38.2114	0.1016	-0.14178	-42.282	-0.1183	-0.1286	-41.39396
-0.17026	-0.17022	-38.2114	-0.1479	-0.16454	-42.2226	-0.1181	-0.1247	-41.33912
-0.17026	-0.17022	-38.2114	-0.1551	-0.1866	-42.2003	-0.1183	-0.1243	-41.31452

المستخلص

يعرض هذه البحث استراتيجية ES- $(\mu + \lambda)$ - المتعددة في حساب القيمة العظمى لدالة الإمكان الأعظم $\ln(L(\phi,\theta))$ للنموذج المختلط $ARMA(1,1)$ حيث أن القيم التي تعظم هذه الدالة هي مُخمّنات معلمات النموذج والتي تعظم مشتقة دالة الإمكان . تتكون استراتيجية -ES من ثلاث خطوات رئيسية\recombination, mutation و selection صممت التجارب بالاعتماد على محاكاة خوارزمية ESمع قيم مختلفة لمعلمات النموذج وحجم العينة .(n). قد أعطت هذه الطريقة نتائج جيدة الى المقدرات وذلك واضح من خلال البيانات المستحصلة وأيضا من خلال المقارنة التي أجريت بين طريقة العزوم المعروفة ب(Yule-Walker) و المستخدمة في حساب مقدري معلمتي النموذج وخوارزمية ES- $(\mu + \lambda)$ وذلك بالاعتماد على مقياس (MSE) الإحصائي ووضح ان ستراتيجية $(\mu+,\lambda)$ ESأعطت نتائج افضل حيث دلت نتائج التجارب أن قيمة MSE باستخدام ES$(\mu+,\lambda)$ افضل من النتائج التي أعطتها طريقة العزوم.