



RESEARCH ARTICLE - MATHEMATICS

Evaluation of Karimi criterion in Markov processes

Salah H Abid^{1*}, Nihad S Khalaf², Israa S Shihab³

¹ Mathematics department, College of Education, Mustansiriyah University, Bagdad, Iraq

^{2,3} Mathematics department, College of Education for Women, Tikrit University, Tikrit, Iraq

* Corresponding author E-mail: abidsalah@gmail.com

Article Info.

Abstract

Article history:

Received
February 2023

Accepted
April 2023

Publishing
30 June 2025

One of the ways that can be exercised in autoregressive model order chosen is to select the order that reduces the error of prediction. The final prediction error criterion employs this technique in order selection. Regrettably, this criterion has poor performance in case of finite samples. Karimi 2007 derived a criterion to address this problem. In this research, the Karimi criterion will be evaluated through the use of some distributions. These distributions are Discrete Uniform, Cauchy, t and Log normal, in addition to Gaussian distribution which is the basis of the Karimi criterion..

This is an open-access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>)

The official journal published by the College of Education at Mustansiriya University

Keywords: Use about five keywords and separate them by a semicolon.

1. Introduction

Consider the real Gaussian autoregressive (AR) process $w(\cdot)$ which is defined by [7],

$$w_t = \gamma_1 w_{t-1} + \gamma_2 w_{t-2} + \dots + \gamma_p w_{t-p} + u_t \quad , t = 1, 2, \dots, T \quad (1)$$

Where T is the sample size and u_t is the white noise of the AR model, independent, identically distributed (i.i.d.) Gaussian random process with zero mean and variance σ_u^2 . p is the order of the AR process and $\gamma_1, \gamma_2, \dots, \gamma_n$ are the real coefficients (parameters) of the process. We assume that $w(\cdot)$ is ergodic in terms of mean and covariance, so the poles of the AR model are inside the unit circle.

The AR model is a forecasting technique. It aims to forecast the observation sample based on previous observation samples by using the AR parameters as coefficients.

In Least square (LS) method, the AR model parameters in equation (1) are estimated by minimizing the error sum of squares, $Min (\sum_{t=1}^T u_t^2)$. It gives the linear systems equation from least squares normal equation as follows,

$$\begin{pmatrix} \sum_{t=p+1}^T w_{t-1}^2 & \sum_{t=p+1}^T w_{t-1}w_{t-2} & \dots & \sum_{t=p+1}^T w_{t-1}w_{t-p} \\ \sum_{t=p+1}^T w_{t-1}w_{t-2} & \sum_{t=p+1}^T w_{t-2}^2 & \dots & \sum_{t=p+1}^T w_{t-2}w_{t-p} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{t=p+1}^T w_{t-1}w_{t-p} & \sum_{t=p+1}^T w_{t-2}w_{t-p} & \dots & \sum_{t=p+1}^T w_{t-p}^2 \end{pmatrix} \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \\ \vdots \\ \hat{\gamma}_p \end{pmatrix} = \begin{pmatrix} \sum_{t=p+1}^T w_t w_{t-1} \\ \sum_{t=p+1}^T w_t w_{t-2} \\ \vdots \\ \sum_{t=p+1}^T w_t w_{t-p} \end{pmatrix} \quad (2)$$

This system can be solved using the orthonormal-upper triangular (QR) factorization method (Golub and Van Loan, 2013) [6].

The first step of AR modeling is order selection. A reasonable approach for order selection is to estimate the prediction error for each candidate order and to select the order that gives the minimum prediction error. "Prediction error" means the one-step prediction mean squared error for a realization of the process independent of the one observed. The Akaike's final prediction error (FPE) criterion was designed as an estimator of the prediction error. It is well-known that FPE is strongly biased in the

finite sample case, i.e., in the case that the number of given data (T) is not large compared to the maximum candidate order. The Akaike's final prediction error (FPE) criterion is [1,2],

$$FPE(k) = \frac{T+k}{T-k} \tilde{\sigma}_k^2, k = 1,2, \dots, m \tag{3}$$

Where m is the upper bound which can p take it and $\tilde{\sigma}_k^2$ is the residual variance, which is a measure of the fitness of the above model to the given data, is defined as follows,

$$\tilde{\sigma}_k^2 = \frac{1}{T-k} \sum_{i=k+1}^T (w_i - \sum_{j=1}^k \hat{\gamma}_j w_{i-j})^2 \tag{4}$$

Using this criterion, an estimate \hat{p} of the true process order p , that minimizes the prediction error is chosen such that,

$$FPE(\hat{p}) = \text{Min}\{FPE(k), k = 1,2, \dots, m\} \tag{5}$$

There are some modifications of the Akaike's final prediction error (FPE). Akaike in 1970b [3] modified his criterion to be,

$$FPE^\beta(k) = \frac{1+k/T^\beta}{1-k/T} \tilde{\sigma}_k^2 \tag{6}$$

Where $0 < \beta < 1$.

McClave 1975 [9] and Bhansali and Downham 1977 [5] suggested the following modification of Akaike's final prediction error criterion as,

$$FPE_\alpha(k) = (1 + \alpha k/T) \tilde{\sigma}_k^2 \tag{7}$$

Where $\alpha > 0$ and the increasing of α reduces the probability of fitting too high an order but for $\alpha \leq 1$ the asymptotic probability of overfitting is substantial.

Karimi in 2007 [8] derived new approximations for the expectations of residual variance and prediction error in the case that the AR parameter estimation method is least square estimation (LSE). These approximations are derived using the theoretical descriptions given in Akaike 1974 [4] for residual variance and prediction error of the LS method. Based on these new theoretical approximations, a modified FPE criterion is developed for AR model order selection to be valid for finite sample cases also. Karimi criterion is,

$$Kr(k) = \frac{1+k/(T-k)}{1-k/(T-k)} \tilde{\sigma}_k^2 \tag{8}$$

The paper aim is to evaluate of modified final production error criterion (MFPE) which is proposed by Karimi to determine the order of Autoregressive process (AR) according to different distributions of error term variate. These distributions are Gamma, Poisson, Exponential, Gumbel and Continuous Uniform Along with using the normal distribution as a distribution of error as a basis in deriving the criterion under consideration.

2. THE EMPIRICAL STUDY

For the purpose of evaluating the performance of Karimi criterion used to estimate the order of autoregressive model, a simulation experiment was done according to the following assumptions:

1. The following sample sizes $T = 10,25,50,100,250$ were used.
2. The Markov model was used with the values of the parameters that make the series in a different cases , stationary case with $\gamma = -0.9, -0.7, -0.5, -0.4, -0.3, -0.1, 0.1, 0.3, 0.5, 0.7, 0.8, 0.9$), random walk case with $\gamma = -1, 1$ and nonstationary case with $\gamma = -1.6, -1.1, 1.1$.
3. Discrete Uniform, Cauchy, t and Log normal and Gaussian Distributions were used as error distributions.
4. A lot of experiments were performed for all possible combinations of the above assumptions with a run size 500 for each time.

The following criteria were used for the purpose of investigating the performance of Karimi criterion in estimating the order of the autoregressive model,

1. The true selection ratio (TSR) from all 500 trials and for each studied case is calculated according to the following formula,

$$TSR = \frac{\text{number of times the estimated order matches the actual order of the model}}{500}$$

2. The mean squared error of estimating the model score

$$MSE = \frac{1}{500} \sum_{i=1}^{500} (\hat{p}_i - p_i)$$

Where \hat{P}_i represents the estimated order of autoregressive model according to Karimi criterion.

The analysis of the results after conducting these experiments will be done according to the assumptions stated previously. The results of each of the 500 trials will be noted and discussed deeply for each case in terms of the performance criteria values mentioned above.

3. THE RESULTS DISCUSSION

Bellow the results discussion for each case according to the error variable distribution.

(a) When the error distribution behavior is **Gaussian**, we note from the results presented in Table (1) that,

- 1- There is stability in the true selection ratio (TSR) and mean square error (MSE) in estimating the model order when the time series is stationary for medium and large sample sizes. We also notice a decrease in the mean square error values (MSE) and an increase in the true selection ratio (TSR) for small the samples, the further away the absolute value of the original parameter from zero.
- 2- We notice a distinguished performance in the nonstationary series for all sample sizes. The quality of that is an increase as the sample size increases.
- 3- In the case of the random walk, we notice the quality of this criterion for negative values is better than in the case of positive, and there is stability in the values of TSR and MSE for all sizes.
- 4- In general, when the residuals are normally distributed, the quality of the performance of the Karimi criterion in the cases of nonstationary series and random walk series, is better than of the case of stationary series.

(b) When the error distribution behavior is **Discrete Uniform**, we note from the results presented in Table (2) that,

1. Excellent robustness for the FPE criterion in small sample sizes for stationary series, and it begins to diminish with decreasing sample sizes.
2. The quality increases at non-stationary series, and everyone is equal in performance at small sizes of series.
3. This is also valid when the series undergoes a random walk.

(c) When the error distribution behavior is **Cauchy**, we note from the results presented in Table (3) that,

1. According to the two criteria, the mean square error (MSE) and the correct selection ratio (TSR) to estimate the order of the model, we notice a distinct and clear performance in small sample sizes in the stationary series case.
2. The performance of the criterion, and as a result, excellent robustness in the case of non-stationary series and for all sample sizes.
3. In the case of the random walk, when the default values of γ is negative, we notice distinct performance as the sample size decreases, but in the case of positive default values of γ , the performance increases with increasing sample size.

(d) When the error distribution behavior is **t**, we note from the results presented in Table (4) that,

1. the FPE criterion may possess high robustness for different types of series, when small sample sizes, and that robustness weakens as the sample size increases in the stationary series and to a lesser extent when the default value of γ is positive. The strength robustness decreases with increasing sample size and increasing non-stability.
2. There is constancy in the correct selection ratio (TSR) and the mean squared error (MSE) in estimating the model order when the series undergoes a random walk.

(e) When the error distribution behavior is **Log normal**, we note from the results presented in Table (5) that,

1. There is fixity in the correct selection ratio (TSR) and mean squared error criteria to estimate the order of the autoregressive model, and its quality increases in large samples. We also notice that there is a slight decrease in the percentage of the correct selection and a slight increase in the mean squared error as the default value of γ moves away from zero.
2. The FPE criterion is robust in non-stationary series in general because of the decrease in the mean squared error and the increase in the percentage of the correct selection.

3. In the case of the random walk series, we note that the performance of the FPE criterion is robust according to the MSE and TSR criteria for all sample sizes.
4. In general, it can be said that the robustness of FPE criterion is better for the cases mentioned in (1), (2) and (3) according to the (MSE) and (TSR) criteria in the case of positive default value of γ values than in the case of negative default value of γ .

SUMMARY

In this paper we present an evaluation of Karimi criterion to determine the order of autoregressive process. Karimi criterion can be seen as a modification of the famous Akaike criterion. To evaluate the performance of Karimi criterion, A simulation experiment was conducted in different cases for a Markov series model: stationary, random walk and nonstationary. Different sample sizes and different distributions of errors variable is used with run size 500 for each one trail. The distributions of errors were Gaussian, Discrete Uniform, Cauchy, t and Log normal. Two criteria were used to make the evaluation: true selection ratio and mean square error. Several conclusions were obtained in this paper. We advise researchers to study this criterion in other multivariate and univariate statistical models.

Table (1): The empirical values of the true selection ratio (TSR) and the empirical values of the mean squares error (MSE) to estimate the model order by using the Karimi criterion at different sample sizes T and different values of the Markov model parameter γ , when the errors series distributed as standard Normal.

γ	T	10	25	50	100	250
-0.9	TSR	0.661	0.681	0.663	0.609	1
	MSE	0.435	0.451	0.547	0.715	0.033
-0.7	TSR	0.635	0.587	0.601	0.539	1
	MSE	0.647	0.743	0.789	0.959	0.033
-0.5	TSR	0.517	0.581	0.533	0.555	0.535
	MSE	1.041	0.953	1.169	1.015	1.215
-0.4	TSR	0.559	0.585	0.565	0.555	0.515
	MSE	1.131	1.075	1.125	1.099	1.307
-0.3	TSR	0.555	0.509	0.531	0.565	1
	MSE	1.189	1.265	1.219	1.251	0.033
-0.1	TSR	0.571	0.529	0.525	0.511	0.563
	MSE	1.257	1.389	1.237	1.263	1.091
0.1	TSR	0.455	0.563	0.479	0.519	0.515
	MSE	1.443	1.217	1.469	1.285	1.259
0.3	TSR	0.513	0.525	0.523	0.545	0.581
	MSE	1.357	1.399	1.275	1.139	1.109
0.5	TSR	0.519	0.551	0.539	0.547	0.617
	MSE	1.267	1.265	1.229	1.179	1.061
0.7	TSR	0.565	0.557	0.535	0.579	0.665
	MSE	1.131	1.115	1.142	1.063	0.785
0.8	TSR	0.555	0.545	0.601	0.683	0.683
	MSE	1.131	1.223	0.939	0.785	0.773
0.9	TSR	0.535	0.595	0.605	0.711	0.735
	MSE	1.101	1.035	1.055	0.751	0.673
-1	TSR	0.715	0.729	0.721	0.667	0.649
	MSE	0.351	0.367	0.441	0.669	0.711
1	TSR	0.553	0.611	0.765	0.765	0.807
	MSE	1.431	1.265	0.679	0.541	0.421
-1.6	TSR	1	1	1	1	1
	MSE	0.033	0.033	0.033	0.033	0.037
-1.1	TSR	1	1	0.735	0.735	0.697
	MSE	0.033	0.121	1.009	0.793	0.723
1.1	TSR	1	1	0.905	0.893	0.873
	MSE	0.033	0.033	0.389	0.323	0.391

Table (2): The empirical values of the correct selection ratio (TSR) and the empirical values of the mean squares error (MSE) to estimate the model order by using the FPE criterion at different sample sizes T and different values of the Markov model parameter γ , when the series residuals is distributed as Discrete Uniform.

γ	T	10	25	50	100	250
-0.9	TSR	0.603	0.709	0.711	0.777	0.813
	MSE	0.393	0.239	0.201	0.069	0.027
-0.7	TSR	0.573	0.617	0.647	0.743	0.777
	MSE	0.645	0.463	0.415	0.205	0.087
-0.5	TSR	0.545	0.617	0.685	0.755	0.807
	MSE	0.769	0.589	0.418	0.205	0.069
-0.4	TSR	0.563	0.581	0.641	0.707	0.787
	MSE	0.727	0.601	0.451	0.307	0.167
-0.3	TSR	0.559	0.615	0.633	0.773	0.881
	MSE	0.815	0.609	0.513	0.175	0.145
-0.1	TSR	0.489	0.575	0.621	0.727	0.787
	MSE	0.921	0.715	0.507	0.335	0.197
0.1	TSR	0.537	0.573	0.631	0.731	0.803
	MSE	0.897	0.765	0.527	0.313	0.109
0.3	TSR	0.573	0.617	0.655	0.741	0.789
	MSE	0.771	0.595	0.527	0.285	0.147
0.5	TSR	0.609	0.621	0.681	0.757	0.817
	MSE	0.693	0.591	0.483	0.215	0.095
0.7	TSR	0.659	0.657	0.629	0.737	0.755
	MSE	0.493	0.495	0.595	0.253	0.217
0.8	TSR	0.669	0.651	0.667	0.759	0.741
	MSE	0.375	0.483	0.467	0.261	0.225
0.9	TSR	0.681	0.663	0.679	0.757	0.769
	MSE	0.339	0.453	0.431	0.209	0.155
-1	TSR	0.593	0.679	0.717	0.807	0.835
	MSE	0.469	0.299	0.135	0.02	0.06
1	TSR	0.751	0.725	0.723	0.749	0.785
	MSE	0.185	0.253	0.333	0.385	0.247
-1.6	TSR	0.885	0.895	0.899	0.857	0.899
	MSE	0.06	0.099	0.03	0.043	0.02
-1.1	TSR	0.643	0.813	0.707	0.869	0.899
	MSE	0.424	0.075	0.553	0.019	0.021
1.1	TSR	0.791	0.775	0.819	0.897	0.899
	MSE	0.103	0.143	0.111	0.097	0.022

Table (3): The empirical values of the correct selection ratio (TSR) and the empirical values of the mean squares error (MSE) to estimate the model order by using the FPE criterion at different sample sizes T and different values of the Markov model parameter γ , when the series residuals is distributed as Cauchy.

γ	T	10	25	50	100	250
-0.9	TSR	0.437	0.519	0.673	0.775	0.845
	MSE	1.497	1.331	0.949	0.762	0.669
-0.7	TSR	0.391	0.407	0.593	0.751	0.777
	MSE	1.975	1.755	1.323	0.973	0.833
-0.5	TSR	0.594	0.491	0.679	0.691	0.791
	MSE	1.607	1.557	1.189	1.129	0.801
-0.4	TSR	0.487	0.579	0.685	0.841	0.773
	MSE	1.633	1.493	1.111	0.649	0.933
-0.3	TSR	0.541	0.595	0.647	0.751	0.811
	MSE	1.513	1.447	1.215	0.973	0.769
-0.1	TSR	0.549	0.601	0.685	0.737	0.829
	MSE	1.523	1.357	1.195	1.001	0.727
0.1	TSR	0.449	0.625	0.743	0.781	0.869
	MSE	1.605	1.369	0.921	0.829	0.585
0.3	TSR	1	0.597	0.719	0.797	0.891
	MSE	1.631	1.511	1.053	0.813	0.521
0.5	TSR	0.619	0.585	0.755	0.811	0.909
	MSE	1.489	1.595	1.017	0.715	0.491
0.7	TSR	0.619	0.645	0.769	0.841	0.967
	MSE	1.489	1.571	1.015	0.715	0.319
0.8	TSR	0.605	0.637	0.797	0.895	0.995
	MSE	1.869	1.705	1.705	0.571	0.255
0.9	TSR	0.631	0.645	0.915	1	1
	MSE	1.903	1.835	0.785	0.251	0.121
-1	TSR	0.635	0.703	0.791	0.839	0.897
	MSE	0.585	0.589	0.603	0.597	0.635
1	TSR	1	1	1	1	1
	MSE	0.245	0.125	0	0	0
-1.6	TSR	1	1	1	1	1
	MSE	0	0.109	0.101	0.109	0.129
-1.1	TSR	1	1	0.757	0.855	0.899
	MSE	0.101	0.159	1.117	0.677	0.579
1.1	TSR	1	1	1	1	1
	MSE	0	0	0	0	0

Table (4): The empirical values of the correct selection ratio (TSR) and the empirical values of the mean squares error (MSE) to estimate the model order by using the FPE criterion at different sample sizes T and different values of the Markov model parameter γ , when the series residuals is distributed as t.

γ	T	10	25	50	100	250
-0.9	TSR	0.732	0.71	0.764	0.826	0.83
	MSE	0.592	0.386	0.278	0.162	0.182
-0.7	TSR	0.632	0.662	0.758	0.778	0.824
	MSE	0.626	0.584	0.362	0.33	0.206
-0.5	TSR	0.592	0.642	0.78	0.852	0.864
	MSE	0.9	0.754	0.388	0.232	0.226
-0.4	TSR	0.582	0.69	0.798	0.826	0.866
	MSE	0.958	0.724	0.364	0.312	0.212
-0.3	TSR	0.582	0.668	0.756	0.734	0.798
	MSE	1.054	0.752	0.55	0.5	0.376
-0.1	TSR	0.6	0.632	0.72	0.782	0.77
	MSE	1.006	0.836	0.592	0.434	0.416
0.1	TSR	0.566	0.63	0.704	0.74	0.756
	MSE	1.04	0.874	0.596	0.584	0.49
0.3	TSR	0.616	0.63	0.7	0.738	0.776
	MSE	0.942	0.826	0.618	0.544	0.434
0.5	TSR	0.576	0.638	0.7	0.72	0.806
	MSE	0.988	0.86	0.72	0.586	0.467
0.7	TSR	0.7	0.736	0.706	0.748	0.744
	MSE	0.63	0.54	0.714	0.51	0.562
0.8	TSR	0.756	0.746	0.78	0.836	0.864
	MSE	0.454	0.56	0.466	0.374	0.262
0.9	TSR	0.81	0.756	0.714	0.774	0.768
	MSE	0.358	0.466	0.556	0.484	0.448
-1	TSR	0.75	0.746	0.818	0.816	0.88
	MSE	0.538	0.416	0.188	0.178	0.084
1	TSR	0.822	0.77	0.776	0.82	0.838
	MSE	0.286	0.482	0.602	0.54	0.438
-1.6	TSR	0.97	0.978	0.982	0.93	0.982
	MSE	0.024	0.014	0	0.154	0
-1.1	TSR	0.758	0.742	0.738	1	0.982
	MSE	0.588	0.558	0.934	0.03	0
1.1	TSR	0.842	0.896	0.954	0.978	0.982
	MSE	0.254	0.146	0.07	0.015	0

Table (5): The empirical values of the correct selection ratio (TSR) and the empirical values of the mean squares error (MSE) to estimate the model order by using the FPE criterion at different sample sizes T and different values of the Markov model parameter γ , when the series residuals is distributed as Lognormal.

γ	T	10	25	50	100	250
-0.9	TSR	0.542	0.58	0.65	0.708	0.738
	MSE	0.576	0.502	0.426	0.5	0.464
-0.7	TSR	0.36	0.454	0.548	0.684	0.864
	MSE	1.334	1.114	0.96	0.71	0.128
-0.5	TSR	0.418	0.478	0.57	0.654	0.67
	MSE	1.342	1.108	0.938	0.77	0.718
-0.4	TSR	0.45	0.51	0.562	0.654	0.67
	MSE	1.214	1.082	0.928	0.722	0.694
-0.3	TSR	0.436	0.528	0.59	0.64	0.678
	MSE	1.258	0.974	0.948	0.76	0.704
-0.1	TSR	0.46	0.558	0.586	0.628	0.682
	MSE	1.258	1.022	0.91	0.79	0.658
0.1	TSR	0.472	0.56	0.622	0.66	0.844
	MSE	1.342	1.026	0.754	0.626	0.538
0.3	TSR	0.464	0.518	0.562	0.684	0.741
	MSE	1.314	1.164	0.988	0.626	0.356
0.5	TSR	0.496	0.532	0.598	0.716	0.83
	MSE	1.337	1.204	0.898	0.54	0.216
0.7	TSR	0.49	0.572	0.618	0.756	0.862
	MSE	1.594	1.29	0.98	0.476	0.152
0.8	TSR	0.47	0.546	0.636	0.822	0.902
	MSE	1.8	1.496	1.088	0.314	0.065
0.9	TSR	0.53	0.592	0.666	0.824	0.92
	MSE	1.686	1.42	1.112	0.36	0.042
-1	TSR	0.624	0.686	0.718	0.718	0.772
	MSE	0.332	0.276	0.304	0.496	0.412
1	TSR	0.842	0.848	0.872	0.932	0.962
	MSE	0.492	0.468	0.372	0.102	0.02
-1.6	TSR	0.972	0.894	0.972	0.972	0.962
	MSE	0.021	0.242	0.024	0.02	0
-1.1	TSR	0.962	0.952	0.654	0.748	0.774
	MSE	0.012	0.04	1.214	0.484	0.476
1.1	TSR	0.972	0.972	0.972	0.972	0.972
	MSE	0.021	0.021	0.021	0.021	0.021

REFERENCES

1. H. Akaike, “*Fitting Autoregressive Models for Prediction*”. Annals of the Institute of Statistical Mathematics, 1969, 21, 243-247.
2. H. Akaike, “*Statistical predictor identification*,” Ann. Ins.Statist. Math., 1970a, vol. 22, pp. 203-217.
3. H. Akaike, “*A fundamental relation between predictor identification and power spectrum estimation*. Ann. Inst. Statist. Math., 1970b, Vol. 22, pp. 219-223.
4. H. Akaike, “*A new look at the statistical model identification*,” IEEE Trans. Autom. Contr., 1974, vol. 19, no. 6, pp. 716–723.
5. R. Bhansali and D. Downham, “Some properties of the order of an autoregressive model selected by a generalization of Akaike’s EPF criterion”, *Biometrika*, 1977, Volume 64, Issue 3, pp. 547–551.
6. G. Golub and C. Van Loan “*Matrix Computations*”The Johns Hopkins Press, London, 1996.
7. J. Hamilton “*Time Series Analysis*”, Princeton University Press, Princeton New Jersey, 1994.
8. M. Karimi, “*A corrected FPE criterion for autoregressive processes*” 15th European Signal Processing Conference (EUSIPCO), Poznan, Poland, , 3-7 Sept. 2007, pp. 803-806.
9. J. McClave “*subset autoregression*” Technometrics, 1975, [Vol. 17, No. 2](#), pp. 213-220.