International Review of Business Research Papers
Volume 6. Number 4. September 2010. Pp. 501 – 511

Model Selection Criteria in Residual sum of Squares form

M. Shafiqur Rahman*

In this article eight existing model selection criteria are expressed in residual sum of squares form with a multiplicative penalty function. Practically these new forms have greater speed and greater scope compared to their original forms. Expressing all existing criteria in one common form makes it very easy to compare their performances among themselves. Marginal penalty functions are derived for all criteria and numerical values are computed for some arbitrary values of sample size and number of parameters. It is observed that in general $\bar{R}^2$ (RBAR) criterion favours the higher parametric model and Bayesian information criterion (BIC) favours the lower parametric model.

Field of Research: Econometrics, Applied Economics

1. Introduction

Model selection criteria are used in applied Economics, Econometrics and Statistics. At present a large inventory of model selection criteria are available in the literature including *AIC* (Akaike,1973), *BIC* (Schwartz,1978), $\bar{R}^2$ (*RBAR*) (Theil,1961), generalized cross validation (*GCV*) criterion (Craven and Wahba,1979), Hannan and Quinn's criterion (*HQC*) (Hannan and Quinn,1979), S_p criterion (Hocking,1976), joint information criterion (JIC) (Rahman and King, 1999), and C_p criterion (Mallows,1964) for choosing an appropriate model from a number of competing alternative models for a particular data set. The performance of any criterion varies from situation to situation. None of them is better in all situations. So there is a need to compare all available criteria with each other to investigate which one performing better in which situation. Fox expressed many model selection criteria as penalized log likelihood functions (Fox, 1995), ranked them in terms of the penalties paid for the addition of an extra parameter and interpreted them as χ^2-statistics. In this article a general form of a new model selection criterion with multiplicative penalty function is proposed. Eight existing model selection criteria were expressed in this new form with a multiplicative penalty function. Marginal penalty functions are derived for each criterion and applied to compare the performances of all criteria among themselves.

The plan of this paper is as follows. General form of a new model selection criterion is proposed in section 2. Section 3 reports the residual sum of squares (*RSS*) form with multiplicative penalty function of eight existing model selection criteria and their corresponding marginal penalty function. Section 4

*M. Shafiqur Rahman, Department of Operations Management and Business Statistics, College of Commerce and Economics, Sultan Qaboos University, Muscat, Sultanate of Oman.

deals with the comparison of the performances of existing model selection criteria with the help of penalty and marginal penalty functions. The final section contains some concluding remarks.

2. Information Criterion with multiplicative penalty

Suppose we are interested in selecting a model from m alternative regression models $M_1, M_2, \ldots, M_m$ for a given data set. Let the model M_j ($j = 1,2, \ldots, m$) be represented by

$$Y = X_j\beta_j + U_j \qquad (2.1)$$

where Y is an $n \times 1$ vector of observations on the dependent variable, X_j is an $n \times (k_j\text{-}1)$ matrix of observations on regressors, β_j is a $(k_j\text{-}1) \times 1$ vector of regression coefficients and U_j is a vector of random disturbances following N(0, σ_j^2 I). Then the log-likelihood function for the model M_j is given by

$$L_j(\beta_j, \sigma_j^2) = -\frac{n}{2}\left[\ln\sigma_j^2 + \ln(2\pi) + \frac{1}{n\sigma_j^2}\left(Y - X_j\beta_j\right)'\left(Y - X_j\beta_j\right)\right]. \qquad (2.2)$$

The log likelihood can be regarded as an estimator of the expected log likelihood. The mean expected log likelihood is the mean with respect to the data, of the expected likelihood of the maximum likelihood model and is a measure for the goodness of fit of a model. The model with the largest mean expected log likelihood can be considered to be the best model. The mean expected log likelihood can be estimated by the maximum log likelihood which for the jth model is given by

$$L_j(\hat{\beta}_j, \hat{\sigma}_j^2) = -\tfrac{n}{2}\left[\ln\hat{\sigma}_j^2 + \ln(2\pi) + 1\right] \qquad (2.3)$$

where, $\hat{\sigma}_j^2 = \frac{E_j^2}{n}$ is the maximum likelihood estimator (MLE) of σ_j^2,

$E_j^2 = (Y - X_j\hat{\beta}_j)'(Y - X_j\hat{\beta}_j)$ is the residual sum of squares and

$\hat{\beta}_j = (X_j'X_j)^{-1}X_j'Y.$

Unfortunately, the maximum log likelihood has a general tendency to over estimate the true value of the mean expected log likelihood. This tendency is more prominent for models with a large number of parameters and implies that if we choose the model with the largest maximum log likelihood, a model with an unnecessarily large number of parameters is likely to be chosen.

It is evident from (2.3) that choosing the model with the largest maximum log likelihood is equivalent to choose the model with the smallest residual sum of squares (E_j^2). Therefore, the model with the smallest E_j^2 can be considered to be the best model. Also if we choose the model with the smallest E_j^2, a model with an unnecessarily large number of parameters is likely to be chosen. To overcome this problem we need some adjustment in the E_j^2 before using it for model selection. This is done by using a penalty function dependent on the number of parameters, among other things. Let P_j be the penalty function for the model M_j. Then we usually select the model with the smallest T_j given by

$$T_j = E_j^2 P_j. \qquad (2.4)$$

This new model selection criterion will be called Criterion with multiplicative penalty (*CMP*). All existing model selection criteria can be expressed in the *CMP* form.

3. Existing criteria in *RSS/CMP* form

Let us consider the problem of selecting a model from m alternative regression models $M_1, M_2, \ldots, M_m$ for a given data set and the j^{th} model is given in (2.1). For convenience of presentation we denote total sum of squares of the j^{th} model by S_j^2. Therefore for the j^{th} model,

$$S_j^2 = \left(Y - \bar{Y}\right)'\left(Y - \bar{Y}\right), \qquad (3.1)$$

where $\bar{Y} = \sum_{i=1}^{n} \frac{Y_i}{n}$.

The residual sum of squares for the j^{th} model is given by

$$E_j^2 = (Y - X_j\hat{\beta}_j)'(Y - X_j\hat{\beta}_j), \qquad (3.2)$$

where $\hat{\beta}_j = (X_j'X_j)^{-1}X_j'Y$.

The coefficient of determination usually denoted by R^2 for the j^{th} model is defined by

$$R^2 = 1 - \frac{E_j^2}{S_j^2}. \qquad (3.3)$$

The proposed model selection criterion is given by

$$CMP = E_j^2 P_j \qquad (3.4)$$

It suggests that

if $E_j^2 P_j < E_i^2 P_i, \quad \forall i = 1,2,..,(j-1),(j+1),..,m$ (3.5)

then the model M_j will be our choice of the best model.

At present a large inventory of model selection criteria are available. Eight of them are listed below and expressed them in *CMP* form.

(a) Theil's adjusted $\bar{R}^2$ criterion

Theil suggested the adjusted $\bar{R}^2$ criterion (Theil, 1961) for model comparison and is given by

$$\bar{R}^2 = 1 - \frac{nE_j^2}{\left(n - k_j\right)S_j^2}. \qquad (3.6)$$

This criterion will select the model with the largest $\bar{R}^2$. The probability of correct selection under this criterion when the model M_j is true can be written as

$$P(CS|\, M_j, \bar{R}^2) = P\left\{\frac{nE_j^2}{\left(n-k_j\right)} < \frac{nE_i^2}{\left(n-k_i\right)}; \forall i = 1,2,...,(j-1),(j+1),..,m\right\}. \qquad (3.7)$$

The equation (3.7) implies that for model selection the $\bar{R}^2$ criterion can be written in the following *CMP* form

$$\bar{R}^2 \approx \frac{nE_j^2}{(n-k_j)} \tag{3.8}$$

Therefore, (3.8) is a particular case of (3.4) with $P_j = \frac{n}{(n-k_j)}$. That is $\bar{R}^2$ criterion is a special case of the *CMP*. Also using $\bar{R}^2$ criterion we can select the model with the smallest error variance. Theil showed that this criterion will select the true model at least as often as any other model. Later on Schmidt showed that $\bar{R}^2$ criterion will not help us in selecting the true model from a set of competing alternative regression models if a regression model contains both the variables of the true model with some extra, irrelevant regressors (Schmidt, 1973, 1975). The marginal penalty for this criterion is $\frac{\partial P_j}{\partial k_j} = \frac{n}{(n-k_j)^2}$ = MP_{RBAR}.

(b) Mallows C_p criterion

Mallows proposed a criterion for model selection (Mallows, 1964) which can be written for the j^{th} model in the following form,

$$C_p = \frac{(n+k_j)E_j^2}{(n-k_j)}. \tag{3.9}$$

Also (Rothman, 1968), (Akaike, 1969) and (Amemiya, 1980) have suggested this criterion. Rothman calling it J_p, Akaike calling it Final Prediction Error (*FPE*) and Amemiya calling it Prediction Criterion (*PC*).

Therefore the C_p criterion is a special case of *CMP* with $P_j = \frac{(n+k_j)}{(n-k_j)}$ and the marginal penalty for C_p criterion is $\frac{\partial P_j}{\partial k_j} = \frac{2n}{(n-k_j)^2}$ = MP_{CP}.

(c) Hocking S_p criterion

Hocking suggested S_p criterion (Hocking, 1976) for model selection and is given by

$$S_p = \frac{n^2 E_j^2}{(n-k_j)(n-k_j-1)} \tag{3.10}$$

Therefore the S_p criterion is a special case of the proposed criterion with $P_j = \frac{n^2}{(n-k_j)(n-k_j-1)}$ and the marginal penalty for this criterion is

$\frac{\partial P_j}{\partial k_j} = \frac{n^2(2n-2k_j-1)}{(n-k_j)^2(n-k_j-1)^2}$ = MP_{SP}.

(d) Generalised Cross Validation criterion

Craven and Wahba proposed Generalised Cross Validation (*GCV*) criterion (Craven and Wahba, 1979) and can be written as

$$GCV = \frac{E_j^2}{\left(1 - \frac{k_j}{n}\right)^2} . \tag{3.11}$$

Therefore the *GCV* criterion is a special case of *CMP* with $P_j = \frac{1}{\left(1 - \frac{k_j}{n}\right)^2}$ and the marginal penalty for *GCV* criterion is $\frac{\partial P_j}{\partial k_j} = \frac{2}{n\left(1 - k_j / n\right)^3} = MP_{GCV}$.

(e) Hannan and Quinn criterion (*HQC*)

Hannan and Quinn suggested a model selection criterion (Hannan and Quinn, 1979) and can be expressed as

$$HQC = E_j^2 (\ln n)^{\frac{2k_j}{n}} . \tag{3.12}$$

Therefore the *HQC* is a special case of the proposed criterion with $P_j = (\ln n)^{\frac{2k_j}{n}}$ and the marginal penalty for *HQC* criterion is $\frac{\partial P_j}{\partial k_j} = \frac{2(\ln n)^{2k_j / n} \ln(\ln n)}{n} = MP_{HQC}$.

.

(f) Akaike information criterion (*AIC*)

Akaike proposed a model selection criterion (Akaike, 1973) usually denoted by *AIC* and can be expressed as

$$AIC = L_j(\hat{\beta}_j, \hat{\sigma}_j^2) - k_j \tag{3.13}$$

The probability of correct selection under *AIC* when the model M_j is true can be written as

$$\begin{aligned} P(CS| M_j, AIC) &= P\left\{L_j\left(\hat{\beta}_j, \hat{\sigma}_j^2\right) - k_j > L_i\left(\hat{\beta}_i, \hat{\sigma}_i^2\right) - k_i ; \forall i = 1,2,.., (j-1), (j+1),.., m\right\} \\ &= P\left\{-\tfrac{n}{2}\left[\ln \hat{\sigma}_j^2 + \ln(2\pi) + 1\right] - k_j > -\tfrac{n}{2}\left[\ln \hat{\sigma}_i^2 + \ln(2\pi) + 1\right] - k_i ; \forall i = 1,2,.., (j-1), (j+1),.., m\right\} \\ &= P\left\{\tfrac{n}{2} \ln \hat{\sigma}_j^2 + k_j < \tfrac{n}{2} \ln \hat{\sigma}_i^2 + k_i ; \forall i = 1,2,.., (j-1), (j+1),.., m\right\} \end{aligned}$$

$$P(CS| M_j, AIC) = P\left\{E_j^2 e^{\frac{2k_j}{n}} < E_i^2 e^{\frac{2k_i}{n}} ; \forall i = 1,2,.., (j-1), (j+1),.., m\right\}. \tag{3.14}$$

Therefore the *AIC* criterion can be written in the following equivalent form

$$AIC \approx E_j^2 e^{\frac{2k_j}{n}} . \tag{3.15}$$

Therefore *AIC* is a special case of *CMP* with $P_j = e^{\frac{2k_j}{n}}$ and the marginal penalty for *AIC* criterion is $\frac{\partial P_j}{\partial k_j} = \frac{2e^{\frac{2k_j}{n}}}{n}$ = MP_{AIC}.

(g) Schwarz's Bayesian Information Criterion (*BIC*)

Schwarz proposed the Bayes Information Criterion (Schwarz, 1978) usually denoted by *BIC* and can be expressed as

$$BIC = L_j(\hat{\beta}_j, \hat{\sigma}_j^2) - \frac{k_j}{2}\ln n \tag{3.16}$$

The probability of correct selection under *BIC* when the model M_j is true can be written as

P(CS| M_j, *BIC*)

$$= \mathrm{P}\left\{L_j\left(\hat{\beta}_j, \hat{\sigma}_j^2\right) - \frac{k_j}{2}\ln n > L_i\left(\hat{\beta}_i, \hat{\sigma}_i^2\right) - \frac{k_i}{2}\ln n; \forall i = 1,2,..,(j-1),(j+1),..,m\right\}$$

$$= \mathrm{P}\left\{-\tfrac{n}{2}\left[\ln\hat{\sigma}_{\mathrm{j}}^2 + \ln(2\pi) + 1\right] - \frac{k_j}{2}\ln n > -\tfrac{n}{2}\left[\ln\hat{\sigma}_{\mathrm{i}}^2 + \ln(2\pi) + 1\right] - \frac{k_i}{2}\ln n; \forall i = 1,2,..,(j-1),(j+1),..,m\right\}$$

$$= \mathrm{P}\left\{-\tfrac{n}{2}\ln\hat{\sigma}_{\mathrm{j}}^2 - \frac{k_j}{2}\ln n < -\tfrac{n}{2}\ln\hat{\sigma}_{\mathrm{i}}^2 - \frac{k_i}{2}\ln n; \forall i = 1,2,..,(j-1),(j+1),..,m\right\}$$

$$= \mathrm{P}\left\{-\tfrac{n}{2}\ln\left(\frac{E_j^2}{n}\right) - \frac{k_j}{2}\ln n < -\tfrac{n}{2}\ln\left(\frac{E_i^2}{n}\right) - \frac{k_i}{2}\ln n; \forall i = 1,2,..,(j-1),(j+1),..,m\right\}$$

$$\mathrm{P(CS}|\ M_j, BIC) = \mathrm{P}\left\{E_j^2 n^{\frac{k_j}{n}} < E_i^2 n^{\frac{k_i}{n}}; \forall i = 1,2,..,(j-1),(j+1),..,m\right\}. \tag{3.17}$$

Hence the *BIC* criterion can be written in the following equivalent form

$$BIC \approx E_j^2\ n^{\frac{k_j}{n}} \tag{3.18}$$

Therefore the *BIC* criterion is a special case of *CMP* with $P_j = n^{\frac{k_j}{n}}$ and the marginal penalty for *BIC* criterion is $\frac{\partial P_j}{\partial k_j} = (n)^{\frac{k_j}{n}-1}(\ln n)$ = MP_{BIC}.

(h) Joint Information Criterion (*JIC*)

Rahman and King proposed the Joint Information Criterion (Rahman and King, 1999) usually denoted by *JIC* and can be expressed as

$$JIC = L_j(\hat{\beta}_j, \hat{\sigma}_j^2) - \frac{1}{4}\left\{k_{\mathrm{j}}\ln n - n\ln\left(1 - \frac{k_{\mathrm{j}}}{n}\right)\right\}. \tag{3.19}$$

The probability of correct selection under *JIC* when the model M_j is true can be written as P(CS| M_j, *JIC*)

$$= \mathrm{P}\left\{ \begin{array}{l} L_j\left(\hat{\beta}_j, \hat{\sigma}_j^2\right) - \frac{1}{4}\left\{k_j \ln n - n \ln\left(1 - \frac{k_j}{n}\right)\right\} > L_i\left(\hat{\beta}_i, \hat{\sigma}_i^2\right) - \frac{1}{4}\left\{k_i \ln n - n \ln\left(1 - \frac{k_i}{n}\right)\right\}; \\ \forall i = 1,2,.., (j-1), (j+1),.., m \end{array} \right\}$$

$$= \mathrm{P}\left\{ E_j^2 \left(n - k_j\right)^{-\frac{1}{2}} (n)^{\frac{k_j}{2n}} < E_i^2 \left(n - k_i\right)^{-\frac{1}{2}} (n)^{\frac{k_i}{2n}} ; \forall i = 1,2,.., (j-1), (j+1),.., m \right\}. \qquad (3.20)$$

Hence the *JIC* criterion can be written in the following equivalent form

$$JIC \approx \frac{E_j^2 (n)^{\frac{k_j}{2n}}}{\sqrt{\left(n - k_j\right)}}, \qquad (3.21)$$

which is of the form (3.4) with $P_j = \frac{(n)^{\frac{k_j}{2n}}}{\sqrt{\left(n - k_j\right)}}$. Hence the *JIC* is a special case of the proposed criterion and the marginal penalty for *JIC* criterion is $\frac{\partial P_j}{\partial k_j} = \frac{(n)^{\frac{k_j}{2n}}}{\sqrt{\left(n - k_j\right)}} \left(\frac{\ln n}{2n} + \frac{1}{2\left(n - k_j\right)} \right) = MP_{JIC}$.

Therefore the *AIC*, *BIC*, *HQC*, $\overline{R}^2$, S_p, C_p, *JIC* and *GCV* criteria are special cases of the proposed criterion and can be easily obtained by suitable choices of P_j. All criteria are compared with the help of penalties and marginal penalties.

4. Comparison of the penalties and Marginal penalties

The penalty functions for eight existing model selection criteria are given in Table 4.1

TABLE 4.1: Penalty functions P_j for eight model selection criteria

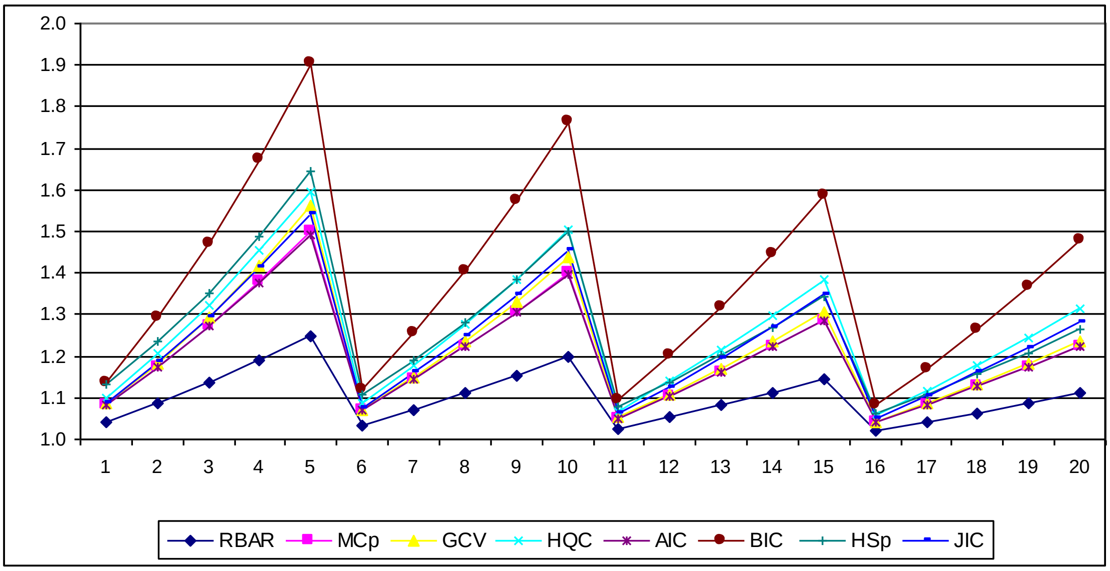
Criteria	P_j	Criteria	P_j
AIC	$e^{\frac{2k_j}{n}}$	*BIC*	$n^{\frac{k_j}{n}}$
GCV	$\frac{1}{\left(1 - \frac{k_j}{n}\right)^2}$	*HQC*	$(\ln n)^{\frac{2k_j}{n}}$
JIC	$\frac{(n)^{\frac{k_j}{2n}}}{\sqrt{\left(n - k_j\right)}}$	$\overline{R}^2$	$\frac{1}{\left(1 - \frac{k_j}{n}\right)}$
C_p	$\frac{(n + k_j)}{\left(n - k_j\right)}$	S_p	$\frac{n^2}{\left(n - k_j\right)\left(n - k_j - 1\right)}$

Values of P_j for some arbitrary values of n and k_j are calculated for each criteria and presented in Table 4.2. Graph 4.1 is the graphical representation of these values.

TABLE 4.2 : Values of P_j for arbitrary values of n and k_j for eight criteria

n	kj	RBAR	MCp	GCV	HQC	AIC	BIC	HSp	JIC
25	1	1.0417	1.0833	1.0851	1.0980	1.0833	1.1374	1.1322	1.0885
25	2	1.0870	1.1739	1.1815	1.2057	1.1735	1.2937	1.2352	1.1858
25	3	1.1364	1.2727	1.29132	1.3239	1.2712	1.4715	1.3528	1.2931
25	4	1.1905	1.3810	1.41723	1.4537	1.3771	1.6737	1.4881	1.4115
25	5	1.2500	1.5000	1.56250	1.5962	1.4918	1.9037	1.6447	1.5426
30	1	1.0345	1.0690	1.07015	1.0850	1.0689	1.1200	1.1084	1.0764
30	2	1.0714	1.1429	1.14796	1.1773	1.1426	1.2545	1.1905	1.1594
30	3	1.1111	1.2222	1.23457	1.2774	1.2214	1.4051	1.2821	1.2495
30	4	1.1538	1.3077	1.33136	1.3860	1.3056	1.5738	1.3846	1.3476
30	5	1.2000	1.4000	1.44000	1.5039	1.3956	1.7627	1.5000	1.4544
40	1	1.0256	1.0513	1.05194	1.0674	1.0513	1.0966	1.0796	1.0605
40	2	1.0526	1.1053	1.10803	1.1394	1.1052	1.2025	1.1380	1.1251
40	3	1.0811	1.1622	1.16874	1.2163	1.1618	1.3187	1.2012	1.1940
40	4	1.1111	1.2222	1.23457	1.2983	1.2214	1.4461	1.2698	1.2676
40	5	1.1429	1.2857	1.30612	1.3859	1.2840	1.5858	1.3445	1.3462
50	1	1.0204	1.0408	1.04123	1.0561	1.0408	1.0814	1.0629	1.0505
50	2	1.0417	1.0833	1.08507	1.1153	1.0833	1.1694	1.1082	1.1037
50	3	1.0638	1.1277	1.13173	1.1778	1.1275	1.2646	1.1563	1.1599
50	4	1.0870	1.1739	1.18147	1.2439	1.1735	1.3675	1.2077	1.2192
50	5	1.1111	1.2222	1.23457	1.3137	1.2214	1.4788	1.2626	1.2818

Graph 4.1: Penalties for some arbitrary values of n and k_j for eight criteria



From Table 4.2 and Graph 4.1 it is obvious that the penalties are in the following order

$$P_{BIC} > P_{HQC} > P_{GCV} > P_{MCP} > P_{AIC} > P_{RBAR}$$

Therefore, in general the penalty is higher for *BIC* and lower for *RBAR* criterion.

The marginal penalties for eight existing model selection criteria are given in Table 4.3

TABLE 4.3: Marginal penalty functions for eight model selection criteria

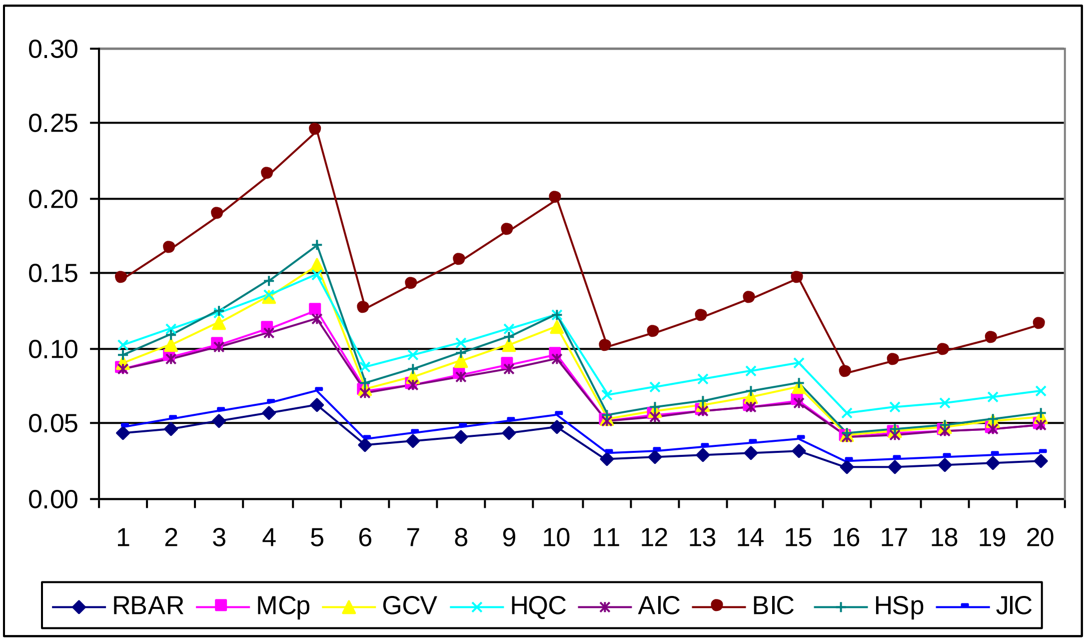
Criteria	*Marginal penalty*	*Criteria*	*Marginal penalty*
AIC	$\frac{2e^{\frac{2k_j}{n}}}{n}$	*BIC*	$(n)^{\frac{k_j}{n}-1}(\ln n)$
GCV	$\frac{2}{n\left(1-{k_j}/{n}\right)^3}$	*HQC*	$\frac{2(\ln n)^{2k_j/n}\ln(\ln n)}{n}$
$\bar{R}^2$	$\frac{n}{(n-k_j)^2}$	*JIC*	$\frac{(n)^{\frac{k_j}{2n}}}{\sqrt{(n-k_j)}}\left(\frac{\ln n}{2n}+\frac{1}{2(n-k_j)}\right)$
C_p	$\frac{2n}{(n-k_j)^2}$	S_p	$\frac{n^2(2n-2k_j-1)}{(n-k_j)^2(n-k_j-1)^2}$

Values of marginal penalties for some arbitrary values of n and k_j are calculated for each criteria and presented in Table 4.4. Graph 4.2 is the graphical representation of these marginal penalties for all criteria.

TABLE 4.4: Marginal penalties for arbitrary values of *n* and k_j for eight criteria

n	kj	RBAR	MCp	GCV	HQC	AIC	BIC	HSp	JIC
25	1	0.0434	0.0868	0.0904	0.1027	0.0867	0.1464	0.0964	0.0483
25	2	0.0473	0.0945	0.1027	0.1128	0.0939	0.1666	0.1098	0.0531
25	3	0.0517	0.1033	0.1174	0.1238	0.1017	0.1895	0.1259	0.0585
25	4	0.0567	0.1134	0.1350	0.1360	0.1102	0.2155	0.1453	0.0645
25	5	0.0625	0.1250	0.1563	0.1493	0.1193	0.2451	0.1688	0.0714
30	1	0.0357	0.0713	0.0738	0.0885	0.0713	0.1270	0.0778	0.0403
30	2	0.0383	0.0765	0.0820	0.0961	0.0762	0.1422	0.0866	0.0438
30	3	0.0412	0.0823	0.0914	0.1042	0.0814	0.1593	0.0968	0.0475
30	4	0.0444	0.0888	0.1024	0.1131	0.0870	0.1784	0.1086	0.0517
30	5	0.0480	0.0960	0.1152	0.1227	0.0930	0.1998	0.1225	0.0563
40	1	0.0263	0.0526	0.0539	0.0697	0.0526	0.1011	0.0561	0.0306
40	2	0.0277	0.0554	0.0583	0.0744	0.0553	0.1109	0.0607	0.0327
40	3	0.0292	0.0584	0.0632	0.0794	0.0581	0.1216	0.0658	0.0348
40	4	0.0309	0.0617	0.0686	0.0847	0.0611	0.1334	0.0716	0.0372
40	5	0.0327	0.0653	0.0746	0.0905	0.0642	0.1462	0.0780	0.0398
50	1	0.0208	0.0416	0.0425	0.0576	0.0416	0.0846	0.0438	0.0249
50	2	0.0217	0.0434	0.0452	0.0609	0.0433	0.0915	0.0467	0.0263
50	3	0.0226	0.0453	0.0482	0.0643	0.0451	0.0989	0.0497	0.0277
50	4	0.0236	0.0473	0.0514	0.0679	0.0469	0.1070	0.0531	0.0293
50	5	0.0247	0.0494	0.0549	0.0717	0.0489	0.1157	0.0568	0.0309

Graph 4.2: Marginal penalties for some arbitrary values of *n* and k_j



From Table 4.4 and Graph 4.2 it is obvious that the marginal penalties are in the following order

$$MP_{BIC} > MP_{HQC} > MP_{GCV} > MP_{CP} > MP_{AIC} > MP_{RBAR}$$

Therefore, in general the marginal penalty is also higher for *BIC* and lower for *RBAR* criterion.

5. Conclusion

The proposed criterion *CMP* is the general form of all existing criteria and is in residual sum of squares form. The model fitting output obtained by using any standard statistical package contains residual sum of squares. Therefore it is very easy to use and less time consuming for any model selection criteria in residual sum of squares form. Comparative study among model selection criteria also becomes very simple. It is observed that in general BIC criterion favours lower parametric model while *RBAR* favours the higher parametric models.

References

Akaike, H. 1969, "Fitting Autoregressive Models for Prediction', *Annals of the Institute of Statistical Mathematics"*, vol. 21, pp. 243-247.

Akaike, H. 1973, "Information theory and an extension of the maximum likelihood principle', Proceedings of the Second International Symposium on Information Theory, B.N. Petrov and F. Csaki, Akademial Kiado, Budapest, pp. 267-281.

Amemiya, T. 1980, "Selection of Regressors', *International Economic Review"*, Vol. 21, pp. 331-354.

Craven, P. and Wahba, G. 1979, "Smoothing Noisy data with Spline Functions: Estimating the Correct Degree of Smoothing by the Method of Generalized Cross Validation', *Numerische Mathematik"*, vol. 31, pp.377-403.

Fox, K.J. 1995, "Model Selection Criteria : A Reference Source", University of British Columbia and University of New South Wales, School of Economics, Sydney, Australia.

Hannan, E.J. and Quinn, B.G. 1979, "The determination of the order of an auto-regression', *Journal of the Royal Statistical Society", Series B, vol.*41, pp.190-195.

Hocking, R.R. 1976, "The analysis and selection of variables in linear regression', *Biometrics",* vol. 32, pp. 1-49.

King, M.L. 1981, "The Durbin-Watson bounds test and regressions without an intercept', *Australian Economic Papers"*, vol. 20, pp. 161-170.

Mallows, C.L. 1964, "Choosing Variables in a Linear Regression: A graphical Aid", presented at the Central Regional Meeting of the Institute of Mathematical Statistics, Manhattan, Kansas (May).

Mallows, C.L. 1973, "Some Comments on C_p', *Technometrics"*, vol. 15, pp. 661-676.

Rahman, M.S. and King, M.L.1999, "Improved Model Selection Criterion', *Communications in Statistics*, *Simulation and Computation*", vol. 28(1), pp. 51-71.

Rothman, D. 1968, "Letter to the Editor', *Technometrics"*, vol. 10, pp. 432.

Schmidt, P. 1973, "Calculating the Power of the Minimum Standard Error Choice Criterion', *International Economic Review"*, vol. 14, pp. 253-255.

Schmidt, P. 1975, "Choosing Among Alternative Linear Regression Models: A correction and some Further Results', *Atlantic Economic Journal"*, vol. 3, pp. 61-63.

Schwartz, G. 1978, "Estimating the dimension of a model', *The Annals of Statistics"*, vol. 6, pp. 461-464.

Theil, H. 1961, *Economic Forecasts and Policy*, 2nd edition, North-Holland, Amsterdam.