A criterion for adaptive autoregressive models

Alois Schlögl[1], S.J. Roberts[2], and Gert Pfurtscheller[1,3]

***Abstract* - A criterion, similar to the information criterion of a stationary AR model, is introduced for an adaptive (non-stationary) autoregressive model. It is applied to non-stationary EEG data. It is shown that the criterion can be used to determine the update coefficient, the model order, and the estimation algorithm.**

***Key words* – modeling, non-stationarity, EEG analysis**

I. INTRODUCTION

An adaptive autoregressive (AAR) model addresses the problem of non-stationary spectral analysis. Adaptively estimated autoregressive parameters are useful in many applications, examples are the on-line spectral analysis of heart rate variability [1], as well as EEG-based brain-computer interfacing [2].

Many different AAR estimation algorithms are available LMS, [3,4], RLS [5,6] and Kalman filtering [6,7], recursive AR algorithm [1,3]; Furthermore, information criteria like AIC [8], etc. are applicable only to stationary AR models.

II. METHOD

An AAR model with order p is written as

$$y(t) = a_1(t)*y(t-1) + ... + a_p(t)*y(t-p) + x(t) =$$

$$= \mathbf{a}(t)^T*\mathbf{Y}(t-1) + x(t) \quad i=1..p \qquad (1)$$

The difference with (stationary) autoregressive (AR) model being, that the AAR parameters vary with time. The one-step prediction error is

$$e(t) = y(t) - \hat{\mathbf{a}}(t-1)^T*\mathbf{Y}(t-1) \qquad (2)$$

The difference between the prediction error $e(t)$ and the innovation process $x(t)$ *is* that in the former one the estimated parameters rather than the "true" model parameters are used.

AAR parameters are estimated using a variety of adpative algorithms. Taking UC to be the update coefficient and k(t) the update gain the following algorithms have been proposed:

LMS 1 [3]:

$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + UC/MSY*e(t)*\mathbf{Y}(t-1)$$

LMS 2 ([4], modified):

$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + UC/\sigma^2_x(t)*e(t)*\mathbf{Y}(t-1)$$

RAR 1 ([1, 3], exponential forgetting function):

$$\mathbf{A}(t) = (1-UC)*\mathbf{A}(t)+UC*\mathbf{Y}(t)*\mathbf{Y}(t)^T$$
$$\mathbf{k}(t) = UC*\mathbf{A}(t)*\mathbf{Y}(t)/(UC*\mathbf{Y}(t)^T*\mathbf{A}(t)*\mathbf{Y}(t)+1)$$
$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + \mathbf{k}(t)^T*e(t)$$

RAR 2 ([1], whale forgetting function):

$$\mathbf{A}(t) = c1*\mathbf{A}(t-1)+c2*\mathbf{A}(t-2)+c3*\mathbf{Y}(t)*\mathbf{Y}(t)^T$$
$$(1-(1-2*UC)*z\hat{}-1)^2 = 1 + c1*(z\hat{}-1) + c2*(z\hat{}-2)$$
$$c3 = 1+c1+c2$$

Kalman Filtering (KF) [6, 7]:

$$\mathbf{a}(t) = \mathbf{a}(t-1) + \mathbf{w}(t)$$
$$\mathbf{w}(t) = N(0,\mathbf{W}(t))$$

(KF1): RLS [2,5,6];
(KF2): [7],
(KF3): $\mathbf{W}(t)=UC*trace(\mathbf{A}(t-1))/p$,
(KF4): $\mathbf{W}(t)=UC*\mathbf{I}$,
(KF5): $\mathbf{W}(t)=UC^2*\mathbf{I}$,

$$\mathbf{Q}(t) = \mathbf{Y}(t-1)^T * \mathbf{A}(t-1)*\mathbf{Y}(t-1) + V(t)$$
$$\mathbf{k}(t) = \mathbf{A}(t-1) * \mathbf{Y}(t-1) / \mathbf{Q}(t)$$
$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + \mathbf{k}(t)^T * e(t)$$
$$\mathbf{A}(t) = \mathbf{A}(t-1) - \mathbf{k}(t)*\mathbf{Y}(t-1)^T*\mathbf{A}(t-1) + \mathbf{W}(t)$$

RLS = KF1 (special form of KF) [2, 5]

$$\mathbf{V}(t) = 1/(1-UC)$$
$$\mathbf{W}(t) = UC*\mathbf{A}(t-1)$$

The various AAR estimation algorithms differ in how $k(t)$ is calculated.

The mean squared error (MSE) is used to measure how well the AAR estimates describe the observed process y. Normalizing the MSE by the variance of the signal (MSY), gives a relative error variance REV, being a criterion for the goodness-of-fit.

$$0 < REV = MSE/MSY \leq 1 \qquad (3)$$

Note, that $y(t)$ at time t is used firstly for calculating $e(t)$; especially $y(t)$ was not used for estimating any $\hat{a}_i(t)$. Therefore, $e(t)$ is uncorrelated to all previous samples $y(t-i)$ $i>1$. Thus, the REV criterion is an objective measure for the goodness-of-fit.

III. RESULTS

Fig.1 shows how the REV depends on UC for the various algorithms. In all cases, one single minimum value can be identified. The curves are different for the different algorithms. The lowest values across all algorithms were observed for Kalman filtering.

Figure 2 displays relative error variance depending on the model order and the update coefficient. Again, $REV(p,UC)$ is quite smooth, especially one global minmum can be observed.

At p=9 and $UC=2^{-8}$ a minimum with a value of REV_{min} = 0.0572 can be identified.

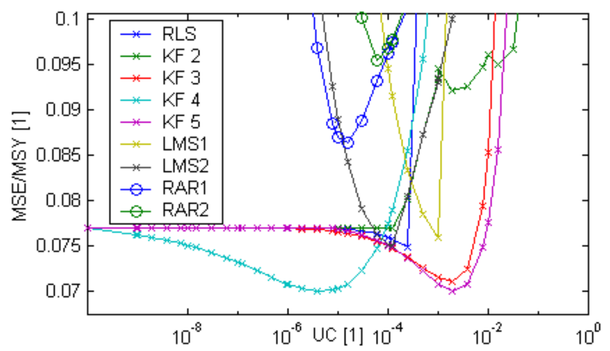

Figure 1: Comparison of AAR estimation algorithms. *REV* of different algorithms depending on the update coefficient UC. The model order was p=10. All algorithms were applied to a non-stationary EEG of 1000s length and sampled with 100Hz. A model order of p=10 was used. The update coefficient *UC* was varied between 10^{-k} k=1..10 and 2^{-k}, k=1..30. The lowest error rate is reached with Kalman filtering (KF2 and KF3 with *UC*=2e-3, KF4 and *UC*=4e-6). (See also [6])

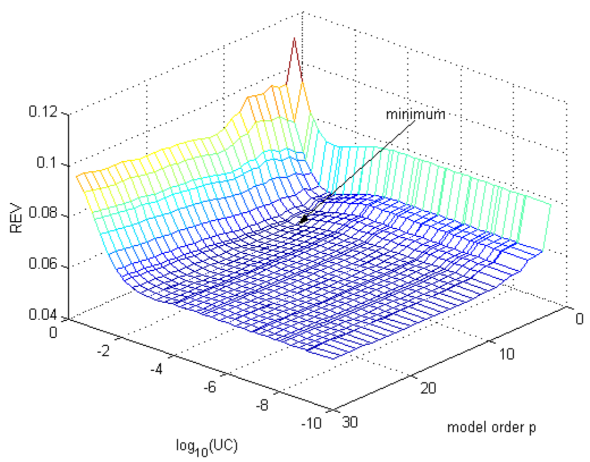

Figure 2: *REV(p,UC)* depending on model order p and update coefficient UC. The algorithm KF5 was applied to EEG with varying spectrum [2], sampled with 128Hz of a length of 407.5s, derived from electrode position C3 during repetitive imaginary left and right hand movement. The model order was varied from 2 to 30.

Generally, it can be observed that the optimal update coefficient decreases with increasing model order. This might have also some theoretical importance. Intuitively, it is clear that a larger number of parameters require a larger observation time, and consequently a smaller update coefficient. However, it can be also related to the principle of uncertainty between time and frequency domain. [9].

IV. SUMMARY

The smaller the REV is, the smaller is the residual. Generally, the aim is that most of the signal is explained by the model. Hence, it is the aim to reduce the REV as much as possible; the smallest REV provides the optimal setting.

It is suggested, that the REV is used as criterion for determining free parameters. The advantage of the REV criterion compared to other model order selection criterion (e.g. [8]) is that no penalty term is needed and it is suitable for non-stationary models.

ACKNOWLEDGMENT: The SIESTA project Biomed-2 BMH4-CT97-2040 was funded by the European Commission, DG XII and by the Austrian Ministry for Science and Technology (BMWF) GZ 650 133/2-VIII/6/97

REFERENCES

[1] Bianchi A., Mainardi L., Meloni C., Chierchia S., Cerutti S. Continuous monitoring of the Sympatho-Vagal Balance through spectral analysis. *IEEE Engineering in Medicine and Biology*. 16(5): 64-73, 1997.

[2] Pfurtscheller G, Neuper C, Schlögl A, Lugger K. Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters. *IEEE Trans Rehabil Eng*. 6(3):316-25, 1998.

[3] Akay M. *Biomedical Signal Processing*, Acad. Press, San Diego, 1994.

[4] B. Schack, H. Witte and G. Grießbach, "Parametrische Methoden der dynamischen Spektralanalyse und ihre Anwendung in der Biosignalanalyse." *Biomedizinische Technik*, 38, 79-80, 1993.

[5] L. Patomäki, J.P. Kaipio, and P.A. Karjalainen, "Tracking of nonstationary EEG with the roots of ARMA models." in *Proc. 17th Intl. Conf. IEEE EMBS*, 1995.

[6] Haykin S. *Adaptive Filter Theory*. Prentice Hall, Englewood Cliffs, NJ, 1996.

[7] Jazwinski A. H. Adaptive filtering, *Automatica,* 5(4): 475-485, 1969.

[8] Akaike, H. A new look at the statistical model identification, *IEEE Trans. Autom. Control*, vol. AC-19, pp.716-723, 1974.

[9] A. Schlögl and G. Pfurtscheller, Considerations on adaptive autoregressive modelling in EEG Analysis, *Proc. of First International Sysmposium on Communication Systems and Digital Signal Processing CSDSP'98*, Sheffield, UK, p. 367-370, 1998

[1] Institute of Biomedical Engineering,
University of Technology,
Graz, A-8010, Austria
a.schloegl@ieee.org

[2] Robotics Research Group in the Department of Engineering Science, University of Oxford
Oxford, UK
sjrob@robots.ox.ac.uk

[3] Ludwig Boltzmann Institute for Medical Informatics and Neuroinformatics
Graz, Inffeldgasse 16a
A-8010, Austria
pfu@dpmi.tu-graz.ac.at

A Criterion for Adaptive Autogressive Models

A Schlögl[1], S Roberts[2], G Pfurtscheller[1], (1)University of Technology, Graz, Austria, (2)Oxford University, Oxford, UK

Scientific Session: ***WE-A328-01*** Adaptive Filtering
Track: 23 Signal Processing

Short Paper Available

An adaptive autoregressive (AAR) model addresses the problem of non-stationary spectral analysis. Adaptively estimated autoregressive parameters are useful in many applications, examples are the on-line spectral analysis of heart rate variability (Binachi et al. 1997), as well as EEG-based brain computer interfacing (Pfurtscheller et al 1998). Many different estimation algorithms are available LMS (Akay 1994, Schack 1993) RLS (Patomaki et al. 1997, Haykin, 1996), and Kalman filtering (Haykin, 1996, Jazwinski, 1969), recursive AR algorithm (Akay, 1994, Bianchi et al. 1997); Furthermore, information criteria like AIC, etc. (Akaike, 1974) are applicable only to stationary AR model.

All AAR estimation algorithms calculate the one-step prediction error (i.e. residual part of the signal not explained by the AAR model parameters). The mean squared error (MSE) can be used to measure how well the AAR estimates describe the observed process *y*. Normalizing the MSE by the variance of the signal (MSY), gives a relative error variance REV, being a criterion for the goodness-of-fit.

$$0 < REV = MSE/MSY \ ?\ 1$$

The one-step prediction error $e(t)$ is uncorrelated to all previous samples $y(t-i)\ i>1$. Thus, the REV criterion is an objective measure for the goodness-of-fit. The REV can be used to determine the model order, the update coefficient or any other free parameters. The advantage of the REV criterion compared to other model order selection criteria is that no penalty term is needed and it is suitable for non-stationary models.