Time varying Least Squares with ARCH Errors for Environmental Time Series Modelling

Alessandro Fassò
University of Bergamo, Dept. of Engineering
Via Marconi 5
24044 *Dalmine BG, Italy*
fasso@unibg.it

1. Introduction

In environmental time series analysis, a number of dynamical stochastic models have been used so far. In particular, for modelling high frequency air pollution and meteorological data, e.g. daily or hourly data, both parametric and nonparametric models of increasing complexity have been proposed. Recently Fassò & Negri (2000) considered nonlinear modelling with long memory and heteroscedasticity for hourly ground ozone time series. For a more detailed discussion of recent literature see Fassò (2000).

In this paper, using the approach of Fassò (2000), we consider recursive estimation of smoothly changing autoregressive models with heteroscedastic innovations. Its performances and flexibility seem adequate both for monitoring air quality data from mobile air pollution stations and for managing missing data at the an early stage of building a more detailed and possibly time-invariant dynamical model for data coming from a static air pollution station.

2. RLS-ARCH Approach

In order to track and forecast a time-varying heteroscedastic model, we use the $ARCH$ extension of the standard recursive least square approach (RLS), based the following time-varying $ARX - ARCH$ equations $y_t = \theta_t' \varphi_{t-1} + \varepsilon_t h_{t-1}$ and $h_{t-1} = \beta_{t-1}' \eta_{t-1}$. Here, ε_t is an independent sequence with $E\varepsilon = 0$ and $E\left|\varepsilon\right| = 1$, φ_{t-1} is the regressor vector including stochastic covariates, deterministic components and lagged process observations. Moreover $\eta_{t-1} = (1, \left|e_{t-1}\right|, \ldots, \left|e_{t-r}\right|)'$, and $e_t = y_t - \theta_t' \varphi_{t-1}$. The quantities θ_t and β_t are unknown slowly time-varying coefficient vectors. In this model, the time-varying conditional mean function $\hat{y}_t = \theta_t \varphi_{t-1}$ has time-varying precision given by its conditional mean absolute error h_t, which is a function of previous forecast errors. When no heteroscedastic component is present or h_t is known, assuming Gaussian innovations and some stochastic linear dynamics for θ_t, it is possible to get the optimal estimate $\hat{\theta}_t$ using the Kalman filter, see e.g. Mosca (1997). In our case, adapting the empirical WLS approach introduced for time-invariant $ARCH$ models, we have closed form recursive solutions given by $\hat{\theta}_t = \hat{\theta}_{t-1} + R_t^{-1}\varphi_{t-1}e_t\omega_{t-1}$ and $\hat{\beta}_t = \hat{\beta}_{t-1} + S_t^{-1}\eta_{t-1}(|\tilde{e}_t| - \hat{\beta}'_{t-1}\eta_{t-1})$ where the matrices R_t and S_t are discussed in Fassò (2000). The dimension of θ_t is then controlled by static or dynamic versions of standard AIC or BIC statistics. With these values the one step ahead forecast is simply given by $\hat{y}_{t+1} = \hat{\theta}'_t\varphi_t$ and its precision can be evaluated by $\hat{h}_t = \hat{\beta}'_t\eta_t$.

3. Case study: the ground ozone hourly data set of Fassò & Negri (2000)

In order to illustrate the method, we use the data set summarized in Table 1. Although some estimation techniques for λ have been appeared in the literature, we have chosen $\lambda = 0.995$ by jointly looking at the one-step forecasting capability and at the periodic behavior of the data. The model adopted here contains four AR components and three covariates given by the lagged values of total solar radiation, relative humidity and nitrogen dioxide. From the related plot one can see a marked seasonal nonlinearity. Similarly the $ARCH$ components indicate a strong seasonal nonlinear heteroscedasticity. Plotting the maximum of absolute roots of the AR polynomial against t shows that the estimated model refers to a stable system (see e.g. Grillenzoni (1997)). A similar stable variance diagnostic was passed by the $ARCH$ component. The model is satisfactory from the fitting point of view as well as from the residual and squared residual autocorrelation point of view.

Table 1. Ozone and residuals statistics

	Mean	Variance	Skewness	Kurtosis
Ozone: y_t	19.064	510.16	2.482	12.049
residuals: e_t	0.367	37.774	0.896	13.647
studentized residuals: e_t/h_t	0.037	1.287	0.344	4.339

REFERENCES

Fassò A. (2000) Recursive least square with ARCH errors and nonparametric modelling of environmental time series. GRASPA w. paper n.6
(web version: http://sirio.stat.unipd.it/graspa/graspa.htm).
Fassò A., Negri I. (2000) Nonlinear statistical modelling of high frequency ground ozone data. In printing on *Environmetrics* (web version: http://dipinge.unibg.it/fasso/bibliografia.htm).
Grillenzoni C. (1997) Optimized Adaptive Prediction. *J. Italian Statist. Soc.*, 6, 1, 37-58.
Mosca E. (1997) *Optimal, Predictive and Adaptive Control.* Prentice Hall.

RESUME

Dans ce travail on étude des modèles dynamiques variables dans le temps hétéroschedastiques appliqué a des séries chronologiques sur l'environnement a haute fréquence.