Research Statement*

Alexander Schmitz†
Mathematical Institute
University of Cologne
50931 Cologne, Germany

Overview

My research interests focus on change-point analysis for time series data and limit theorems under dependence conditions.

This research statement is organized as follows: In Section 1 a brief synopsis of Schmitz and Steinebach (2010) is presented. Section 2 contains an account of my current research and results from my PhD thesis. Finally, my future research plans are outlined in Section 3.

1 Monitoring Change-Points

In testing time series data for structural stability two different approaches can be chosen. There are *retrospective* procedures which deal with the detection of structural breaks within an observed data set of fixed size, whereas *sequential* procedures check the stability hypothesis each time a new observation is available. Based on schemes proposed by Chu et al. (1996), Horváth et al. (2004) derived a **sequential testing procedure for detecting a change in the parameters of a linear regression model**, after a stable training period of size m (say). Their testing procedure is based on the first excess time of a detector over a boundary function, where the detector

*date: April 2011
†e-mail: schmitza@math.uni-koeln.de

is a cumulative sum (CUSUM) type statistic of the residuals. The boundary function can be suitably chosen such that the test attains a prescribed asymptotic size α (say) and asymptotic power one as m tends to infinity. Since they modeled the errors of the linear regression to be independent and identically distributed, *my motivation* is to show that the sequential monitoring procedure for the linear model

$$y_i = \mathbf{x}_i^T \boldsymbol{\beta}_i + \epsilon_i = x_{1,i}\beta_{1,i} + \cdots + x_{p,i}\beta_{p,i} + \epsilon_i, \quad 1 \le i < \infty,$$

which was discussed in Horváth et al. (2004), **continues to hold under dependence conditions on the error sequence**.

We assume the parameter vectors to be constant over a training period of length m, i.e.

$$\boldsymbol{\beta}_i = \boldsymbol{\beta}_0, \quad 1 \le i \le m.$$

This period is used as a reference for comparisons with future observations.

The monitoring procedure leads to a decision between the *"no change"* null hypothesis and the *"change at unknown time k^*"* alternative, i.e.

$$H_0: \ \boldsymbol{\beta}_{m+i} = \boldsymbol{\beta}_0 \quad \text{for all } \ i \ge 1 \quad \textit{versus}$$

$$H_A: \ \boldsymbol{\beta}_{m+k^*+i} = \boldsymbol{\beta}_* \ne \boldsymbol{\beta}_0 \quad \text{for some } \ 1 \le k^* < \infty \ \text{ and for all } \ i \ge 0.$$

The parameter k^* is called the change-point which is assumed to be unknown as well as the values of the parameters $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_*$.

The monitoring procedure is defined via a stopping rule based on the first excess time τ_m of a change detector $\hat{Q}_m(\cdot)$ over a boundary function $g_m(\cdot)$, i.e.

$$\tau_m = \inf\left\{k \ge 1: \ |\hat{Q}_m(k)| > g_m(k)\right\}.$$

Following Horváth et al. (2004) the detector $\hat{Q}_m(k) = \sum_{i=m+1}^{m+k} \hat{\varepsilon}_i$ is a CUSUM type statistic of the residuals $\hat{\epsilon}_i = y_i - \mathbf{x}_i^T \hat{\boldsymbol{\beta}}_m$, where the unknown regression parameter $\boldsymbol{\beta}_0$ is estimated by the least squares estimator $\hat{\boldsymbol{\beta}}_m$ based only on the observations from the training period. The boundary function is chosen as $g_m(\cdot) = \sigma c g_m^*(\cdot)$, where the parameter σ is a normalizing constant. A way to obtain a *controlled asymptotic level* α is to *fix the critical constant* $c = c(\alpha)$ such that, under the null hypothesis H_0,

$$\lim_{m \to \infty} P(\tau_m < \infty) = \alpha.$$

Since the methods proposed here are essentially *nonparametric*, the main goal is to **derive a limiting distribution via invariance principles**.

For each m, let $S_m(k) = \sum_{i=m+1}^{m+k} \epsilon_i$. Then, under an appropriate moment condition, we can find two standard Wiener processes $\{W_{0,m}(t),\ t \geq 0\}$ and $\{W_{1,m}(t),\ t \geq 0\}$ and positive constants σ and δ, such that we have a uniform weak invariance principle (IP) over the training period ,i.e.

$$\sup_{1 \leq k \leq m} k^{-1/(2+\delta)} \left| S_0(k) - \sigma W_{0,m}(k) \right| = O_P(1) \quad (m \to \infty), \tag{1}$$

together with a uniform weak IP for the monitoring sequence, i.e.

$$\sup_{1 \leq k < \infty} k^{-1/(2+\delta)} \left| S_m(k) - \sigma W_{1,m}(k) \right| = O_P(1) \quad (m \to \infty). \tag{2}$$

Furthermore, we assume that there is a positive-definite $p \times p$ matrix $\mathbf{C}$ and a constant $\tau > 0$ such that

$$\left| \frac{1}{\ell} \sum_{i=1}^{\ell} \mathbf{x}_i \mathbf{x}_i^T - \mathbf{C} \right| = O\left(\ell^{-\tau}\right) \quad P - \text{a.s.} \quad (\ell \to \infty). \tag{3}$$

With the parameters σ and τ introduced above we define the boundary function

$$g_m(k) = \sigma c g_m^*(k) = \sigma c m^{1/2} \left(1 + \frac{k}{m}\right) \left(\frac{k}{m+k}\right)^{\gamma}, \tag{4}$$

where $0 \leq \gamma < \min\{\tau,\ 1/2\}$ is a certain tuning constant.

Then, under the "no change" null hypothesis H_0 and under an *asymptotic independence condition*, Schmitz and Steinebach (2010) proved

$$\lim_{m \to \infty} P\left(\frac{1}{\sigma} \sup_{1 \leq k < \infty} \frac{\left|\hat{Q}_m(k)\right|}{g_m^*(k)} > c \right) = P\left(\sup_{0 < t \leq 1} \frac{|W(t)|}{t^{\gamma}} > c \right), \tag{5}$$

where $\{W(t),\ t \geq 0\}$ is a standard Wiener process.

2 Research Summary

In my **current research** the asymptotic independence condition needed to prove (5) is resolved via **coupling techniques**, which are, in particular, suitable for strong mixing conditions on the errors. Moreover, according to Horváth et al. (2007), in the case of $\gamma = 1/2$, which is excluded in (4), an asymptotic **extreme value distribution** can be derived via proving a **Darling-Erdős type limit theorem** for independent errors.

In my PhD thesis (Schmitz, 2011) it is shown that certain extreme value asymptotics, **related to retrospective change-point testing procedures**, can be derived via coupling techniques also under strong mixing conditions. The novel feature is that the construction is based on standardized Brownian bridge type approximations, assuming only a **logarithmic decay of the mixing coefficients**. These limit theorems were originally proposed by Csörgő and Horváth (1997) for independent random variables. In the future, these results and techniques shall be applied to extend a change-point test of Ling (2007) towards **near-epoch dependent (NED)** data on an underlying error sequence which obeys a strong mixing condition.

Moreover, the uniform weak invariance principles in (1) and (2) can be derived from strong invariance principles available for stationary processes. In Schmitz (2011) another method to prove **strong invariance principles for linear time series with dependent errors** is presented. This method is based on mixingale approximations. Moreover, some new "backward" strong invariance principles for linear processes with strongly mixing errors are derived. As a consequence, we are able to establish limit theorems for certain weighted tied-down partial sums within an **ARMA-GARCH framework**. In particular, Aue et al. (2006) proposed weight functions to detect structural breaks with better power. In Schmitz (2011) we consider a complementary class of weight functions and derive related limit theorems via using the "backward" approximations. In the future, these results and techniques shall be **extend to the multivariate case**.

Although the methods to derive (5) are essentially nonparametric, an application of the results in practice requires the estimation of the unknown parameter σ. In Schmitz and Steinebach (2010) a class of **consistent estimators for the long-run variance σ** is established via using a nonoverlapping-blocks technique. The investigation of further properties of these estimators in our general setting is an interesting task. This research direction is of course connected with the study of covariance matrices and sample correlation matrices in the multivariate case.

Assumption (3) is a condition on the large sample behavior of the (dependent) stochastic regressor sequence and can be derived from **Marcinkiewicz-**

Zygmund type laws of large numbers. In a recent talk [1], I presented a suitable law of large numbers for NED sequences via improving the rate of convergence in a result of Ling (2007). Moreover, it was shown that moderate versions of the threshold in (4) allow for a more general dependence among the observations including non-stationary martingale differences.

3 Future Research

My future research plans are related to the further investigation of structural breaks in multivariate time series and nonparametric inference of regression functions with jumps. In particular, my research will consider methods of change-point analysis based on invariance principles, empirical processes, permutation and rank based methods. Therefore, in addition to the asymptotic theory of (nonparametric) mathematical statistics, my focus is on strong approximation results and empirical process techniques, especially for (multivariate) dependent data.

References

Aue, A., Berkes, I. and Horváth, L. (2006). *Strong approximation for the sums of squares of augmented GARCH sequences*. Bernoulli, **12**, 583–608.

Chu, C.-S. J., Stinchcombe, M., and White, H. (1996). *Monitoring structural change*. Econometrica, **64**, 1045-1065.

Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*. John Wiley & Sons, Chichester.

Horváth, L., Hušková, M., Kokoszka, P., and Steinebach, J. (2004). *Monitoring changes in linear models*. J. Statist. Plann. Inference, **126** , 225-251.

Horváth, L., Kokoszka, P., and Steinebach, J. (2007). *On sequential detection of parameter changes in linear regression*. Statist. Probab. Lett., **77** , 885-895.

Ling, S. (2007). *Testing for change points in time series models and limiting theorems for NED sequences*. Ann. Statist., **35** , 1213-1237.

[1] slides: `http://www.mi.uni-koeln.de/~schmitza/`

Schmitz, A. (2011). *Limit Theorems in Change-Point Analysis for Dependent Data*. Dissertation. University of Cologne.

Schmitz, A. and Steinebach, J. (2010). *A note on the monitoring of changes in linear models with dependent errors*. In: Dependence in Probability and Statistics (eds. P. Doukhan, G. Lang, D. Surgailis und G. Teyssière), Lecture Notes in Statistics, **200**, 159-174, Springer-Verlag, Berlin.