HYBRID RESAMPLING CONFIDENCE INTERVALS FOR CHANGE-POINT OR STATIONARY HIGH-DIMENSIONAL STOCHASTIC REGRESSION MODELS

Wei Dai and Ka Wai Tsang

The Chinese University of Hong Kong, Shenzhen

Supplementary Material

S1. Proof of Theorem 1 and simulation study in Section 3.1

We first combine the results of Lee and Lai (2009) with those of Lai, Xu, and Yuan (2020) to prove Theorem 1. As pointed out in the first paragraph of Section 2.1, the consistency of $\hat{\boldsymbol{\beta}}_{\hat{J}}$ (using the notation therein) has been established by Lai, Xu, and Yuan (2020). This implies that with probability approaching 1 as $n \to \infty$, $\hat{w}_t := y_t - \sum_{j \in \hat{J}} \hat{\beta}_j x_{tj} \to \epsilon_t$, where x_{tj} and ϵ_t are the same as those in (1.1) and $\hat{\beta}_j (j \in \hat{J})$ are the components of $\hat{\boldsymbol{\beta}}_{\hat{J}}$, hence the asymptotic theory of Lee and Lai (2009, Section 3 and in particular pp.430–431) is applicable to the double block bootstrap in Algorithm 1 under their Assumptions 1–5. Theorem 1 (with the choice $\ell \propto n^{-1/3}$ and the studentized statistics (2.1) together with the choice (2.2) for $m(=m_n)$ in Algorithm 1) then follows from Corollary 1 of Lee and Lai (2009).

We next give in Table 1 the results of the simulation study of the average coverage errors of post-selection confidence intervals described in Section 3.1, for the conditional approach, denoted by C, and the hybrid resampling approach, denoted by H.

Table 1: Average coverage errors for $\beta_j, j \in \hat{J}$ with $2\alpha = 0.2$, $n = 400, p_n = 500$ using the hybrid resampling (H) or conditioning (C) way

(ω, a, b, ν)	Way	$\beta_1 = 0.71$	$\beta_2 = 0.69$	$\beta_3 = 0.52$	$\beta_4 = 0.46$	$\beta_5 = 0.32$	$\beta_6 = 0.28$
$(0.1, 0.1, 0.1, \infty)$	H	0.2020	0.1940	0.2140	0.2040	0.2100	0.1740
	C	0.9828	0.9778	0.9462	0.9407	0.8438	0.8889
$(0.1, 0.1, 0.1, 4)$	H	0.2240	0.2220	0.2320	0.1940	0.1884	0.2345
	C	0.9500	0.9509	0.8626	0.8346	0.8966	0.8000
$(0, 0.55, 0.35, \infty)$	H	0.2060	0.2020	0.2100	0.2000	0.1960	0.2020
	C	1	1	1	1	1	1
$(0, 0.55, 0.35, 4)$	H	0.2191	0.2279	0.2261	0.2343	0.2308	0.2160
	C	0.9716	0.9776	0.9583	0.9587	0.9524	1
$(0.01, 0, 0, \infty)$	H	0.2020	0.2040	0.2140	0.2240	0.2200	0.2040
	C	0.9977	0.9951	0.9848	0.9862	0.9310	0.9286

S2. Proof of Theorem 2 and empirical study in Section 3.2

We first give the proof of Theorem 2, for which we also review the background literature on the EB approach to change-point stochastic regression models. As pointed out in the last sentence of Section 2.2, the estimates of the hyperparameters based on moving windows are consistent. Hence we can apply Algorithm 2 in Section 2.3 "in lieu of the oracle" that reveals the time segments between two consecutive change-times, using the segmentation functionality of the EB change-point modeling approach of Lai and Xing (2013, Section 4) under the following enhancement of (2.5):

Assumption. $p \to 0$ as $n \to \infty$ such that $np \to \lambda \in (0, \infty)$.

This assumption implies a limiting Poisson process of change-times up to time nt, hence the number $\#_n$ of change-times up to time n has a limiting Poisson distribution with mean λ and the interarrival times are asymptotically i.i.d. exponential. If the hyperparameters $\boldsymbol{\eta}, p$ and the parameters $\mathbf{z}, \mathbf{V}, \rho$ and d of the Bayesian model were known, then the oracle procedure could simulate $I_{t_0:n}$ directly for the " exact method" to construct credible intervals for ν_t^2 and scalar functions of $\boldsymbol{\beta}_t (t_0 \le t \le n)$. The last paragraph of Section 2.2 gives consistent estimates $\hat{\boldsymbol{\eta}}, \hat{p}, \hat{\mathbf{z}}, \hat{\mathbf{V}}, \hat{\rho}$ and $\hat{d}$. Hence under the **Assumption** that we have just stated for Theorem 2, the estimates converge to the actual values as $n \to \infty$ and therefore Algorithm 2 to generate $I_{t_0:n}^{(b)} (b = 1, \ldots, B)$ and its associated hybrid resampling method to construct credible intervals differ from the oracle credible intervals by $o_p(1)$, proving Theorem 2.

We next provide the underlying developments of the EB approach to change-point stochastic regression models, beginning with Yao (1984) on the EB approach to estimation of a "noisy discrete-time step function" that was an outgrowth of his PhD thesis at MIT under Herman Chernoff who had initiated the Bayesian approach to estimation of time-varying normal means in Chernoff and Zacks (1964). Lai had been working on sequential change-point detection in quality control and time series in the 1990's, culminating in the discussion paper Lai (1995) presented to the Royal Statistical Society and Lai (1998, 2000) together with Lai and Shan (1999) in *IEEE Transactions*. He then worked with his PhD students to extend the detection and estimation problems from single to multiple change-points. The PhD thesis of Tongwei Liu (2000), Yuguo Chen (2001), Haiyan Liu (2003) and Haipeng Xing (2005) at Stanford under his supervision were in this area of research and led to the papers Chen and Lai (2007) and Lai and Xing (2011,2013) in the References and Lai, Liu and Xing (2005) and Lai, Liu and Xing (2009) in the Additional References. In fact, the last paper was written up first but publication was delayed because of Tongwei Liu's move from Hewlett Packard in the Stanford Research Park to take up a management position at D.E. Shaw Global Investment and Technology Development Company in New York City. On the other hand, Tongwei Liu's thesis already introduced the key ideas that were used by Lai, Liu and Xing (2005) and Chen and Lai (2007), Morever, the delay to 2009 in the *Special Issue of Communication of Statistics in Honor of Shelemyahu Zacks* actually worked well as a tribute to Zacks because the paper was inspired by Chernoff and Zacks (1964) and Yao (1984). We are grateful to Professor Lai for sharing with us the chronological development of the ideas behind these papers.

We now present Tables 2 and 3 together with Figure 2 associated with the empirical study in Section 3.2. Table 2 gives the estimates $\hat{\beta}_1, \hat{\beta}_2, \hat{a}$ and $\hat{b}$ in model (3.1) and the time period covered, for which $\hat{a}+\hat{b} = 1.0$ indicates volatility persistence. It also gives corresponding estimates of the piecewise constant $\beta_{t,1}, \beta_{t,2}$ between two identified consecutive change-times in model (3.2) together with the $\hat{a}, \hat{b}$ values during each time period. Table 3 gives 95% credible intervals for these piecewise constant parameters $\beta_{t,1}, \beta_{t,2}$ and ν_t. Figure 2 plots these estimates over the time period of 416 weeks.

Table 2: Estimates of regression coefficients and a, b in ARX-GARCH model

Model	Period	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{a}$	$\hat{b}$
(3.1)	1/01/01-12/31/08	-0.0778	-0.2229	0.1354	0.8646
(3.2)	1/01/01-12/26/02	-0.0739	-0.0946	0.2904	0.4609
	12/27/02-4/26/07	-0.0216	-0.2296	0.0139	0.9254
	4/27/07-12/22/07	-0.4894	0.0616	0.1184	0.4511
	12/23/07-7/05/08	-1.6337	0.0850	0.1234	0.5721
	7/06/08-12/31/08	0.1215	-0.3164	0.1616	0

Table 3: 95% of credible intervals for piecewise constant parameters

Period	$\beta_{t,1}$		$\beta_{t,2}$		ν_t	
01/01/01-12/31/08	-0.3939	0.1661	-0.3346	0.2254	0.0285	0.0635
12/27/02-4/26/07	-0.1816	0.1384	-0.3896	-0.0696	0.0127	0.0227
4/27/07-12/22/07	-1.2894	0.3906	-0.4184	0.6216	0.0277	0.0627
12/23/07-7/05/08	-2.6737	-0.5137	-0.4750	0.5650	0.0370	0.1320
7/06/08-12/31/08	-0.9985	1.0015	-0.9564	0.2436	0.0985	0.1585

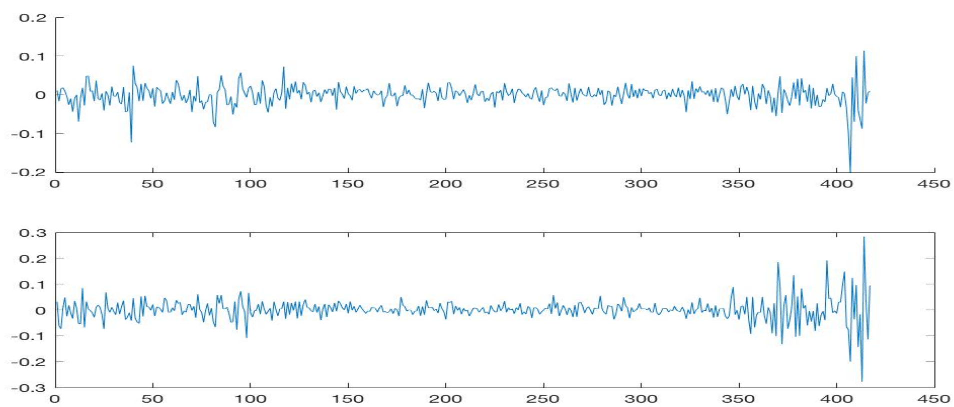

Figure 1: Weekly returns of SP500 index(top),WFC(bottom)

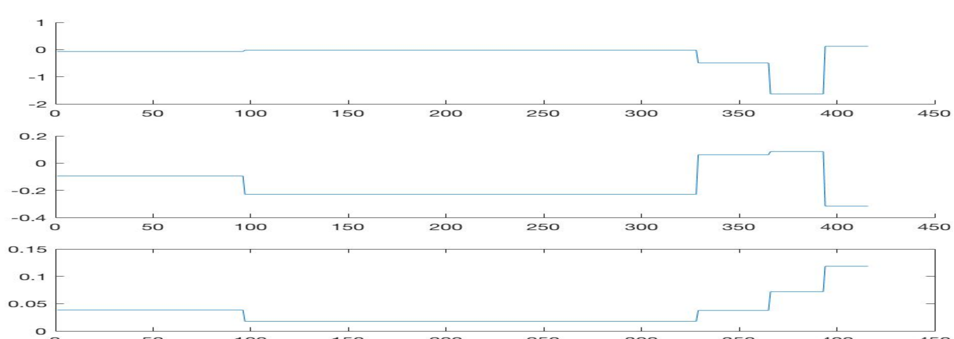

Figure 2: Estimates of time-varying $\beta_{t,1}$(top), $\beta_{t,2}$(Middle), ν_t(bottom)

Additional References

Chernoff, H. and Zacks, S. (1964). Estimating the current mean of a normal distribution which is subjected to changes in time. *Ann. Math. Statist.* **35**, 999–1018.

Davis, M.H.A. (1975). The application of nonlinear filtering to fault detection in linear systems. *IEEE Trans. Automatic Control* **20**, 257–259.

Diaconis, P. and Ylvisaker, D. (1979). Conjugate priors for exponential families. *Ann. Statist.* **7** 269–281.

Lai, T. L. (1995). Sequential change-point detection in quality control and dynamical systems (with discussion). *J. Roy. Statist. Soc. Ser. B* **57**, 613–658.

Lai, T. L. (1998). Information bounds and quick detection of parameter changes in stochastic systems. *IEEE Trans. Information Theory* **44**, 2917–2929.

Lai, T. L (2000). Sequential multiple hypothesis testing and efficient fault detection-isolation in stochastic systems. *IEEE Trans. Inform. Theory* **46**, 595-608.

Lai, T. L., Liu, H. and Xing, H. (2005). Autoregressive models with piecewise constant volatility and regression parameters. *Statistica Sinica* **15**, 279–301.

Lai, T. L., Liu, T. and Xing, H. (2009). A Bayesian approach to sequential surveillance in exponential families. *Communications in Statistics – Theory and Methods* **38**, 2958–2968.

Lai, T. L. and Shan, J. Z. (1999). Efficient recursive algorithms for detection of abrupt changes in signals and control systems. *IEEE Transactions on Automatic Control* **44**, 952–966.

Yao, Y. (1984). Estimation of a noisy discrete-time step function: Bayes and empirical Bayes approaches. *Ann. Statist.* **12** 1434–1447.