Extraction of a local linear trend from physiological time series

Dr. rer. nat. Roland Fried, Prof. Dr. rer. nat. Ursula Gather
Department of Statistics, University of Dortmund, 44221 Dortmund, Germany
Dr. Michael Imhoff,
Surgical Department, Community Hospital Dortmund, 44137 Dortmund, Germany

Abstract

We discuss methods for robust signal extraction from noisy physiological time series as measured in intensive care. The aim is a method which allows a fast and reliable de-noising of the data and separation of artifacts from relevant changes in the patients condition[1]. For approximating local linear trends we use robust regression estimators. We examine the performance of the L_1 regression, the repeated median[2] and the least median of squares[3] for this task.

Introduction

In intensive care, clinical information systems acquire and store physiological variables online at least every minute. Reliable automatic monitoring systems are needed to process these data in real time and to support decision-making at the bedside in time critical situations. Clinically relevant changes such as sudden level shifts and trends need to be detected and to be distinguished from noise and irrelevant artifacts. Median filtering is frequently applied for signal extraction from a noisy time series[1]. However, the performance of the median worsens markedly when a trend occurs[4]. Therefore, we use robust regression estimators as these are able to adapt to local trends[5]. Our main interests are the reproduction of a linear trend, the detection of level shifts, the detection of trend changes, and the computational demands.

Methods

Robust regression. For robust extraction of a local linear trend we apply robust estimators designed for the simple linear regression model $y = \mu + \beta x + e$. An intuitive approach to robust estimation of μ and β is to replace squared by absolute distances using L_1 regression. The finite sample replacement breakdown point denotes the minimal percentage of outliers which suffices to carry the estimate beyond all bounds. In our situation, for L_1 regression it is about 29.3%. The repeated median[2] and the least median of squares[3] (LMS) both share the breakdown point of the median, which is approximately 50%. For calculation of the exact LMS solution[6] we use an algorithm which has a computational complexity of $O(n^4)$. Hence, the time needed for computation of the LMS is much larger than for the other methods and increases rapidly with the window width.

Simulation model. In order to compare the regression estimators in a single time window we generate data from the model $Y_t = \mu + \beta t + e_t$, $t=-m, \ldots , m$, with $\mu=0$ and for several slopes β. The error e_t is simulated from an autoregressive model of order one [AR(1)] $e_t = \phi e_{t-1} + u_t$ with mean zero, ϕ in (-1,1) and Gaussian innovations u_t with unit variance $\sigma^2=1$. Estimates of μ and β allow to approximate the level and the slope of the signal at the center $t = 0$ of the time window. The resulting time delay m is determined by the required stability (m large) and the time delay possible in an online application (m small). We consider the cases m=10, 15, 25.

Results and Conclusions

We find that the L_1 regression offers little advantage in comparison to the repeated median. The LMS is the least influenced by outliers if ϕ is close to zero, but this advantage disappears as $|\phi|$ increases. Moreover, the LMS is very variable and computationally demanding. The repeated median does also resist some outliers, is computationally much less expensive and does not show severe instabilities. Hence, we consider it to be a promising candidate for robust signal extraction when computation time is critical.

References

1. Mäkivirta A, Koski E, Kari A Sukuvaara T The median filter as a preprocessor for a patient monitor limit alarm system in intensive care. Computer Methods and Programs in Biomedicine 1991; 34: 139-144.
2. Siegel AF Robust regression using repeated medians. Biometrika 1982; 68: 242-244.
3. Rousseeuw PJ Least median of squares regression. J Amer Statist Assoc 1984; 79: 871-880.
4. Fried R, Gather U Fast and robust filtering of time series with trends. In: Härdle W, Rönz B editors. COMPSTAT 2002. Proceedings in Computational Statistics. Heidelberg: Physica 2002, pp. 367-372.
5. Davies PL, Fried R, Gather U Robust signal extraction for on-line monitoring data. J Statistical Planning and Inference, to appear.
6. Stromberg AJ Computing the exact least median of squares estimate and stability diagnostics in multiple linear regression. SIAM J Sci Comput 1993; 14: 1289-1299.

Supported, in part, by the DFG (SFB475)