Analysis of Online Monitoring Data from Intensive Care

Ursula Gather
University of Dortmund, Department of Statistics
Vogelpothsweg 87
D-44221 Dortmund, Germany
gather@statistik.uni-dortmund.de

1. Introduction

In critical care, nowadays clinical information systems generate extremely high-dimensional time series. More than one hundred time dependent vital variables are recorded permanently by online monitoring systems. This yields new and challenging tasks for physicians but also for statisticians. The natural goal is to make sensible and profitable use of this mass of information in medical decision making. As humans are not able to extract the most important structure out of such high-dimensional time series just by pure visual impression, statistical methods are necessary to provide a reasonable bedside decision support.

Here, besides the aim of recognizing outliers, level changes and intervention effects in each univariate time series of a relevant vital variable, it is important to find ways of reducing the dimension of the high-dimensional monitored time series. It is known from medical background knowledge that many of the recorded variables are strongly dependent, so it is near at hand to look for procedures which are able to detect the underlying dependence structure automatically and to reduce the dimension of the time series in an appropriate way.

2. Online pattern detection in time series

A basic step for monitoring critically ill patients and providing suitable bedside decision support for the physician is the detection of abnormal patterns of change in the observed time series. Patterns such as level shifts and slow trends have to be detected quickly and correctly, and they have to be distinguished from outliers caused by measurement artifacts and clinically irrelevant short term fluctuations. Typically, there are large positive autocorrelations within physiological time series. These should be included into the data analysis (Endresen and Hill, 1977) since otherwise we will often incorrectly detect a monotone trend in the data (Woodward and Gray, 1993).

Bauer, Gather and Imhoff (1999) develop an automatic procedure for online detection of patchy outliers and level shifts. This approach is based on an embedding of the time series into an m-dimensional Euclidean space. The marginal distribution of m subsequent observations is considered and rules for outlier detection in multivariate data are thus transferred into the context of time series. The dimension m should be chosen according to the sample partial autocorrelations. Rules for distinguishing between outliers and level shifts can be formulated easily if the differenced series is used since a single outlier in the observed series induces two outliers in the differenced series, while a level shift induces just one outlying difference. More complicated patterns can be distinguished by monitoring the movement of the vectors of time delayed observations outside an estimated control ellipsoid. This control ellipsoid can be constructed using a Mahalanobis type distances of the vectors from the current estimated level of the process, where all estimates should be calculated from a moving time window since physiological time series can only be regarded as locally stationary. Simulation studies show that this approach is more powerful for patchy outliers and level shifts than outlier detection based on one-step-ahead prediction for instance.

Methods for trend detection often assume a trend to be linear. However, in possibly life-threatening situations we should not rely on this assumption since non-linear trends might not be detected this way. Brillinger (1989) modifies a mini-max approach developed by Abelson and Tukey (1963) for the detection of any monotone trend to autocorrelated data. This approach can be adapted to the situation of online monitoring by calculating the test statistic from a current time window. Time intervals which are not very long should be used since physiological time series can only be regarded as locally stationary. Therefore some simplifying assumptions have to be applied (Fried, Gather and Imhoff, 2001). The autocovariances within the noise process can be modelled locally by a low order AR(p) model as is common practice in statistical process control (cf. Lu and Reynolds, 1999). A simple possibility to estimate the autocovariances is to eliminate firstly a linear trend by usual regression methods and thereafter to calculate the sample autocovariances of the residuals. Simulation studies as well as applications to real physiological time series have been performed. Both linear and non-linear trends and could be detected at least as early as by an experienced physician.

3. Dimension reduction

In intensive care a multitude of physiological variables is measured in the course of time. It is well-known that there are strong dependencies between these variables. Physicians usually select one variable out of a group of closely related variables and base their decisions on the patterns found in this variable only. Alternatively, we can use statistical methods for dimension reduction, e.g. dynamic factor analysis (Peña and Box, 1987), to compress the relevant information into a few important variables. However, the latent variables found in the analysis have to be interpretable since physicians need to understand the meaning of an alarm. Therefore we should use methods such as graphical models for multivariate time series to learn about the dynamic relations between the variables and to deduce appropriate assumptions for the construction of a factor model from the data (Gather, Imhoff and Fried, 2000). In a case-study, a ten-variate time series consisting of arterial and pulmonary arterial pressures (diastolic, systolic, mean), central venous pressure, heart rate and blood temperature was analyzed. Application of graphical models revealed that there were four groups of closely related variables and thus four interpretable factors could be obtained for this ten-variate time series of vital signs. These latent factors were found to be more adequate for detection of patterns in the observed variables than each single variable (Gather et al., 2001). Moreover, application of graphical models to time series sampled for individual patients during different clinical states such as pulmonary hypertension or congestive heart failure shows that these states can be characterized by different partial correlation structures. For instance, during the state of pulmonary hypertension there are mainly strong partial correlations between the arterial and the pulmonary arterial pressures, while during the status of vasopressure support there are strong partial correlations between the heart rate and the arterial pressures, too. These findings agree with and supplement existing medical knowledge. Therefore we expect to gain further insights into the causes of clinical complications as well as for detecting such complications by applying graphical models.

4. Perspective

Statistical methods for the analysis of physiological variables offer an opportunity for a more reliable evaluation of the individual treatment. A future task is the construction of intelligent bedside decision support systems. Such a system will incorporate statistical techniques for pattern recognition within multivariate time series as we have outlined here. These techniques could be combined with methods of artificial intelligence which use the patterns found in the statistical analysis to assess the current state of the patient. By classifying these patterns according to existing knowledge (gained from physicians and further data analysis) the physician can then be given options of how to respond (Morik et al., 2000) .

REFERENCES

Abelson, R.P. and Tukey, J.W. (1963). Efficient utilization of non-numerical information in quantitative analysis: general theory and the case of simple order. *Ann. Math. Statist.*, **34**, 1347-1369.

Bauer, M., Gather, U. and Imhoff, M. (1999). The identification of multiple outliers in online monitoring data. Technical Report 29/1999, SFB 475, Department of Statistics, University of Dortmund, Germany.

Brillinger, D.R. (1989). Consistent detection of a monotonic trend superposed by a stationary time series. *Biometrika*, **76**, 23-30.

Endresen, J. and Hill, D.W. (1977). The present state of trend detection and prediction in patient monitoring. *Intensive Care Medicine*, **3**, 15-26.

Fried, R., Gather, U. and Imhoff, M. (2001). The online detection of a monotone trend in a time series. Preprint, Department of Statistics, University of Dortmund, Germany.

Gather, U., Fried, R., Lanius, V. and Imhoff, M. (2001). Online monitoring of high-dimensional physiological time series - a case-study." *Estadistica*, to appear.

Gather, U., Imhoff, M. and Fried, R. (2000). Graphical Models for Multivariate Time Series from Intensive Care Monitoring. Technical Report 33/2000, SFB 475, University of Dortmund, 44221 Dortmund, Germany.

Lu, C.-W. and Reynolds, M.R. Jr. (1999). EWMA control charts for monitoring the mean of autocorrelated processes. *Journal of Quality Technology*, **31**, 166-187.

Morik, K., Imhoff, M., Brockhausen, P., Joachims, T. and Gather, U. (2000). Knowledge discovery and knowledge validation in intensive care. *Art. Int. Med.*, **19**, 225-249.

Peña, D. and Box, G.E.P. (1987). Identifying a Simplifying Structure in Time Series. *J. Americ. Stat. Assoc.*, **82**, 836-843.

Woodward, W.A. and Gray, H.L. (1993). Global warming and the problem of testing for trend in time series data. *Journal of climate*, **6**, 953-962.

RESUME

Dans les unités de soins intensifs, les systèmes d'information fournissent des séries chronologiques de dimension extrêmement élevée. Ces séries servent de support aux processus de decision. Des méthodes statistiques sont proposées en vue de permettre la détection des observations aberrantes, des changements de niveau et des effets d'interventions dans chacune des séries prises de façon univariée. Des procédures de réduction de la dimension des données par l'étude des dépendances croisées sont également discutées.