Johannes Schmidt-Hieber
Mathematical Institute
Leiden University
schmidthieberaj@math.leidenuniv.nl

Column Tenure-tracker

Reconstruction of risk measures from financial data

In this column holders of a tenure track position introduce themselves. The tenure track positions in mathematics became available in 2013. Excellent researchers could apply in several expertise areas of mathematics. Johannes Schmidt-Hieber has a tenure track position at Leiden University.

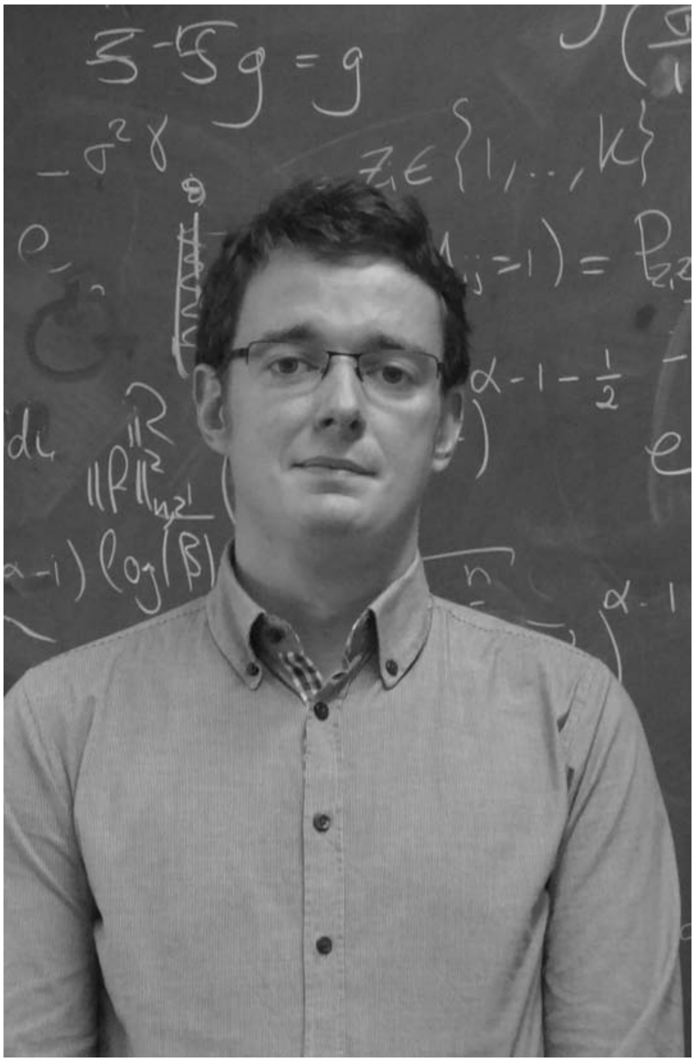
The last years have seen a growth in statistical applications in which the data and the unobserved quantity of interest are linked by a complex mechanism. Additionally, the object that we wish to reconstruct from the data can have in its own a complicated structure. To handle such problems requires the development of new methods and a detailed mathematical understanding of the underlying models which are currently subject of intensive research.

To illustrate the power of modern statistical procedures, we give a short survey on spot volatility estimation, which is a very intricate problem within nonparametric statistics. Here, the data come from financial assets, traded on short time intervals, such as milliseconds. The goal is to reconstruct from these observations the so called spot volatility which is a (random) function measuring the local variability of the price over time. The spot volatility is an important quantity for risk management and since it cannot be observed directly, stable reconstruction methods are crucial.

The main issue with these data is that there are two layers of randomness. The first one originates from the price dynamic; under the efficient price hypothesis, this dynamic describes the evolution of a financial asset over time. But the observed prices are additionally perturbed by a second layer of randomness. One reason for that are rounding errors; prices are usually given in full cent. But there are various other effects, which are summarized under the generic term 'microstructure noise'. From a statistical perspective, microstructure noise is very unpleasant as it makes the spot volatility almost invisible in the data. As the name indicates, microstructure effects are comparably small. But, if we channel the data through a reconstruction procedure for the spot volatility, the outcome will typically be far away from the truth. In statistics, errors should cancel out, microstructure effects, however, tend to add up and to dominate the reconstructions.

To find a reconstruction of the spot volatility which is unaffected by microstructure effects, one can study the following simple toy model first. Suppose we observe n prices $Y_{1,n},\dots,Y_{n,n}$ at time points i/n with

$$Y_{i,n} = \sigma W_{i/n} + \epsilon_{i,n}, \quad i = 1,\dots,n.$$



Here, $(W_t)_{t\geq 0}$ denotes a Brownian motion modeling the efficient price and σ is the (constant) spot volatility. The random variables $\epsilon_{i,n}$ are independent standard normal distributed and represent the microstructure noise. The two sources of randomness (efficient price and microstructure noise) are assumed to be independent. What we observe is therefore a Gaussian process and all information is contained in its covariance structure. Looking at the eigenvalues of the covariance, one finds that only few of these eigenvalues (of the order of $\sqrt{n}$ out of n) contain information about the quantity of interest σ. Taking this into account, σ can now be reconstructed from the data $Y_{1,n},\ldots,Y_{n,n}$ by a filtering and weighting scheme. Although a solution in a restricted model does not automatically provide a solution in a more general setting, it might nevertheless lead to new insights. And here it does. In fact, understanding the toy model above leads to a general reconstruction method that is unaffected by microstructure noise. One even can show that if more and more data are available, the so reconstructed spot volatility approximates the truth with the best possible speed which any reconstruction method in this problem theoretically could achieve.

The following real data example relies on prices of Euro-BUND futures (FGBL) for which up to 30.000 trades per day are available. In Figure 1, the data and the reconstructed spot volatility for one trading day are displayed. Apparently, the spot volatility is quite stable and increases slightly until 2 p.m. During this time period, we see therefore an increased market uncertainty and higher fluctuations of the underlying efficient price. In the early afternoon the method detects large variations of the spot volatility. These changes are typically linked to external events, such as macroeconomic announcements. Notice that we do not observe the classical volatility pattern, where the volatility is high at the opening and low during a 'lunch break'. For our reconstructions we use a notion of volatility that is independent of the underlying trading intensity which causes these standard patterns. The representation of the spot volatility chosen in Figure 1 is in particular suitable for detection of events that are specific to one day.

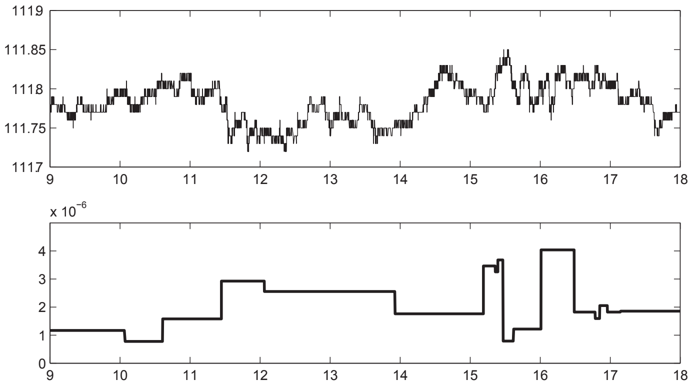


Figure 1 Application to real data. Upper: Price on June, 4th 2007 between 9 a.m. and 6 p.m. Lower: Reconstructed spot volatility. (Figure taken from [1])

Day	Market Uncertainty	maximal spot volatility during announcement
Jan-11	0	0.459
Feb-08	0	0.541
Mar-08	0	0.490
Apr-12	0	0.331
May-10	0	0.330
Jun-06	0	0.191
Jul-05	0	0.587
Aug-02	0.05	1.286
Sep-06	0.1	0.906
Oct-04	0.03	0.621
Nov-08	0	0.869
Dec-06	0	1.119
average over days above		0.644
average over all days in 2007		0.551

Table 1 Market uncertainty and maximal spot volatility during monthly announcements on key interests rates. (Table taken from [1])

As application, we study the effect of macroeconomic events on the spot volatility. Table 1 shows the maximal value of the reconstructed spot volatility during the monthly announcements of the European Central Bank on key interest rates in 2007, which are of major importance for economy. Few days before the release, the predictions of analysts on the outcome are recorded. The first column in Table 1 shows the standard deviation of these predicted values. A high standard deviation refers to a high market uncertainty and should therefore be reflected by a larger spot volatility. Indeed the values in Table 1 suggest such a relationship. Secondly, we find that the average maximum of these days is higher than the average over all days in 2007 giving further evidence that key interest announcements lead to an increase in uncertainty.

An implementation of the presented method and further examples can be found in the Matlab based SpotvolToolbox available from www.stochastik.math.uni-goettingen.de/SpotvolToolbox.

Biography

Johannes Schmidt-Hieber studied mathematics and theoretical physics in Freiburg, Göttingen, and at University of California, Davis. In 2010, he received his PhD from University of Göttingen and University of Bern. Since then, he has been a postdoc at Vrije Universiteit Amsterdam and at ENSAE in Paris. His research interests are in nonparametric statistics and in particular on nonparametric methods for stochastic processes. He has also worked on fractional processes and deconvolution problems. His current focus is on understanding the Bayesian method for high-dimensional and nonparametric models.

Reference

1 T. Sabel, J. Schmidt-Hieber and A. Munk (2014), Spot volatility estimation for high-frequency data: adaptive estimation in practice, to appear in *Modeling and Stochastic Learning for Forecasting in High Dimension*, Springer Lecture Notes in Statistics.