Time-varying spectral analysis on Hilbert spaces

Valorization

Since 2014, a valorization addendum is a mandatory part of all doctoral theses at Dutch Universities. Roughly speaking, the addendum should provide social justification of the research undertaken. More precisely, according to the National Valorization Committee the term valorization stands for *" The process of creating value from knowledge by making knowledge suitable and/or accessible for economic and/or social exploitation and translating it into competitive products, services, processes and new activities"*. In this addendum, the knowledge valorization of the topic of this thesis – i.e., practical and theoretical aspects of time–varying spectral analysis on Hilbert spaces – is outlined. I shall do this by discussing how the topic and the respective findings of my doctoral thesis are of social and economic relevance.

This thesis has been centered around the concept of serially correlated stochastic processes – i.e., time series – in particular those with a changing correlation structure over time. With nonzero serial correlation we mean that a stochastic process – a process that is subject to random variation and evolves over time – is possibly influenced by its own past. Equivalently, a nonzero serial correlation structure implies that, if we decompose the process in terms of its cyclical behavior by representing it in the spectral domain, the energy that is dissipated by the process is not equally spread over frequencies. A socially and economically relevant example where the frequency content of a stochastic process is of importance can come from engineering. Decomposition of earthquake vibrations into its components allows to determine which vibrations, of different speeds and amplitudes, account for most variation in the signal. Buildings can then be designed to avoid interaction with the strongest components.

As stated in the introduction of this thesis, many physical phenomena, including the above example, exhibit nonstationary behavior as a result of smooth changes in their second–order structure. For example, in meteorology the atmospheric turbulence shows clear changes when measured over time. Atmospheric turbulence are the small, irregular air motions characterized by winds that vary in speed and direction. These turbulences affect how water vapor, energy and other substances such as smoke are distributed and therefore cause for instability in the atmosphere. Moreover, daily records of temperature, precipitation and cloud cover over a region as three related surfaces may change over time due to global climate changes. Anticipation to weather conditions is vital in different areas of society and can help reduce economic as well as social damages to a minimum. In order to optimally anticipate to these conditions, weather prediction models need to take this nonstationary behavior into account.

In economics, empirical studies have shown that macroeconomic data, such as interest rates or variables related to the gross domestic product, exhibit smooth changes when measured over longer periods of time or on a fine enough time resolution. Central banks and other research institutions will have to integrate this behavior in their development of prediction models. This is of foremost importance since these models form the basis of an economic outlook and are consequently used by policy makers to adjust government spending and fiscal policy. Another important related example comes from finance, where implied volatility of an option as a function of moneyness changes over time. This information is relevant among others to investors, banks and insurance companies.

In medicine, different recording techniques of brain activity such as electroencepahlograms (EEG's), functional magnetic resonance images (fMRI) and local field potentials (LFP's) show that the underlying dynamics of the brain process have spectral properties that evolve over time. These changing dynamics are important to take into consideration in order to correctly model how, for example, different regions of the brain such as the nucleas accumbens or the hippocampus, are involved in certain cognitive processes. In chapter 2, the introduced data–adaptive method was applied to local field potentials recordings.

These are only a few applications where the process is characterized by time–varying spectral characteristics. The list of disciplines in which such processes occur is however much longer and includes geophysics, astronomy, sound analysis and electrical engineering. Just as for the examples provided above, these disciplines are all involved in answering socially and economically relevant questions. To answer these questions, they rely on statistical tools and techniques that are made available by the statistical or econometric community or yet related fields. Failing to properly take into account the time–varying second–order characteristics will make the underlying model inappropriate and accordingly any inferences drawn from it invalid.

Despite of its importance, the majority of methods applied is still based on the assumption that the underlying data generating process is either time–invariant or that the nonstationarity is caused by abrupt changes. Another type of nonstationary processes that has received a considerable amount of attention in the (econometric) time series literature, are so–called integrated processes. These can be classified by the property that the differenced series are stationarity. These types of nonstationary processes have in common that their analysis generally does not require a completely different framework to derive statistical properties. This in contrast to the case where the second–order structure is varying over time.

As made clear by the wide range of disciplines listed above, changes in the second–order structure are however the rule rather than the exception. It was already mentioned in the introduction of this thesis that the extension of many estimation methods to processes of which the data generating mechanism is in a constant state of change, is not a natural one. Not only will the classical theoretical framework – on which many statistical inference procedures are based – became meaningless, the decomposition in terms of frequency components might no longer have the same physical interpretation. In many applications, the research question however requires knowledge on how the process dissipates energy over different frequency components. A clear example of this was provided at the beginning of this addendum, where we motivated the relevance to correctly decompose earthquake vibrations. Even when the research question is not directly formulated in terms of its frequency content, it can often still be advantageous to proceed the analysis of the series in the spectral domain. The second–order dependence structure of a time series can be completely characterized in the spectral domain via the spectral density (operator). Especially when one is facing high–dimensional data, when parametric modeling cannot be justified or when the process exhibits cyclical behavior, the spectral domain can provide a useful alternative to the time domain.

Relatively recent (Dahlhaus, 1996a) a framework for finite–dimensional time series with time–varying characteristics was introduced that allows for both meaningful statistical inference as well as for a spectral theory in which concepts such as 'frequency' and 'energy' keep the same physical interpretation. Although this has led to a surge in the literature on processes with time–varying spectral characteristics, there are still many open problems. An important one is the problem to estimate the time–varying spectrum in practice. Because this object is a function of both time and frequency and because theoretical results are not directly applicable, the practitioner is forced to face the uncertainty principle. That is, the practitioner needs to find the right balance between the required estimation precision in time and frequency direction, given that more precision in one direction directly means a loss of precision in the other direction.

Another more recent problem is that the surge in data storage techniques has led to the need for models that can deal with data that are intrinsically infinite–dimensional. In effect, many modern datasets can be viewed as sampled recordings from complex mathematical structures such as curves or surfaces. The extension of methods that are suitable in finite–dimensional spaces to methods that are suitable in infinite–dimensional spaces is far from straightforward and requires careful consideration of various convergence concepts as well as the consideration of appropriate dimension reduction techniques. The field of research involved with this during the past few decades is known as functional data analysis. A particular subfield, known as functional time series, is concerned with ordered collections of functional data and focuses on the development of methods that take into account the second–order dependence structure. These methods rely however on the assumption that this dependence structure remains constant over time and thus that the series is functional weakly stationary. Not unlike the finite–dimensional setting, this assumption turns out to be too restrictive in many applications. The aforementioned meteorological elements as related surfaces as well as the implied volatility surface are examples of functional time series with changing second–order characteristics. The development of statistical techniques to model this type of functional data correctly has however remained an open problem.

This thesis has addressed these two open problems. That is, this thesis has been concerned with both the practical problem of the estimation of time–varying spectral densities as well as the need for the development of theory and methodology for infinite–dimensional stochastic processes that are characterized by changing second–order characteristics. In Chapter 2, a data–adaptive approach was introduced for the estimation of time–varying spectra. Without going into details again, the method allows the data to tell the practitioner what would be the 'correct' balance between time– and frequency resolution for each time–varying spectral variate in the time–frequency plane. The algorithm as developed will be provided as a software package and will therefore be easily accessible to any practitioner that is facing this problem. Chapter 3 was devoted to develop spectral theory and methodology to model functional time series that have time–varying spectral characteristics. Although the results of Chapter 3 are theoretical and not of a nature that the practitioner can apply directly, they are a necessary step in understanding the behavior of these type of processes. This in turn is a necessary step in the development of methods that will become available to the practitioner. An important aspect of research is that it paves the way for future research. This chapter does exactly that by providing the foundation for a framework that enables statistical inference of nonstationary functional time series and that allows the derivation of large sample approximations of estimators and test statistics. Chapter 4 builds in fact on the framework of Chapter 3 by providing a test statistic that allows to test for the presence of time–varying spectral characteristics. This test allows the practitioner to determine whether standard inference methods can be used or that these become invalid and that alternative methods based on the framework as proposed in Chapter 3 need to be considered in order to draw correct inferences.

I hope the knowledge valorization provided in this addendum will make clear the importance of the research undertaken in this dissertation. An outlook into possible future research has been provided in both the conclusion of the relevant chapters as well as in the overall conclusion (Chapter 5). These projects, as well as the implementation of the proposed methods into easily accessible software packages, are something I look forward to work on in the near future.