Week 11

Examples of Applications of Spatial and Spatio-Temporal Processes

Overview

Welcome to Week 11 of Spatial Analysis.

This week we will cover three topics:

- Working with Big spatial data

This week concludes our studies of spatial processes and models. We will consider problems to be addressed when one analyses Big spatial data. We will illustrate how to apply the previously discussed methods to the real estate data on Irish sold properties and New York City noise complaints cases. You will see applications of both, random field and point processes approaches.

By the end of this week, you will learn about:

Topic 1: Working with Big spatial data

- Analysis of Big Spatial Data.
- Analysis of Irish Property Prices by using kriging.
- Analysis of the spatial distribution of Irish sold properties

Topic 2: Analysing Ney York Noise Complaints

- Data preparation
- Analysis of helicopter noise complaints
- Analysis of noise in parks complaints

By completing this module, you will be working towards the following subject-intended learning outcomes:

1. Formulate purposeful questions to explore new statistical ideas and subsequently design valid statistical experiments.
2. Present clear, well-structured analysis of important statistical model results.
3. Creatively find solutions to real-world problems consistent with those commonly faced by practising statisticians.
4. Professionally defend or question the validity of existing statistical analyses and associated evidence-based conclusions that are derived via application of sound spatial statistical methodology.

Topic 1: Working with Big spatial data

In this topic, we will demonstrate how the previously developed methods can be applied to the analysis of Big spatial data.

We will start with the general challenges that data scientists have in the analysis of big spatial data. Then, we will outline some of the most common approaches in the modelling and analysis of big spatial data.

As an example, we will use the Property Price Register data about Irish properties from the years 2012-2017. In particular, this dataset has information about GPS coordinates and prices for approximately 220k property sales. We will start with the preparation of this dataset for analysis. In the first part of the analysis, we will use random field models and kriging to produce a prediction map of property values in this region. In the second part, we will investigate the spatial distribution of locations for sold Irish properties.

In real applications, one may need to analyse much larger spatial data. In such cases, the mentioned strategies and methods can be scaled and applied in a similar way.

Read

You will read the topic <u>Working with Big Spatial Data</u> and will apply methods you learned before to the Irish properties' dataset.

From the folder, open and read the document titled **Week_11_Topic_1**.

Topic 2: Analysing NYC noise complaints

This topic concludes this subject. To analyse noise incidents in New York from 2018, we will be use the dataset provided by the New York Open Data Portal. This dataset contains detailed information about reported noise complaints, including their location, type, time, and other attributes for each recorded incident.

Our primary focus will be on the initial stages of analysis for noise complaints related to helicopters and loud music, parties, and talking in New York City parks. We will begin by analysing these two types of noise complaints separately by using a variety of spatial point process methods. On the second stage, we will compare the spatial point patterns of these two types of noise complaints and investigate their joint behaviour.

Read

You will read the topic <u>Analysis of Noise Complaints</u> and will apply methods you learned before to the NYC noise complaints dataset.

From the folder, open and read the document titled **Week_11_Topic_2**.

Workshop

Activity

Workshop

This activity will be completed in R. Repeat the R programming content covered in Week 11. Modify the code and understand the impact of different R parameters on changes in results.

Your task

- Repeat R commands learnt in Week 11 prior to the two-hour workshop session.
- Try to modify the code and understand the impact and meaning of different R function parameters. Interpret the observed changes in plots and analysis results.
- Feel free to discuss questions with other students as you go in the forum, and please also take the time to help others. It is amazing how much we all can learn from each other's questions, and how in helping others we strengthen our own understanding.
- Come along to the two-hour scheduled workshop session and discuss any challenges, seek advice and work through some problems with your peers and facilitator.
- Revisit these problems in later weeks and challenge yourself to get a deeper understanding to build on what you learn later.

Guidelines

- This activity is not graded but is an essential part of your learning. It will facilitated by your instructor.
- You don't need to submit your R code; however, to be successful in this subject it is necessary to work through all R coding materials from this week and understand how to apply the corresponding R commands.
- You should repeat all R commands in this week's materials before the workshop. This will give you an opportunity to efficiently work with the facilitator during the workshop and get your questions answered.
- You should spend around two hours on this activity.

Summary

This week, we finished the studies of statistical and data science models based on spatial processes. We considered applications of the method from the previous topics to the analysis of Big spatial data.

Here's a list of tasks that you should be working on or have completed:

- **Required readings**
- **Workshop**

The following resources provide you with this week's references and additional suggested readings.

Additional suggested readings and resources

While these readings and resources are not essential, they provide greater insight into the concepts covered in the week's lectures and give you the choice to enhance your learning or pursue an area of interest in greater detail.

Software and data:

- CRAN documentation for the spatstat package:
  https://cran.r-project.org/web/packages/spatstat/index.html

Theoretical concepts:

- Spatial point processes:
  https://www.apps.stat.vt.edu/leman/VTCourses/BaddeleyPointProcesses.pdf

Books:

- Bivand, R. S., Pebesma, E., & Gomez-Rubio, V. (2013). Applied spatial data analysis with R (2nd ed.). Springer. https://doi.org/10.1007/s12061-014-9118-y Avalilable on-line in La Trobe EBL ebook Library
- Baddeley, A., Rubak, E., & Turner, R. (2015). Spatial Point Patterns: Methodology and Applications with R. Chapman and Hall/CRC. https://doi.org/10.1201/b19708 Available on-line in La Trobe EBL ebook Library
- Wikle C.K., Zammit-Mangion A., Cressie N. (2019). Spatio-Temporal Statistics with R. Chapman and Hall/CRC. Avalilable on-line from https://spacetimewithr.org/