

國立成功大學理學院

110 學年度第一學期模組化課程

無母數迴歸方法

Nonparametric Regression Methods

授課教師

任職單位

畢業學校

林立祥

Louisiana State University
(路易斯安那州立大學)

Georgia Institute of Technology GA, USA
(喬治亞理工學院)

課程類別

學分數

選必修

開課人數

其他注意事項

Lecture

1.5

選修

20

因 COVID-19 疫情嚴峻，本課程將延至 2022 暑假(111-1)。

Recitation

先修課程或先備能力

建議修過一學期統計學

課程難易度

難 中偏難 中偏易 易

建議修課學生背景

理學院，工學院，電資學院，管理學院

教學方法

講授 50%，實作 25%，討論 15%，報告 10%

補充說明：

每天課程中，9:00-12:00 主要為觀念與方法課堂講解，13:00-15:30 會透過電腦與 R 語言對上午所學的觀念與方法進行實作。課程進行中會對重要觀念，設計討論題與課程報告並提供實際例子與學生討論，加深學生對重要觀念的熟稔度。

評量方式

報告 50%，出席 20%，課堂簡報 30%

補充說明：

為鼓勵學生合作學習，培養團隊精神，課程進行時會讓學生以 2-3 人為一組的分組方式完成課程指定報告與簡報呈現。

(i) 課堂簡報 (30%): 學生需要在最後一天上課時呈現上台選定報告主題的初步資料分析結果。報告內容應包含講解使用的無母數迴歸方法，與分析結果呈現。簡報時間依實際報告組別數有所調整，會在課程進行中另行公布。

(ii) 課程書面報告 (50%): 學生於課程結束後一周內以Email方式繳交一份課程書面報告，該報告需清楚描述報告的研究問題與目標並包含詳細建模及分析無母數資料的流程。報告的目標是訓練學生建立自我學習更多無母數方法的能力，授課老師將會在課程時間以外安排額外討論的時間與各組學生討論適當主題並指導 (非強迫參加)，目前暫定的討論時間為

第二天課程結束後15:30-17:30

第三天課程結束後15:30-17:30

第四天課程結束後15:30-17:30

第五天課程開始前08:00-09:00

對於報告主題，學生亦可選擇自己有興趣的無母數模型主題做延伸探討。報告的初步分析結果

國立成功大學理學院

110 學年度第一學期模組化課程

會在課程最後一天呈現，並依據授課老師的建議，將報告完成修改後於課程結束一周內以 Email 方式繳交。詳細繳交方式與報告格式會在課堂公布。

學習規範

學生應參與所有理論與實作課程並完成報告與簡報

課程概述

無母數迴歸是研究自變量和依變量之間非線性關係的重要工具。此課程由講義課和 R Lab 課程組成。講義課旨在介紹常見的無母數迴歸方法，並討論各種實際使用無母數迴歸的相關議題，包括測試非線性效應，測量模型偏差和變異，評估預測不確定性。R Lab 課程供學生練習從講義課中學到的分析技巧。另外，此課程將指導並協助學生建立自我學習課堂上未涵蓋的無母數方法之能力、撰寫課程報告及上台口頭報告。

課程概述(英文)

Nonparametric regression is a powerful tool to explore nonlinear relationships between input and output variables. The course is composed of lecture sections and R Lab sections. The lecture sections provide an overview of nonparametric regression methods. Various practical aspects of using nonparametric regression, including testing the nonlinear effects, measuring the model bias and variance, assessing the prediction uncertainty, are to be discussed. The R lab sections are designed for students to practice the analysis skills they learned from the lecture sections. In addition, the students will be guided and assisted in learning certain nonparametric regression methods not covered in the classes by themselves, making short course projects, and presenting the projects.

課程進度

| 堂次 | 時間 | 進度說明 |
|----|------------|---|
| | 9:00-15:30 | 09:00-12:00 Introduction to Nonparametric Modeling <ul style="list-style-type: none">• Review on linear modeling and likelihood-based concepts• Examples for nonparametric regression 12:00-13:00 Lunch Break 13:00-14:00 R Lab: Basic Programming 14:00-15:30 R Lab: Basis Expansion Approaches and Penalized Approaches |
| | 9:00-15:30 | 09:00-10:30 Kernel Regression 10:30-12:00 Some Theories for Nonparametric Regression <ul style="list-style-type: none">• Bias-variance tradeoff 12:00-13:00 Lunch Break 13:00-14:30 R Lab: Tuning Smoothing Parameters by Cross-Validation <ul style="list-style-type: none">• Basic concept of cross-validation• Fast computation of cross-validation 14:30-15:30 R Lab: Constructing ANOVA Table for Nonparametric Regression |

國立成功大學理學院

110 學年度第一學期模組化課程

| | | |
|--|------------|---|
| | 9:00-15:30 | 09:00-12:00 Model Evaluation through Bootstrapping <ul style="list-style-type: none">• Basic concept of bootstrapping• Estimate model bias and variance• Construct confidence interval and confidence band 12:00-13:00 Lunch Break 13:00-15:30 R Lab: Practice of Bootstrapping for Nonparametric Regression |
| | 9:00-15:30 | 09:00-12:00 Access Prediction Uncertainty by Conformal Prediction <ul style="list-style-type: none">• Difference between confidence intervals and prediction intervals• Basic concept of conformal prediction 12:00-13:00 Lunch Break 13:00-15:30 R Lab: Practice of Conformal Prediction |
| | 9:00-15:30 | 09:00-12:00 Project Presentations and Discussions 12:00-13:00 Lunch Break 13:00-14:30 Project Presentations and Discussions 14:30-15:30 R Lab: More Nonparametric Regression Techniques |

課程學習目標

1. Learning techniques for exploring nonlinear relationships in regression analyses.
2. Developing data analysis skills in measuring nonlinear model performance.
3. Enhancing storytelling and presenting skills in data analytics.

課程的重要性、跨域性與時代性

無母數迴歸方法在現今資料科學及機器學習的許多應用領域扮演者不可或缺的角色，因為這些應用往往無法透過一般的線性模型解釋，需要透過資料驅動(data-driven)的建模方法找出適合的模型做進一步統計推論，這些統計推論問題及相關解決手法，亦會在課程中介紹，希望藉此幫助學生建立使用無母數迴歸解決實際應用的能力。

其他備註

參考書目：

- Wolfgang Härdle. (1990) *Applied nonparametric regression*. Cambridge: Cambridge university press.
- Grace Wahba. (1990) *Spline models for observational data*. Society for industrial and applied mathematics.
- Jianqing Fan and Irene Gijbels. (1996) *Local polynomial modelling and its applications*. Boca Raton: CRC Press.
- Larry Wasserman. (2006) *All of nonparametric statistics*. New York: Springer.
- Alexandre B. Tsybakov (2008) *Introduction to nonparametric estimation*. New York: Springer.
- Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani (2013). *An introduction to statistical learning*. New York: Springer.
- Paul H.C. Eilers and Brian D. Marx (2021) *Practical Smoothing: The Joys of P-splines*. Cambridge: Cambridge University Press.