

Title

Mathematical Optimization in Data Science

Teacher

Dolores Romero Morales, Copenhagen Business School

Expected literature

B. Baesens (2014), *Analytics in a Big Data World: The Essential Guide to Data Science and its Applications*. Wiley and SAS Business Series.

T. Hastie, R. Tibshirani and J. Friedman (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd Edition. Springer.

F. Provost and T. Fawcett (2013), *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. O'Reilly.

Theme 1

Data Science

Topics

- Data Science
- Examples
- The Process
- Data

Preparatory Reading

B. Baesens (2014), *Analytics in a Big Data World: The Essential Guide to Data Science and its Applications*. Wiley and SAS Business Series. [Chapters 1-2](#).

F. Provost and T. Fawcett (2013), *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. O'Reilly. [Chapters 1-2](#).

Additional Reading

T.H. Davenport and D.J. Patil (2012), Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review* 90(10): 70–76.

M. Schroeder, R. Shockley, J. Smart, D. Romero Morales, P. Tufano (2012), Analytics: The real-world use of big data, *Research Report by the IBM Institute for Business Value in collaboration with Saïd Business School*.

Theme 2

Classification

Topics

- Supervised Learning
- Classification
- Classification Trees
- Nearest Neighbors
- Support Vector Machines

Preparatory Reading

B. Baesens (2014), *Analytics in a Big Data World: The Essential Guide to Data Science and its Applications*. Wiley and SAS Business Series. [Chapter 3 \(Sections 4 and 6\)](#).

T. Hastie, R. Tibshirani and J. Friedman (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd Edition. Springer. [Chapter 4, 9, 13](#).

F. Provost and T. Fawcett (2013), *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. O'Reilly. [Chapter 4-6](#).

Additional Reading

E. Carrizosa and D. Romero Morales (2013), Supervised Classification and Mathematical Optimization. In: *Computers and Operations Research*, Vol. 40, No. 1, p. 150–165.

A. Karatzoglou, D. Meyer and K. Hornik (2006), Support Vector Machines in R. In: *Journal of Statistical Software*, Vol. 15, No. 9, p. 1–28.

PC Workshop

- Support Vector Machines with R

Theme 3

Clustering

Topics

- Unsupervised Learning
- Dissimilarities
- Hierarchical Clustering
- k-means Clustering
- k-medians Clustering

Preparatory Reading

B. Baesens (2014), *Analytics in a Big Data World: The Essential Guide to Data Science and its Applications*. Wiley and SAS Business Series. [Chapter 4 \(Section 3\)](#).

T. Hastie, R. Tibshirani and J. Friedman (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd Edition. Springer. [Chapter 14 \(Sections 1, 3\)](#).

F. Provost and T. Fawcett (2013), *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. O'Reilly. [Chapter 6](#).

PC Workshop

- Clustering with R

Theme 4

Visualization

Topics

- Visualization
- Dimensionality Reduction
- Multidimensional Scaling
- Principal Components Analysis

Preparatory Reading

T. Hastie, R. Tibshirani and J. Friedman (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd Edition. Springer. [Chapter 14 \(Sections 5, 7\)](#).

Recommended Reading

I. Borg and P.J.F. Groenen. *Modern Multidimensional Scaling: Theory and Applications*. Springer. 2005.

I. Borg, P.J.F. Groenen and P. Meir. *Applied Multidimensional Scaling*. Springer. 2013.

E. Carrizosa, A. Nogales Gomez, and D. Romero Morales. Strongly Agree or Strongly Disagree? : Rating Features in Support Vector Machines. *Information Sciences*, 329:256-273, 2016.

B. Johnson and B. Shneiderman. Tree-maps: A space-filling approach to the visualization of hierarchical information structures. In *IEEE Conference on Visualization*, pages 284-291. 1991.

S. Liu, W. Cui, Y. Wu, and M. Liu. A survey on information visualization: recent advances and challenges. *The Visual Computer*, 30(12):1373-1393, 2014.

J. Thomas and P.C. Wong. Visual analytics. *IEEE Computer Graphics and Applications*, 24(5):20-21, 2004.