SCHOOL IN "DIGITISE, VISUALISE AND OPTIMISE"
Data science and convex optimisation methods for empirical finance
Version: May 15, 2020

Time and location: February 3-5 and 8-9, 2021. The school with be delivered either online or in the Executive Center of the Università della Svizzera Italiana (USI), Via Buffi 13, CH-6900 Lugano, Main Building, Second Floor.

Instructor:

Fabio Trojani, Professor of Finance and Statistics, SFI Senior Chair, University of Geneva and SFI; mailto: fabio.trojani@alphacruncher.com; Web page: `http://www.people.usi.ch/trojanif/`

Course description: The school studies recent advances in data science and optimisation methods for studying relevant research questions in finance and economics. A large number of these questions involves solving corresponding convex optimisation problems, derived from suitable objective functions and related constraints, which can be systematically solved in a number of important cases using efficient optimisation techniques. We will provide a broad self-contained introduction to these methods and their applications, with a particular focus on questions related to (i) the solution of relevant financial models and (ii) the use of statistical techniques for analysing the properties of these models with real data. After the course, students should be able to:

- Identify model structures allowing an efficient treatment with convex optimisation methods, or be able to reformulate models and problems accordingly where this is appropriate
- Select and combine existing Python packages to efficiently code solutions for financial models and statistical procedures related to their empirical analysis

Course pre-requisites: Interest in the applications of quantitative modelling, data-science and optimisation methods to finance and economics, as well as interest in exploring key properties of financial and economic research databases. While the course is self-contained, basic knowledge of programming language Python and of principles of optimisation is needed. Regarding the knowledge of programming language Python, students in the school will have free access to a reserved class environment on datacamp (`http://www.datacamp.com`), allowing them to acquire basic knowledge in Python, by completing following courses in datacamp before the start of the school:

- *Introduction to Python*: `http://www.datacamp.com/courses/intro-to-python-for-data-science`
- *Introduction to Matplotlib*: `http://www.datacamp.com/courses/introduction-to-matplotlib`
- *Pandas Foundations*: `http://www.datacamp.com/courses/pandas-foundations`
- *Python Data Science Toolbox (Part 1)*:

  `https://www.datacamp.com/courses/python-data-science-toolbox-part-1`

Regarding the knowledge of principles of optimisation, we strongly encourage students to work our before the start of the school some selected subsections of Stephen Boyd et al.' short course on optimisation, available on-line at `http://web.stanford.edu/~boyd/papers/cvx_short_course.html`. More precisely, please work out following subsections of this course:

1. *Convex optimisation overview*: Slides 1-20 and 24-33, available at:

   `http://web.stanford.edu/~boyd/papers/pdf/cvx_opt_intro.pdf`

   A video of this part of the course is available at: `http://www.youtube.com/watch?v=9sDidkln7R0`

2. *Constructive convex analysis and disciplined convex programming*: Slides 1-18, 25-27, 31-33, 37, and 44, available at:

   `http://web.stanford.edu/~boyd/papers/pdf/cvx_dcp.pdf`

   A video of this part of the course is available at: `https://www.youtube.com/watch?v=PFVOTBQJvSE`

3. *Convex optimisation applications*: Slides 1-11 and 30-33, available at:

   `http://web.stanford.edu/~boyd/papers/pdf/cvx_applications.pdf`

   A video of this part of the course is available at: `https://www.youtube.com/watch?v=IlC7WvfdByo`

Material: Self-contained lecture notes and teaching material will be provided, covering key concepts and ideas necessary to understand and apply selected data-science and optimisation methods to relevant research questions in finance and economics. During the school, students will have full access to Nuvolos (`http://nuvolos.cloud`), an integrated data science and computational platform, on which they will directly implement in Python some of the introduced models and optimisation methods, based on real financial data from established research databases.

Teaching philosophy: The course aims to provide students in a workshop-like style with key skills and competences allowing them to apply selected data-science and optimisation methods to relevant research questions in finance and economics. More broadly, the course provides students with useful experience, intuition and implementation skills that are typically gained after structuring and addressing a research project with a substantial data-analytic and optimisation component. Lectures provide methodological background and context that students need to deepen, both with their reading of the selected material to the course and with their implementation work in groups. Lectures will be interactive, using (i) Python (`http://www.python.org`) and CVXPY (`http://www.cvxpy.org`), as a programming language and an optimisation package, and (ii) Nuvolos (`http://nuvolos.cloud`), a browser-based complete cloud in which files, datasets, code and applications work together, in order to directly implement the introduced models and optimisation methods using real data from established research databases.

Basic schedule: The school will last 7 days, with Days 1-5 of lectures and workshops, and two additional days of independent self-study and group work. On Days 1-5 lectures and workshops will be held according to the following structure. Lectures will be held in the morning from 9:00-12:00. Python workshops will be held from 13:30-15:00, while theoretical and coding exercises in groups will be held from 15:30-17:00. On the two days of independent self-study and group work, students will have the opportunity to further elaborate on the theoretical concepts and coding aspects developed on Days 1-3, to complete unfinished exercises and/or coding assignments, and to start working on the final graded project for the school.

Preliminary table of contents:

1. Day 1
   (a) Introduction and motivation
   (b) Abstract optimisation problems, existence of solutions and Weierstrass theorem
   (c) Convex relaxations, min-max inequalities and strong duality
   (d) Moreau approximations and convex conjugates
   (e) Short recap on basic convexity definitions and properties
   (f) Python workshop: International asset allocation with transaction costs
2. Day 2
   (a) Generalised inequalities and $K-$convexity
   (b) Smallest elements, dual cones and Pareto efficiency
   (c) Canonical convex problems: definition and key properties
   (d) Nesting hierarchy: From quadratic to conic programs
   (e) Python workshop: Factor reduction and sparse principal components
3. Day 3
   (a) Quasi convex and robust optimisation
   (b) Equivalent formulations of optimisation problems and standard forms
   (c) Minimum and minimal elements

   (d) Vector and multi-criterion optimisation
   (e) Python workshop: Thematic portfolio investing with multi-criterion optimisation

⋆ Independent self-study and group work
   (a) Work on completing exercises and code from Day 1-3
   (b) Start of work on final school project

4. Day 4
   (a) Lagrange dual function, Lagrange multipliers and saddle-point property
   (b) Lagrange duality and Slater's constraint qualification
   (c) Karush-Kuhn-Tucker conditions and Wolfe duality
   (d) Reformulations of optimisation problems via duality
   (e) Python workshop: Building cross-sections of sorted portfolios with forests

5. Day 5
   (a) Short recap on Banach spaces, dual spaces and bounded operators
   (b) Semi-continuity, closeness and relative interiors
   (c) Moreau's theorem and its applications
   (d) Convex optimisation and Fenchel duality
   (e) Minimum divergence problems in infinite dimensional spaces
   (f) Python workshop: Model-free asset pricing in large asset markets

Grading and school project: Grading will be based for 30% on active class participation, for 70% on the graded take home project and the corresponding oral exam. The school project consists of (i) a set of selected theoretical and code exercises and (ii) a small research project to be developed in groups.

Selected reading material:

1. Bryzgalova M. Pelger, and J. Zhu, (2019). Forest through the Trees- Building Cross-Sections of Stock Returns,(2019), working Paper.
2. V. De Miguel, A. Martin-Utrera, F. J. Nogales, and R. Uppal (2017). A portfolio perspective on the multitude of firm characteristics, working paper.
3. J. Freiberger, A. Neuhier and M. Weber (2016). Dissecting characteristics nonparametrically, Working Paper.
4. S. Korsaye, A. Quaini and F. Trojani (2018). Smart SDFs, working paper.
5. S. Korsaye, F. Trojani and A. Vedolin (2019). The global factor structure of exchange rates, working paper.
6. E. G. Luttmer (1996). Asset pricing in economies with frictions. Econometrica, 1439–1467.
7. F. Methling and R. von Nitzsch (2018), Thematic Portfolio Optimization: Challenging the Core Satellite Approach, working paper.
8. J.-J. Moreau (1962). Fonctions convexes duales et points proximaux dans un espace hilbertien. Comptes Rendus de l'Académie des Sciences, pages 2897–2899.
9. S. Nagel, S. Kozak, and S. Santosh (2018). Shrinking the cross-section. Journal of Financial Economics.
10. R. Tibshirani. Regression shrinkage and selection via the lasso (1996). Journal of the Royal Statistical Society. Series B, 267–288.

11. R. Uppal and P. Zaffaroni (2016). Portfolio choice with model misspecification: a foundation for alpha and beta portfolios, working paper.

12. R. Uppal, P. Zaffaroni, and I. Zviadadze (2018). Beyond the bound: Pricing assets with misspecified stochastic discount factors, working paper.

13. C. Harvey, Y. Liu and H. Zhu, 2016. ...and the cross-section of expected returns. Review of Financial Studies, vol. 29, issue 1, 5-68.

14. H. Zou and T. Hastie (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B, 67(2):301–320.

Courses and references on convex optimization and convex analysis:

1. Convex optimisation:

   `http://web.stanford.edu/class/ee364a/lectures.html`

   `http://www.stat.cmu.edu/~ryantibs/convexopt/`

2. Disciplined convex programming and CVX

   `http://web.stanford.edu/class/ee364a/lectures/cvx_lecture_slides.pdf`

3. Convex analysis

   `https://dspace.mit.edu/bitstream/handle/1721.1/70523/6-253-spring-2004/contents/index.htm?sequence=4&isAllowed=y`