Statistics

INCOMPLETE DATA ANALYSIS

Missing ~~Data~~

NA

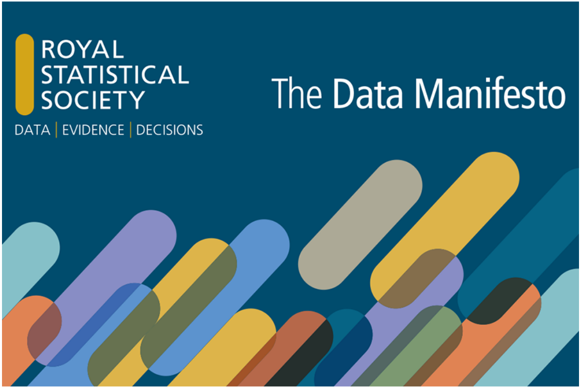
Miguel de Carvalho

THE UNIVERSITY *of* EDINBURGH
School of Mathematics

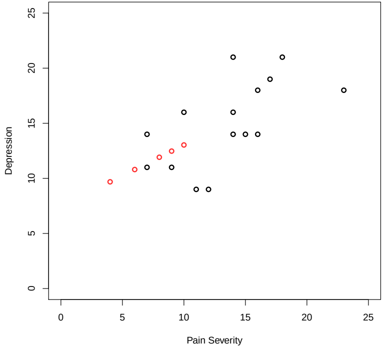
The Data Manifesto

Royal Statistical Society

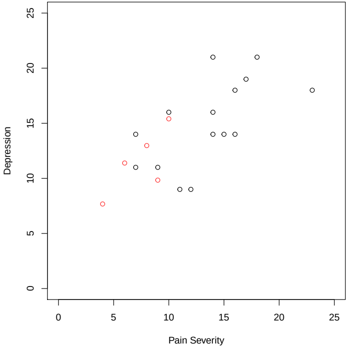
"What steam was to the 19th century, and oil has been to the 20th, data is to the 21st. It's the driver of prosperity, the revolutionary resource, that is transforming the nature of economic activity, the capability that differentiates successful from unsuccessful societies."



- If data are the 21st century driver of prosperity, then a solid understanding of missing data is key to that prosperity.
- Missing data are ubiquitous in practice. They are so common in applications and case studies that most data analysis packages have reserved a special symbol just to encode them: `NA` (which stands for '`Not Available`').
  - a patient might dropout from a clinical trial.
  - a company might decide not to answer a survey—or some questions on that survey.
- The reasons for missingness may be varied and depending on its exact reasons a certain statistical analysis might be sound—or it might be completely flawed.



Regression imputation



Stochastic regression imputation

Incomplete Data Analysis

The Sixth Sense

MATH11176 – Statistical Programming

http://www.drps.ed.ac.uk/23-24/dpt/cxmath11176.htm

Generalised Regression Models (MATH11187)
Course organiser: Bruce Worton

Highlights:

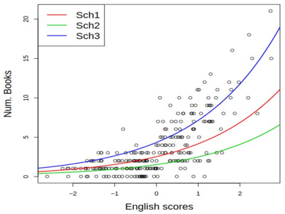
1. **Overview:** Statistical modelling, including linear, GLM, random effects models.
2. **Core:** Blend of theory motivated by case studies and applications.
3. **Applications:** Identify and apply appropriate statistical models to data and interpret the corresponding results.
4. **Modelling:** Extensive use of R to fit models.

Case study – Books!

The file *Books.dat* contains the data from a survey of pupils from three schools, asking how many books they have read outside of school over the last year.

Variable	Description
books	The number of books read in the last year outside of school.
school	The school that the pupil attends (`Sch1/Sch2/Sch3`).
gender	Gender of the pupil identified at birth (`M/F`).
math	Standardized maths test scores.
eng	Standardized English test scores.

$$
\begin{aligned}
Y_i &\sim \text{Poisson}(\lambda_i) \quad \text{for} \quad i = 1, \ldots, n \quad \text{where} \\
\log(\mathbb{E}[Y_i]) &= \beta_0 + \beta_1 x_{1,i} + \cdots + \beta_p x_{p,i}
\end{aligned}
$$



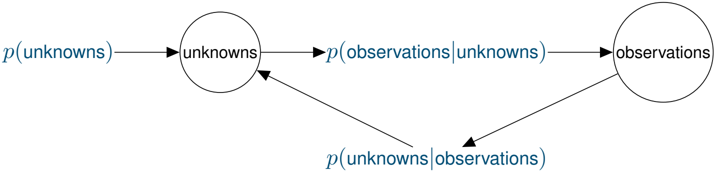
MATH11177: Bayesian Theory

Finn Lindgren
finn.lindgren@ed.ac.uk

2023/24 Semester 1 (Tuesdays)

Bayesian Theory: Probabilistic modelling, estimation, and prediction

- Using prior knowledge about unknown quantities



- Constructing probabilistic models
- Interpreting the results of a Bayesian analysis
- Using probability theory to construct computational methods for Bayesian estimation
- In widespread use in applied statistics and machine learning

THE UNIVERSITY of EDINBURGH

INTRODUCTORY PROBABILITY AND STATISTICS
ACADEMIC YEAR: 2023-2024

Course Organizer: Skarleth Carrales

THE UNIVERSITY *of* EDINBURGH

Course Description:

This course introduces students to the fundamental concepts of probability and statistics, focusing on their applications in various fields. The course equips students with the necessary tools to understand and analyse uncertainty, randomness, and data variability.

Prerequisites:

Prior knowledge of basic algebra and calculus is recommended

Assessment:

This course is delivered 100% online.

1) There is a continuous weekly assessment quiz on Stack (50%).
2) Final report covering all course materials (40%).
3) Peer review (10%).

There is an active discussion forum for any question that can be answered for the CO or the students.

Course Objectives:

- Build a strong foundation in probability theory and statistical principles.
- Learn to analyse and interpret data variability and central tendencies.
- Develop proficiency in calculating probabilities and solving problems using probability models.
- Acquire essential skills for hypothesis testing and drawing confidence intervals.
- Understand key probability distributions and their applications.

THE UNIVERSITY of EDINBURGH

Key Topics Covered:

- Introduction to Probability and Set Theory.
- Conditional Probability and Independence.
- Discrete and Continuous Random Variables.
- Probability Distributions: Binomial, Poisson, Exponential, Normal.
- Expected Values, Moments, and Central Limit Theorem.
- Basics of Statistical Inference, Hypothesis Testing, and Confidence Intervals.
- Simple Linear Regression and Correlation.

Learning outcomes:

- Introductory practical R programming
- Develops essential quantitative reasoning skills.
- Provides a foundation for understanding and interpreting data in various fields.
- Equips students with tools for making informed decisions based on uncertainty.
- Preparatory knowledge for more advanced statistical and data analysis courses.

MATH11311 – Time Series

http://www.drps.ed.ac.uk/23-24/dpt/cxmath11131.htm

Bayesian Data Analysis

- **Learning objective:** The course aims to provide practical experience applying Bayesian analysis to various datasets using interpretable statistical models.
- **Programming:** The statistical analyses will be conducted in R using the widely used computer packages R-INLA, JAGS and Stan.
- **Lecturer**: Daniel Paulin
- **Contents of the course:**
  - Brief review of Bayesian ideas, model checking and convergence diagnostics.
  - Introduction to INLA.
  - Introduction to JAGS.
  - Introduction to STAN.
  - Linear and generalised linear models (fixed effects).
  - Hierarchical Bayesian models, generalised linear models with random effects.
  - Further topics, including spatial and temporal models.
- **Assessment:** 50% coursework (2 assignments) + 50% exam.

THE UNIVERSITY of EDINBURGH
School of Mathematics

MATH11205: Machine Learning in Python

Course Organizer: Sara Wade
sara.wade@ed.ac.uk

Lecturer: Ozan Evkaya
oevkaya@ed.ac.uk

University of Edinburgh
13 September, 2023

Course Description

Pre-requisite: Python Programming (MATH11199) OR Computing and Numerics (MATH08065) → numpy, pandas, matplotlib

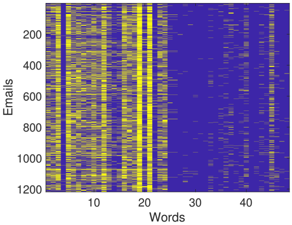
Summary: theory and ideas behind ML techniques, and applications using python libraries, e.g. scikit-learn.

Topics:

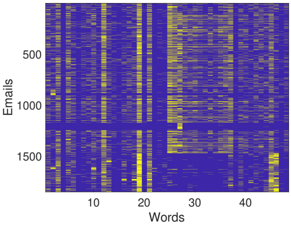
1. Introduction to Machine Learning
2. Model Evaluation: training, testing, cross-validation
3. Regression: least squares, regularization, nonlinear regression
4. Classification: logistic regression, support vector machines, trees
5. Unsupervised Learning: dimension reduction, clustering
6. Introduction to Neural Networks

Course Structure

- Lecture hours: 15; Computer workshops: 15.
- Assignment (40%): two group projects (1-4 students), 20% each.
- Weekly workshop completion (10%).
- Exam (50%).



(a) Spam



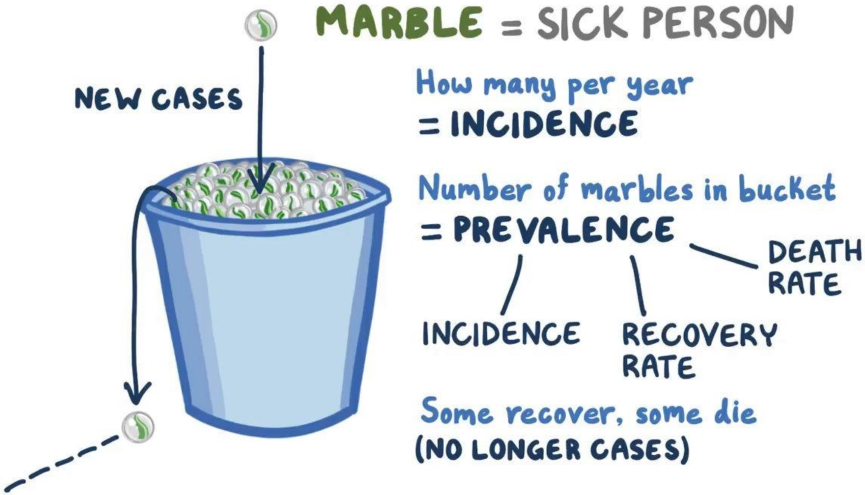
(b) Not Spam

Figure: Spam dataset: binary matrix indicating the presence of words, among a dictionary of 48 words, for spam emails (left) and non-spam emails (right).

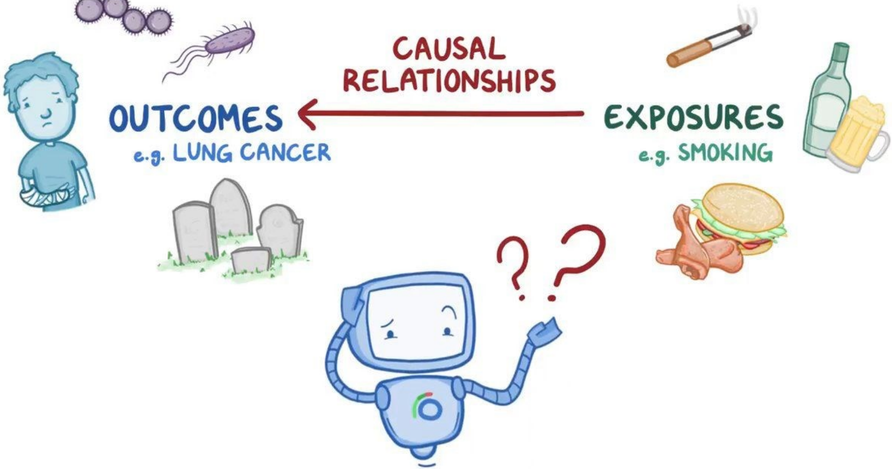
Biostatistics

MATH11230 | 10 credits | Semester 2 | Dr Maarya Sharif

Optional course for SwDS, SOR and OR programmes



https://www.osmosis.org/learn/Incidence_and_prevalence



https://www.osmosis.org/learn/Incidence_and_prevalence

CASE CONTROL STUDY

past exposures

risk factor A

risk factor B

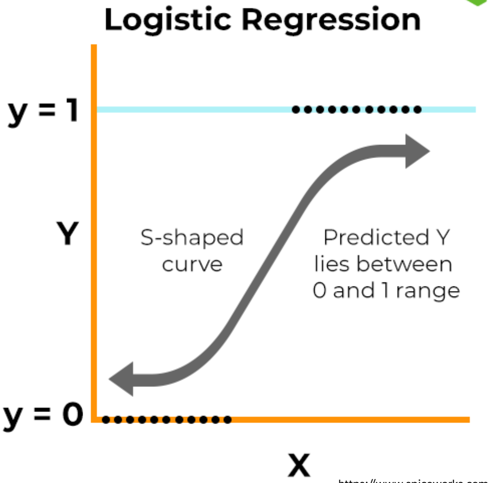
risk factor C

risk factor D

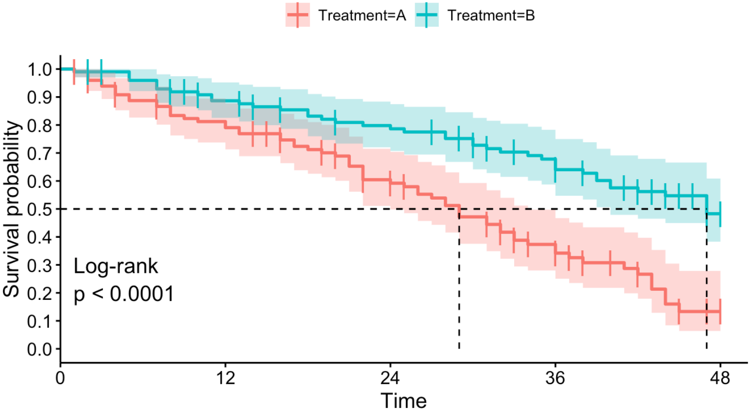
outbreak investigation

$$\frac{\text{odds of being exposed - case}}{\text{odds of being exposed - control}} = \text{Odds Ratio (OR)}$$

https://www.youtube.com/watch?v=Jd3gFT0-C4s



https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-logistic-regression/



https://towardsdatascience.com/kaplan-meier-curves-c5768e349479

Targeted Causal Learning (MATH11238)

Dr Sjoerd Beentjes

sjoerd.beentjes@ed.ac.uk

- Semester 2
- Level 11, 10 credits
- Aimed primarily at SwDS MSc students, others welcome
- Prerequisites: Probability, Statistical Programming (or similar)

Assessment:

- Formative (0%, for feedback)
- Summative (20%)
- Exam (80%)

Targeted Causal Learning (MATH11238)

Aim: Answer **causal** questions based on real-world data, aided by **modern ML algorithms**, whilst preserving statistically **valid inference**

Minimise model-misspecification, target the question of interest

Approach:

- Develop theory underpinning “targeted” estimates
- Learn to apply theory and estimators to relevant examples and case studies using available R packages (“tlverse” ecosystem)
- Examples from biomedicine, policy making, finance, etc.

Learning outcome:

Understanding of theory + practical experience allowing the use of these techniques to obtain statistically valid answers to causal questions

Targeted Causal Learning (MATH11238)

Aim: Answer **causal** questions based on real-world data, aided by **modern ML algorithms**, whilst preserving statistically **valid inference**

Weeks 1-2: Causal questions (Neyman-Rubin, Pearl), (non-)examples e.g., (i) what is the causal effect of a drug on health outcome, (ii) what part of that effect is mediated via increasing/decreasing blood pressure?

Weeks 3-5: How does randomisation relate to causality? From randomised to observational data. Randomisation in temporal data

Weeks 6-10: General Targeted Learning roadmap (semiparametric statistics, plug-in bias, one-step estimator, ensemble machine learning)

Operational Research

MATH11007 – Methodology, Modelling and Consultancy Skills

http://www.drps.ed.ac.uk/23-24/dpt/cxmath11007.htm

Stochastic Modelling

CO 2023-24: Theo Assiotis

What is the course about?

- Course is about stochastic processes $(X(t); t \geq 0)$.
- Models natural phenomena evolving in time in some stochastic way.
- Key notion is that of a **Markov chain**. Special case of stochastic process.
- Special property: "*Future of the process is independent of the past given the present*".

Why study this?

- The mathematics is beautiful.
- Many applications. Markov chains come up in:
  - Mathematical physics.
  - Biology.
  - Finance.
  - Optimization.
  - Machine learning.
  - Pure mathematics.

What will we do?

- Build some theory.
- See some examples.
- Solve some problems.
- Assessment (Homework 20 % + Exam 80%) mainly focused on practical problem solving.
- **WARNING** If you have no background in probability you will need to do some background reading and be very motivated. Otherwise the course will be **very hard.**

MATH11111 – Fundamentals of Optimization

http://www.drps.ed.ac.uk/23-24/dpt/cxmath11111.htm

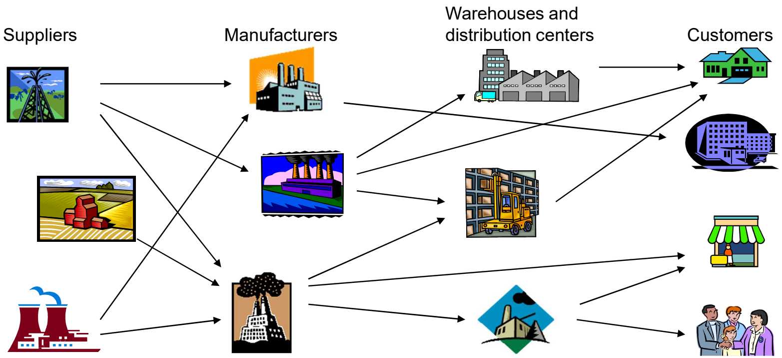
MATH11147 – Large Scale Optimization for Data Science

http://www.drps.ed.ac.uk/23-24/dpt/cxmath11147.htm

MATH11183 – Topics in Applied Operational Research

http://www.drps.ed.ac.uk/23-24/dpt/cxmath11183.htm

Supply Chain



Simchi-Levi et al., Designing and managing the supply chain, 2007

CSCMP – State of Logistics Report 2023

- **U.S. business logistics** costs now stands at a record $2.3 trillion.
  ≈ 9% of national GDP

- The U.S. e-commerce market **grew by 8%, to $1.03 trillion.**
  It is now 14.5% of the entire U.S. retail market.

Challenges in Logistics & SCM

- A supply chain is a very **complex network of facilities** dispersed over a large geographical area and in many cases all over the globe

- Different facilities in the supply chain often have conflicting objectives

- The supply chain is a **dynamic system** that **evolves over time**, e.g. customer demand and supplier capabilities as well as supply network relationships change over time

- Uncertainty is inherent in every supply chain
  - Customer demand can never be **forecasted exactly**
  - Travel times will never be **certain**, and
  - Machines or vehicles **may break down**
  - Facilities may have to **shut down temporarily**

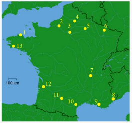
INTEGER AND COMBINATORIAL OPTIMIZATION
MATH11192, Level 11

Lecturer: Sergio García Quiles.
Contact Details: sergio.garcia-quiles@ed.ac.uk, Office 6230.
Prerequisites: Fundamentals of Operational Research (MATH10065) and Fundamentals of Optimization (MATH11111).

Description
In many optimization problems, the solution is found among a set of finite elements. Typical such problems are routing problems, matching problems or scheduling problems. However, the space search can be very large (combinatorial explosion) and, as a consequence, exhaustive search is usually prohibitive. Therefore, specialized mathematical techniques must be used to explore the solution space in an efficient way. In order to study these techniques, it is important to understand fundamental notions from integer programming and graphs theory (total unimodularity, matching, spanning tree, etc.) as well as general techniques (Lagrangean relaxation, branch-and-cut, metaheuristics).

This course will study exact and heuristic methods for solving several of the most important integer and combinatorial optimization problems. We will first cover some basic notions in integer programming and graph theory. Later, they will be applied to the study of specific problems and solution algorithms.

Example
In the comic book Asterix and the Banquet (*Le Tour de Gaul d'Astérix*), Asterix and Obelix have to travel around Gaul and collect several regional culinary specialities for a banquet upon their return in order to save their village from the threat of the Roman Inspector General Overanxious. In their tour, they visit 12 villages and then return home. Given that it is a race against the clock, they visit each village exactly once. What is the optimal route that they have to follow?



1: Village d'Astérix. 2 : Rotomagus (Rouen).
3 : Lutèce (Paris). 4 : Camaracum (Cambrai).
5 : Durocortorum (Reims). 6 : Divodurum (Metz).
7 : Lugdunum (Lyon). 8 : Nicae (Nice).
9 : Massilia (Marseille). 10 : Tolosa (Toulouse).
11 : Aginnum (Agen). 12 : Burdigala (Bordeaux).
13 : Gesocribate (Le Conquet).

Syllabus

1. Integer Programming. Total Unimodularity. Valid Inequalities and Preprocessing.
2. Solution Algorithms. Branch-and-Cut. Lagrangean Relaxation. Metaheuristics.
3. Matching Problems. The Assignment Problem.
4. Network Problems. Spanning Trees.
5. Covering Problems.
6. The Traveling Salesman Problem. Heuristics for the TSP.
7. Other Applications: Knapsack Problems, Scheduling Problems.

Required Knowledge
The course is mostly self-contained. Sometimes there will be references to notions from other courses (simplex method, continuous linear optimization, linear independence) and to user level knowledge of an optimization solver (e.g., Xpress).

Assessment
50% exam, 50% coursework. 2 assignments, each 25%, on Weeks 5 and 10 to be handed in one week later.

Operational Research in the Energy Industry

General Information

The course has two parts, Part-KM and Part-CD.

Part-KM focuses on optimization models and is taught by Ken McKinnon

Part-CD focuses on economics and reliability issues and is taught by Chris Dent.

Time:	Semester 2, 2023-24
Course organiser:	Chris Dent
E-mail:	`K.McKinnon@ed.ac.uk`
	`Chris.Dent@ed.ac.uk`

Assessment

Course Assessment

Full course: Assignments 35%, Exam 65%.
Part-KM: Assignments 35%, Exam 15%
Part-CD: Assignments 0%, Exam 50%

Assignments in Part-KM will involve building and solving optimization models.

- A suitable system for developing these models is mosel/Xpress
- If you already know a different optimization modelling language (e.g. AMPL, julia/JuMP, Python/Pyomo), that could be used instead.
- It will be assumed you are comfortable using some similar optimization system.

Prerequisites

MUST have passed:

Fundamentals of Operational Research (MATH10065) AND Fundamentals of Optimization (MATH11111) AND Methodology, Modelling and Consulting Skills (MATH11007)

OR

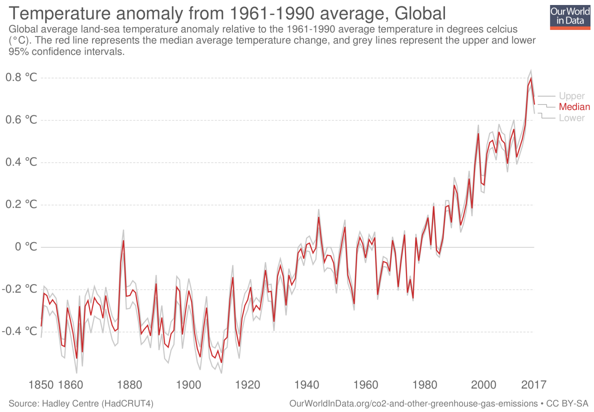
Linear Programming, Modelling and Solution (MATH10073)

Topics

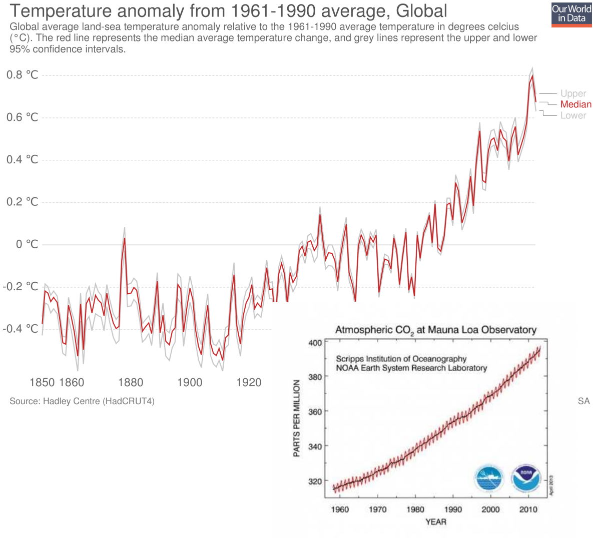
Core topics:

- Optimal power flow (finding the best output levels of electricity generators subject to network constraints)
- Unit commitment (scheduling start up and shut down of generators)
- Power system reliability (avoiding blackouts)
- Demand for Heat and electricity
- Power system investment planning
- Modelling of competition in markets
- Electricity market design

World Temperature



World Temperature and CO_2 Concentrations



Topics

Replace fossil fuelled generation with renewable generaton (wind, solar).

Opens up exciting New research areas:

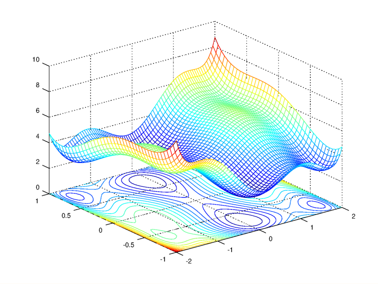
- Optimal power flow (finding the best output levels of electricity generators subject to network constraints) System inertia, mixed AC/DC transmission networks
- Unit commitment (scheduling start up and shut down of generators)
  Problem is much more stochastic because of unpredictability of renewables
- Power system reliability (avoiding blackouts)
  How to deal with intermittancy of renewables and lack of inertia
- Demand for heat and electricity Much more storage needed
- Power system investment planning Network infrastucture now much more important
- Modelling of competition in markets Many more actors ... prosumers
- Electricity market design Challenging to coordinate central and distributed planning

Texts for background reading

1. Wood and Wollenberg, ‘Power Generation, Operation and Control’ (optimization, mathematical aspects).
   Library has electronic version.
2. Lin and Magnago, ‘Electricity Markets: Theories and Applications’
   http://onlinelibrary.wiley.com/book/10.1002/9781119179382
   Free to download when on the University network
3. S. Stoft, ‘Power System Economics: Designing Markets for Electricity’, Wiley-Blackwell. ISBN 978-0471150404. (Online access through library).
4. Kirschen and Strbac, ‘Fundamental of Power System Economics’
   Wiley-Blackwell. ISBN 978-0470845721 (economics and markets, not too mathematical).
   Library has electronic version.
5. David JC MacKay, ’Sustainable Energy – without the hot air’, (good practical insights)
   https://www.withouthotair.com/download.html   Free online.

Nonlinear Optimization (NLO – MATH11194)

- Lecturer: Andreas Grothey
- 10 credits
- 18h lectures
  + 9 hours practical workshops (Python)
- 70% exam, 30% coursework (practicals)



An Optimization Problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in \Omega \subseteq \mathbb{R}^n. \end{array}$$

Set Ω is usually given by constraints, eg. $2x + 3y \leq 1, \quad x \cdot y = 1$

Nonlinear Optimization

- Objective function and/or constraints are nonlinear (beyond convex quadratic)
- Sometimes no constraints
  $\rightarrow$ Nonlinear Unconstrained Optimization

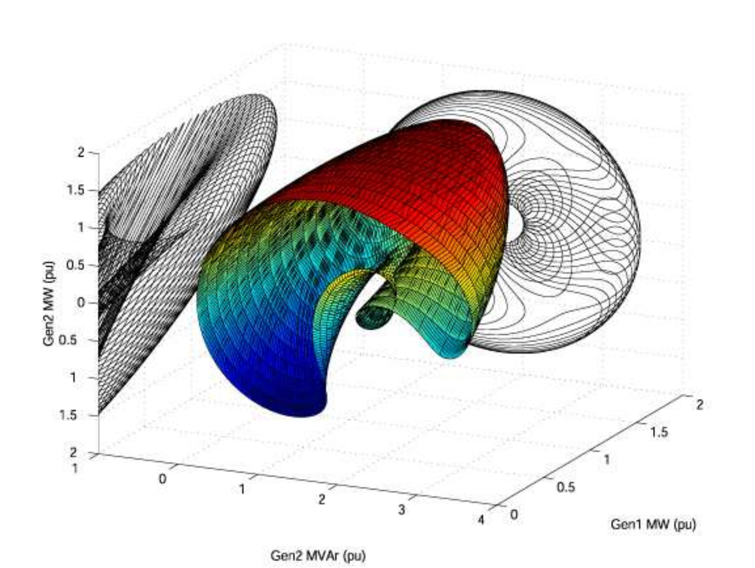
Applications of Nonlinear Optimization

Applications

- Engineering
  (often constrained by laws of physics which are nonlinear)
  - Power Systems Operation
  - Process Systems Engineering (mixing things)
  - Mechanical (Aerospace Engineering)
- Nonlinear Least Squares/Data Fitting/Parameter Estimation
- Machine Learning (training a neural network is really just solving a nonlinear optimization problem)
- Telecommunications

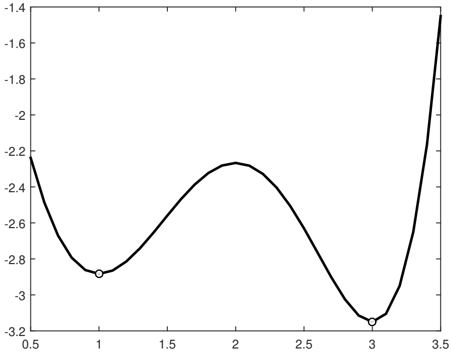
Often the real problem is nonlinear!
(people just solve a linear approximation because it is easier)

Optimal Power Flow: Feasible Region



Example 1.1: Local vs Global minima

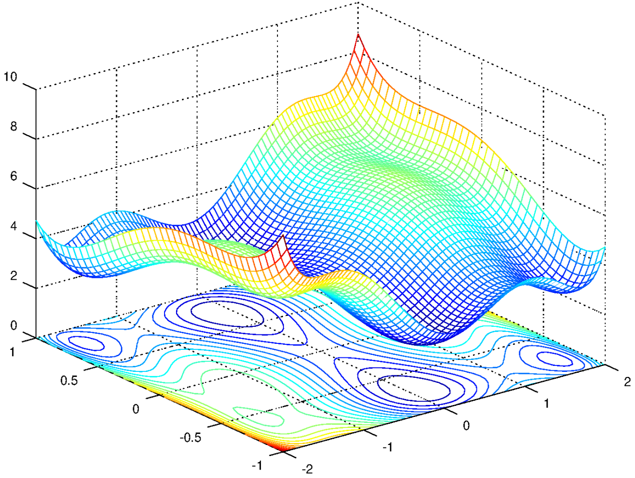
$$f(x) = \frac{1}{5}x^5 - \frac{5}{4}x^4 + \frac{5}{3}x^3 + \frac{5}{2}x^2 - 6x$$



- $x_1 = 1$ and $x_2 = 3$ are both local minima,
- but x_1 is not global.

Example 1.2: Local vs Global minima

$$f = x_1^2(4 - 2.1x_1^2 + \frac{1}{3}x_1^4) + x_1x_2 + x_2^2(-4 + 4x_2^2)$$



What does the course teach?

- Unconstrained and constrained nonlinear optimization
- Theory: Optimality conditions
- Algorithms: Often very geometrical and intuitive.
  - Line search
  - Trust region
  - Newton and Quasi-Newton Methods
  - Nonlinear Least Sqaures
  - Sequential Linear and Quadratic Programming
  - Penalty Methods/Augmented Lagrangians
- Specific Applications
- Using ready-made systems (NEOS, Modelling systems)

You have to see the algorithms in practice!

$\rightarrow$ course includes a large practical component where you will implement algorithms and gain practical experience with them.

Finance

MATH11158 – Optimization Methods in Finance

http://www.drps.ed.ac.uk/23-24/dpt/cxmath11158.htm

X THE UNIVERSITY OF STRATHCLYDE

FINANCIAL RISK AND UNCERTAINTY

SEMESTER 1

PROFESSOR ANDREW MARSHALL

Course Lecturer

Welcome to the course Financial Risk and Uncertainty (FRU).

My name is Andrew Marshall and I am a Professor at the University of Strathclyde in Glasgow.

https://www.strath.ac.uk/staff/marshallandrewprof/

Although I am an external lecturer I have designed and taught on FRU at Edinburgh for over 10 years.

I hope you enjoy this class.

If you have any questions, please do not hesitate to email me at: a.marshall@strath.ac.uk

CLASS CONTENT

Increased volatility in foreign exchange markets is one consequence of record interest rate hikes, geopolitical tensions and sustained inflationary pressures, among other global challenges. What can banks and companies do to stay ahead of the game?

Course Content

1. Introduction to derivatives and financial risk management. Why hedge?
2. Focus on currency and interest rate risk management, also consider commodity risk
3. Forwards/Futures and analysis of derivatives
4. Options
5. Swaps
6. Use of case studies to identify and manage financial risk management

Assessment

Individual Assignment

30%

Submitted Jan 2024

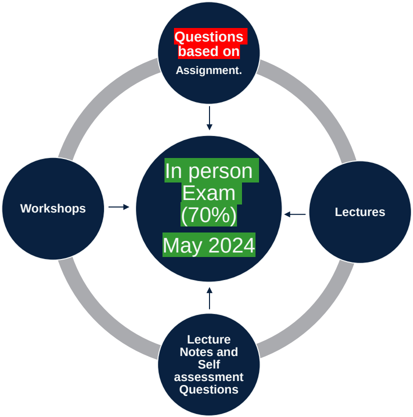
Based on

Lectures

Lecture Notes

Workshops

Independent Reading



Course Material

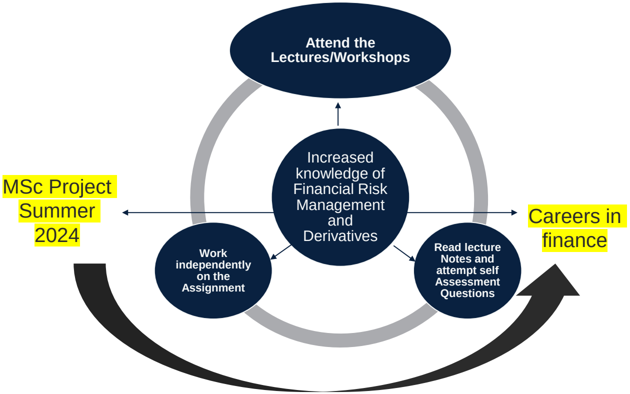
1. Full set of Lecture notes will be posted on Learn.
2. Workshop Questions will be posted on Learn.
3. Textbook Recommendations (No need to purchase but useful to consult via Library for independent reading)

J C Hull. "Options, Futures and Other Derivatives". Prentice Hall, International Inc. 9th Ed. 2017, ISBN13: 9781292212890 (Note previous/future editions are acceptable)

D. A. Dubofsky and T. W. Miller
"Derivatives: Valuation and Risk Management" Oxford Press. ISBN: 9780195114706

Don M. Chance and Robert Brooks "Introduction to Derivatives and Risk Management, Cengage 10th Edition, 2016, ISBN13: 9781305104969.

To succeed in FRU

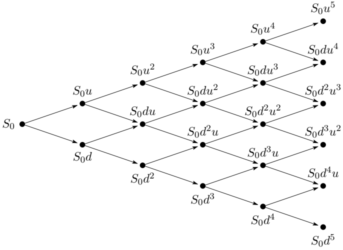


Discrete-Time Finance

Tibor.Antal@ed.ac.uk

Semester 1, Level 11, Two lectures per week, 5 workshops, 5% coursework from 5 assignments

Contents



Bird's eye view on “Stochastic Analysis in Finance”

István Gyöngy

School of Mathematics, University of Edinburgh, JCMB 5612

September 13, 2023

Course content: consists of units I, II and III

Unit III. Black-Scholes Market with two assets

- Riskless asset (bond): B_t, the bond price at t satisfies

$$dB_t = rB_t\,dt, \quad B_0 = 1, \quad \text{i.e., } B_t = e^{rt},$$

  where r is a non-negative constant;

- Risky asset (stock, share): S_t the stock price at t, satisfies a stochastic differential equation (SDE),

$$dS_t = \mu S_t dt + \sigma S_t\,dW_t,$$

  where μ and $\sigma > 0$ (called 'volatility') are constants, and $W = (W_t)_{t\geq 0}$ is a Wiener process.

Option Pricing

- European type options: An EU-type option is a contract according to which the holder of the option pays $g(S_T)$ at a given time T in the future, for a transaction of the risky asset at T, where g is a given function, agreed in the contract.
- American type options: the holder of the option is allowed to make the transaction at any time τ, during the time $[0, T]$, taking into account the the stock prices until τ.
- Exotic options

Main questions: What is the 'fair price' of an option, and how to calculate it for the above types of options?

Unit II. Basics of Stochastic Analysis

- Random processes: Wiener process, martingales;
- Stochastic Itô integrals, Itô processes, stopping times;
- Stochastic chain rule for Itô processes ('Itô formula');
- Stochastic differential equations (SDEs);
- Change of probability measures ('Girsanov theorem');
- Martingale representation theorems;
- SDEs and PDEs, stochastic representation of solutions to PDEs;
- Kolmogorov forward and backward equations.

Unit I. Basics of Probability Theory

- Probability space, random variables (RVs), the law and distribution function of RVs;
- Integral of measurable functions-Expectation of RVs, L_p-spaces;
- Convergence of random RVs and their expectations;
- Conditional expectation;
- Markov-Chebyshev, Hölder and Jensen inequalities;
- Absolute continuity of probability measures.

Simulation
General Information About the Course

Burak Büke

School of Mathematics
University of Edinburgh

January 17th, 2023

Marking Structure

- Continuous Assessment (50% of overall mark)
  - 3 Problem Sets (5 % each)
  - Course Project (35% of overall mark)
- Exam (50% of overall mark)
  - At the end of semester

Course Outline

- Discrete-Event simulation
- Random Number Generation
- Input Analysis
- Output Analysis
- Variance Reduction
- Simul8 and Arena (Simulation Packages)

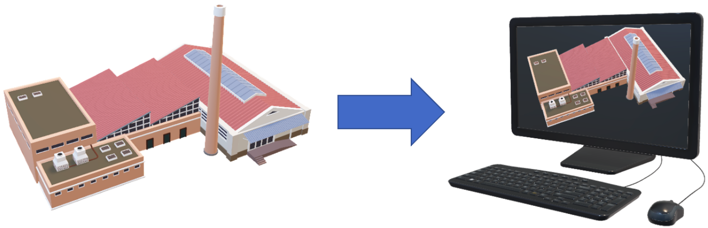
- To capture real world randomness, we need basic tools from probability and statistics

Simulation Software - Simul8 and Arena

- Alternative softwares, both run on Windows machines only
  - Mac users need to install a virtual machine (e.g. VirtualBox) that runs Windows
- Simul8 can be accessed online and hence will be our main tool.
- Instructions for download will be uploaded on Learn for both.

What is Simulation?

- **Definition:** Simulation is the "imitation of a real system".
  - Discrete-event system simulation will be covered in this class.



Why Simulate?

- Why do we use simulation?
    - Gain insight into our problem
    - Characterize the performance of the real system when analytical models are unavailale
    - Design and experiment with new system without actually changing the system
    - Train the users of a system without actually using the real system
    - Validate analytical results
    - Sell our results better using animation

MATH11132 – Financial Risk Theory

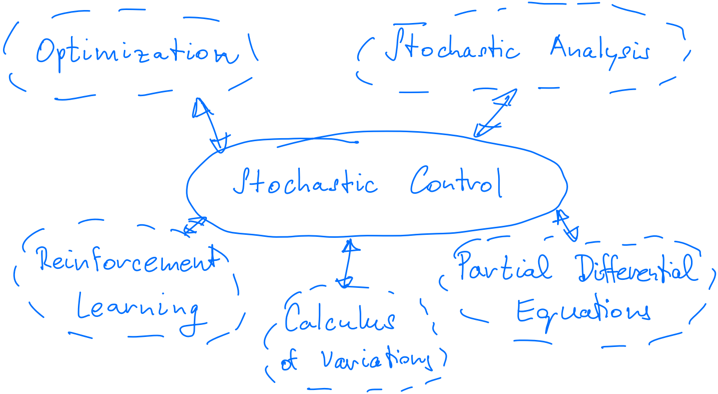
http://www.drps.ed.ac.uk/23-24/dpt/cxmath11132.htm

MATH11148 – Credit Scoring

http://www.drps.ed.ac.uk/23-24/dpt/cxmath11148.htm

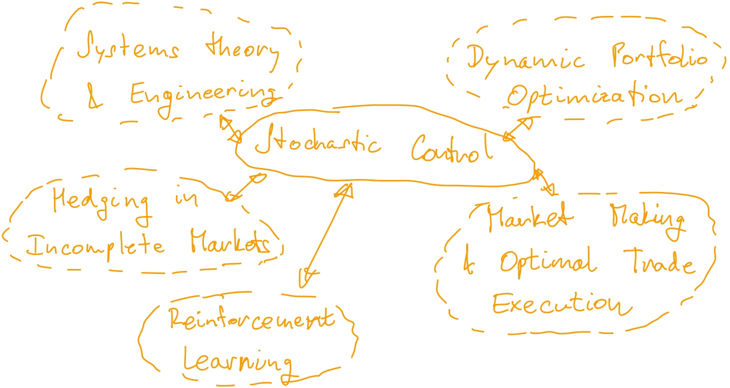
SCDAA - Stochastic Control and Dynamic Asset Allocation

Connections other topics in mathematics



SCDAA - Stochastic Control and Dynamic Asset Allocation

Applications



Prototypical problem

$$\begin{cases} \text{maximize } \ J(t,x,\alpha) = \mathbb{E}\Big[\displaystyle\int_t^T f\big(s, X_s^{\alpha,t,x}, \alpha_s\big)\, ds + g\big(X_T^{\alpha,t,x}\big) \Big| \alpha \Big] \ \text{ over } \ \alpha\,, \\ \text{subject to } \ dX_s^{\alpha,t,x} = b\big(s, X_s^{\alpha,t,x}, \alpha_s\big)\, ds + \sigma\big(s, X_s^{\alpha,t,x}, \alpha_s\big)\, dW_s, \quad X_t = x. \end{cases}$$

SCDAA - Stochastic Control and Dynamic Asset Allocation

Course structure:

- Prerequisites: SAF or PMF or equivalent, Python.
- Weekly 2h lecture, bi-weekly 1h workshop.
- 80% exam in May / June.
- 20% coursework; In the past this was group task, Python implementation of an algorithm solving some control problem running W6-W12 of S2.

Main sources:

- Lecture notes: `https://www.maths.ed.ac.uk/~dsiska/LecNotesSCDAA.pdf`.
- H. Pham. *Continuous-time stochastic control and optimization with financial applications*. Springer, 2009.
- A. Cartea, S. Jaimungal, and J. Penalva. *Algorithmic and High-Frequency Trading*. Cambridge University Press, 2015.
- N. Touzi. *Optimal stochastic control, stochastic target problems and backward SDE* `http://www.cmap.polytechnique.fr/~touzi/Fields-LN.pdf`.

MATH11202 – Numerical Probability and Monte Carlo

http://www.drps.ed.ac.uk/23-24/dpt/cxmath11202.htm

Computational and Applied Mathematics

Python Programming (MATH11199)

Dr Charlotte Desvages

Course summary

- First: **programming fundamentals** using **Python**.
- Then, practical applications in **data manipulation and visualisation**.

No previous programming experience is required!

Students will engage with **professional programming workflows and tools**, and will have the opportunity to **collaborate** with peers to develop their skills.

Learning outcomes

On successful completion of this course, you should be able to:

- Design and implement Python programs to solve a range of mathematical problems.
- Select and use appropriate libraries and data structures to perform computational analyses in Python; consult the relevant documentation.
- Review a Python program to explain the underlying logic, identify and fix bugs, and suggest improvements to structure and style.
- Collaborate with peers on programming tasks, using suitable tools.
- Use Python to carry out investigations on data and extract key insights; display and discuss results in a well-presented report.

Course structure

A typical week:

- In your own time: watch 1-2 short **videos**, and complete a **tutorial notebook** with interactive examples and exercises.
- **Lecture**: Mondays 1.10-2pm. Code-along examples reviewing the previous week. Bring your laptop!
- **Workshop**: Thursdays or Fridays (depending on your group). Work on a programming task in groups of 2-3.
- Short **weekly homework**: either quiz (autograded), or peer-assessed code review.

All materials will be released on the Monday of each week.

Assessment

You will have a short piece of homework every week, and two larger projects (one individual, one in groups). There is no exam.

- **4 quizzes**: best 3 out of 4 count for a total of 15% (5% per quiz).
- **4 code reviews**: best 3 out of 4 count for a total of 5% (1.67% per code review).
- **Project 1**: done individually, halfway through the semester, counts for 40%.
- **Project 2**: done in small teams during workshops starting Week 8, counts for 40%. Extract some interesting information from a given dataset, and produce a report with carefully designed data visualisations.

Contacts

- Course Organiser and lecturer: Dr Charlotte Desvages (`charlotte.desvages@ed.ac.uk`)
- Course Secretary: Ms Gemma Aitchison (`gemma.aitchison@ed.ac.uk`)

Industrial Mathematics (MATH11231)

Dr Ben Goddard
b.goddard@ed.ac.uk

- Semester 1
- Level 11, 10 Credits
- Aimed primarily at CAM MSc students, but open to others
- Prerequisites: differential equations, linear algebra, probability, programming

Industrial Mathematics (MATH11231)

Dr Ben Goddard

b.goddard@ed.ac.uk

- Semester 1
- Level 11, 10 Credits
- Aimed primarily at CAM MSc students, but open to others
- Prerequisites: differential equations, linear algebra, probability, programming

- Focuses on problems which come from industry, aiming for relevant and applicable solutions
- Requires analytical and problem-solving skills
- Includes applied mathematics, computing, statistics, data science, . . .

Industrial Mathematics (MATH11231)

Dr Ben Goddard
b.goddard@ed.ac.uk

- Weeks 1–3 introduce necessary mathematics and skills
- Weeks 4–7 Project 1 (advection-diffusion models)
- Weeks 8–11 Project 2 (agent-based models)

Industrial Mathematics (MATH11231)

Dr Ben Goddard
`b.goddard@ed.ac.uk`

- Weeks 1–3 introduce necessary mathematics and skills
- Weeks 4–7 Project 1 (advection-diffusion models)
- Weeks 8–11 Project 2 (agent-based models)

- Very few traditional lectures;
  lots of interaction including industry guests
- 2-hour workshops each week;
  predominantly to work on group projects
- Assessment is through two 5-page group projects, 50% each

MATH11240: Numerical methods for Data

Desmond J. Higham
Konstantinos C. Zygalakis

School of Mathematics, University of Edinburgh

Course team

Prof. Des Higham

Prof. Konstantinos Zygalakis

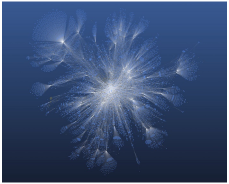
Outline

1. What is this course about?
   - Networks
   - Inverse problems
   - Deep learning

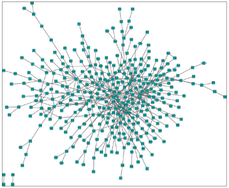
Networks

We aim to understand the connections between objects

1. Do the proteins split naturally into fairly distinct, well-connected, groups?
2. Which parts of the internet are most vulnerable to attack?



webpages network

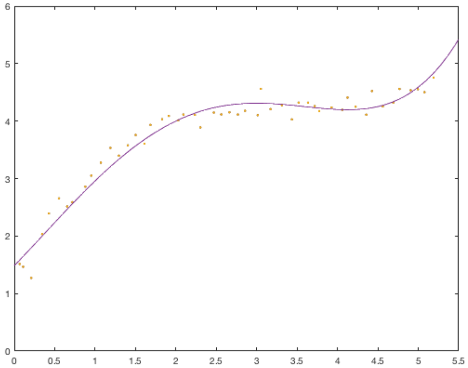


protein network

Inverse problems

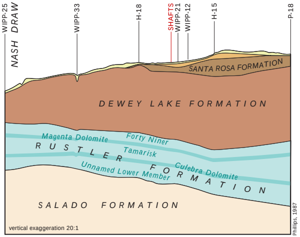
- An inverse problem is concerned with determining causal factors from observed data.
- In mathematical terms, we want to determine system inputs based on (partial and noisy) observations of system outputs.
- Inverse problems appear in many different areas!

Example 1: The regression problem



Goal: reconstruct a function f given noisy point values $\{f(u_i)\}$

Example 2: The diffusion problem



Cross-section at WIPP

Goal: reconstruct the hydraulic conductivity k of the subsurface given noisy measurements of the water pressure $\{p(u_i)\}$.

Example 3: Computational Imaging

Goal: reconstruct an image x given a noisy, partial observation y

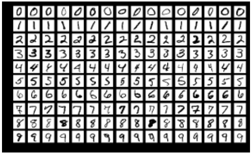
Original

Blurred

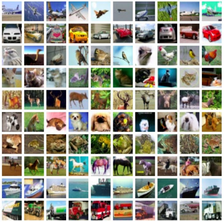
Deep Learning

Many many applications:

1. Classification
   - Given a set of labeled images
   - Put labels on unseen images
2. Image colouring
3. Music composition
4. Natural Language Processing



MNIST data set



CIFAR10 data set

What do these problems have in common?

1. A lot of linear algebra involved in the mathematical formulation of these problems.
   - Eigenvalue decomposition
   - Singular value decomposition
2. To solve the problems discussed above we need to solve an optimization problem of the form

   $$\min f(x),$$

   where f is a continuously differentiable function, sometimes very complicated (deep learning).
   - The chain rule to understand backpropagation.
   - Optimization algorithms such as gradient descent.

Assessment

- Exam 50%
- Coursework 50%
  1. 3 hand in assignments (pen and paper) each worth 10%
  2. 2 computer assessments each worth 10%

Research Skills in Computational Applied Mathematics (RSCAM)

- Mandatory for students on CAM MSc (and only open to them)
- Offers training in programming, numerics, writing, presentation, and group work.
- Credit also awarded for attending seminars (8 required/semester)
- Exercises in Python, and, in 2023-4 Julia and Parallel Computing
- First meeting - Wednesday, 9am-10:50am in JCMB 6206
- Workshops - Fridays in even weeks.