Study guide for MMA291 Project in Mathematics Fall 2014

Course homepage:

On the course homepage you can find the study guide and relevant links. Project descriptions will also be posted there.
`http://www.mdh.se/amnen/matematik/kurser/kurshemsidor/mma291`

Objectives:

The course shall:

- Give a deeper knowledge of a mathematical subject
- Develop the ability to formulate a problem and to work independently
- Develop the ability to present knowledge and the results obtained in speech and writing.

Examination:

PRO1, 7.5 credits, marks Pass (G) or Pass with distinction (VG), Project with oral presentation.

The project will be done in pairs (if possible) and each pair will choose project topic and write, in cooperation with the supervisor, a description of their project which describes the chosen topic and the goals of the project. The description should include a descriptive headline for the project (not necessarily final), a description of the topic with some concrete goals and at least one source that can be used for the project. An example of a project description can be found on the course homepage.
The focus of the project should be mathematical but can be mathematical in different ways. It can consist of

- *mathematical modeling*, that is looking at ways to interpret theories mathematically to build mathematical models or comparing different mathematical models,
- *mathematical problem solving*, where a specific problem and ways of solving it are described in detail are examined or
- *mathematical computation*, where a model or mathematical technique is used to analyze a set of data, focusing on the mathematical method / model, it's advantages, disadvantages and implementation.

Naturally these different areas can also be combined.

Literature:

The literature depends on the chosen topic. Each group or students supervisor may supply the students with some appropriate literature but searching for relevant information on your own is part of the course.

Instruction:

After choosing topic the student is assigned a supervisor. Continuous supervision of the project is mandatory. The student and supervisor should agree on how this supervision should be handled.

Oral presentation

The oral presentation of the project will be 30 minutes long plus some time for discussion and answering questions. The date for it will be decided when a sufficient amount of progress has been made on the project.

Written report

Sources and citing

All sources for the project must be clearly written in the report. If text is copied from any source then it should be given as a quotation and be appropriately attributed. The exception to this is definitions and theorems. Definitions of common and well-known concepts can be given without a source, as can well-known 'standard' theorems. Less well-known definitions and theorems should be attributed but should not be given as quotes. For examples of one way to handle citations see the topic suggestions.
Wikipedia should not be used as a source! It is often a good idea to look at the sources given in the wikipedia article in order to find good sources though. The same goes for most encyclopedias and reference works.

Report formatting

The length of the report depends on the project and is decided upon together with the supervisor. The report should be structured clearly and have a table of contents. All tables and figures should be numbered and given a descriptive caption. Important equations and formulas should be numbered.

LaTeX

The report should be computer written using the LaTeX typesetting software. LaTeX is a free and open source software available for most platforms that is designed to make it easy to write reports, articles and books of a technical nature. The program can be tricky to learn but many journals an conferences require submissions to be written using LaTeX and the program is very powerful in the hands of an experienced user.
On the course homepage you can find links to all the material listed here along with a LaTeX-template for your report (if you prefer you can use another template than the one on the homepage).
Here are a couple of short guides to starting with LaTeX:

LaTeX Tutorial, Jeff Clark
`http://ece.uprm.edu/~caceros/latex/introduction.pdf`

Att skriva rapporter med latex, Per Foreby (in Swedish)
`http://www.ddg.lth.se/perf/handledning/`

LaTeX for Word Processor Users, Guido Gonzato
`http://www.tex.ac.uk/tex-archive/info/latex4wp/latex4wp.pdf`

These guides are somewhat outdated but are all fairly short and easy to understand for beginners.

If you want more comprehensive sources for LaTeX you can check out:

The Not So Short Introduction To LaTeX 2_ε, Tobias Oetiker
`http://tobi.oetiker.ch/lshort/lshort.pdf`

A Simplified Introduction to LaTeX, Harvey J. Greenberg
`http://csrg.inf.utfsm.cl/~rtobar/download/manual-latex.pdf`

If you have a Windows computer the easiest way to get started with LaTeX is to use MiKTeX.
For MAC OS users the simplest option is probably MacTeX.
For Linux users the easiest way to install LaTeX is to use the package manager for your distribution (e.g. the Ubuntu software center) and install a LaTeX editor. A common simple one is TeXmaker. The package manager will then install any extra files you need for LaTeX.
There are many editor programs that you can download to make it easier to work with LaTeX. MiKTeX comes with TeXworks and MacTeX comes with TeXShop. These are relatively simple programs but there are many alternatives that have more features. Some common ones are

- **TeXnicCenter** is an editor for Windows with many features, there is a short guide (in Swedish) with pictures showing how to get started with this program on the course homepage.
- **TeXstudio** is a similar editor that is available for both Windows, Mac and Linux.
- **LyX** is an editor that is a little more similar to standard word processors such as Microsoft Word or Open Office.

All this information about LaTeX is also available on the course homepage.

Topic suggestions

Basic topics

Modeling interest rate and calculating yield curves

Loans are a natural part of modern business, either by taking out loans in order to invest or by lending money to others for profit. A natural consequence of this is that interest rates become very important and having the ability to predict and describe interest rates efficiently and accurately can be very useful. There are many different models to describe how interest rate varies over time in different systems. A possible project would be to analyze one or more of these models. One example could be the Vasiček model which was an early model for interest rates from the 1970's which was followed by many other models. Possible versions of this project include:

- Describing the origin and formulation of the Vasiček, describe some of its strengths and weaknesses and / or compare to some other models.
- Describing the Vasiček model and show how various problems that involves the model can be solved analytically, for example yield curve prediction.
- Using a real data set and make predictions using the Vasiček model, either by using numerical techniques or analytical results. The predictions can then be compared to real data in order to evaluate the quality of the Vasiček model in the given situation.

Examples of interest rate models and some results can be found in [1] and some interesting mathematical results regarding the Vasiček model can be found in [2].

Arnold Tongues

Mode locking, also called *phase locking*, is an interesting phenomenon that occurs when you have two dynamical systems that interact weakly with each other. If the interaction between the systems is of the right kind the systems can synchronize in unexpected ways. An illustrative example of this would be two pendulum clocks that hang next two each other on a wall. If the clocks are started with their pendulums in the opposite direction the pendulums will after some time (usually a few days) swing in perfect sync. Mode-locking is also the reason that there is a light and dark side of the moon. One way to describe and analyze mode-locking is to use *Arnold tongues* which are visualizations of how varying some parameter in the interacting systems affects the (possible) mode-locking. A suitable project could be to study the model and theory presented in [3] and try to replicate their results using numerical or analytical methods.

Identifying copied DNA sequences

For many cancer types there are extra copies of DNA in the tumor cells. In a normal cell each DNA segments is expected to have two copies. Measuring the number of copies can be difficult though as not all DNA in each cell can be measured perfectly. In this project you can explore ways to find a good way of identifying segments that have been copied using statistical models and dynamic programming on a data set. The analysis is most suitably made using a computer. For sources see: [4], [5] and [6]. They can all be found using the databases you can access through the MDH library homepage.

Advanced topics

The Four- and Eight-squares Theorems, and Vector Products

Given two pairs (a, b) and (c, d) of integers, it is always possible to find integers u,v such that $u^2+v^2 = (a^2+b^2)(c^2+d^2)$, and (u, v) depends linearly on (a, b) and (c, d). This result, known as the two-squares theorem, follows naturally from the properties of multiplication of complex numbers. It turns out that this formula has analogues for pairs of n-tuples of integers only when n is 4 or 8, and that the formulae in each case are given by higher-dimensional analogues of the complex numbers, known as the quaternions and the octonions, respectively.
The aim of the proposed project is to gain a deeper understanding of this result, and its connection to the geometry of space. The result for four-tuples, the so-called four-squares theorem, is closely related to the vector product in three-dimensional space, and similarly, the eight-squares theorem is tied to the existence of a vector product in seven-dimensional space. It would be particularly interesting to see if the four-squares theorem could be interpreted geometrically in three-dimensional space, if the formula, literally, could be 'seen'.
Questions/lines of inquiry:

- Clarify why the n-squares formulae only exist for $n = 2, 4, 8$;
- investigate the connection between the formulae and the vector products in four- and eight-dimensional space;
- find geometric interpretations of the formulae, particularly for $n = 4$.

Useful literature: [7] and [8].

The Algebra of Wireless Communication

Space-time codes are codes adapted for wireless transmission with multiple transmitting and receiving aerials. Such codes need to be robust not only against random noise disturbances, but also against interference between signals sent from different aerials. In recent years, an intimate connection has been found between space-time codes and abstract algebra; it turns out that the so-called division algebras provide a rich and natural source of codes of high accuracy and performance.
In this project, we will examine the mathematics behind the construction of space-time codes. This includes an overview of research in the field in recent years, and possibly also exploring some new directions of research based on a generalised concept of a division algebra, which seems well suited for the demands of code construction.

- Understand the concept of a space-times code, and the criteria that make them suitable for MIMO (multiple-input multiple-output) transmission;
- understand the algebraic structures used to construct space-time codes: algebraic number fields, associative and non-associative division algebras;
- study how associative division algebras can be used to construct perfect space-time codes;
- explore ways in which space-time codes can be obtained from non-associative division algebras.

Useful literature: [9], [10] and [11].

Real Division Algebras

The integers are included in the rational numbers, which embed in the reals, and the reals may be extended to the complex numbers... but what happens next? Are there some 'hyper-complex' numbers of which the ordinary complex numbers are a subset? The answer may be either positive or negative, depending on whom you ask, and both viewpoints have their merit. While strong case can be made for the idea that the complex numbers are the 'crown' of the number system, it is equally true that there are several natural ways to view the complex numbers as sitting inside some higher-dimensional system of 'numbers', and that these structures both natural and useful in mathematics and physics.
The idea of this project is to make acquaintance with the classical instances of 'hyper-complex numbers', that is the quaternion and octonion algebras, and then choose a direction of further inquiry. This may be the classification theorems for associative and alternative real division algebras (which, simply speaking, tell you why and in which respect the quaternions and octonions are the 'best' hyper-complex numbers), study of more general classes of division algebras, or applications to other fields (e.g. the use of quaternions in computer animation).
Questions/lines of inquiry:
Beyond the understanding of the basic concepts and theory of real division algebras, some possible themes are

- the classification theorems by Frobenius and Zorn, of associative respectively alternative real division algebras;
- division algebras over other fields, the Brauer group;
- general division algebras over the real numbers;
- quaternions in computer animation.

Useful literature: [7], [12].

The Escalator Boxcar Train

When studying the dynamics of the populations in ecological systems one realizes that the states of individuals, such as age, size and so on may change the reproduction rate, growth rate and death rate. Thus we study physiologically structured population models (PSPM). There are several ways to do this numerically and some analytical example exists. One popular method, the Escalator Boxcar Train (EBT), was introduced by de Roos in [13]. Although used extensively, the EBT method has just recently been proven to converge. There are several possible variants and extensions of the EBT method (e.g. [14], [15]). The following projects concern this subject.

ETB Project 1; Convergence rate of the EBT method

Objective:
In the EBT method one has to introduce boundary cohorts due to the new-born individuals in the population. The dynamics of these boundary cohorts has been defined differently in the literature. The question here would be to compare the convergence rate of different boundary cohorts in the EBT-method.
Task:
Construct a solver in MATLAB for the EBT-Method and compare the rate of convergence of the methods proposed in [13] and [15] and other similar boundary cohort definitions.

Difficulty (7.5hp):
A challenge for one student with sufficient MATLAB knowledge. At most two students can do this project as a group.

ETB Project 2; Stability of merge and splitting operators in the EBT-Method

Objective:
Numerical stability of the EBT-method when using splitting and merging operators. When the cohorts follow the paths of the growth rate, their distance will sometimes become very small, in which case a merging operator is useful to hold down the number of cohorts in the solver. On the other hand, if the distance grew too large, the accuracy is lost, so one would like to split two cohorts into three. Can this be done in a numerical stable way?
Task:
Construct a solver in MATLAB for the EBT-Method where merge and splitting operator functions are defined in a mathematically correct way and then test the model. This project will be in collaboration with Linus Carlsson as a supervisor.
Difficulty (7.5hp):
A challenge for two students with sufficient mathematical and MATLAB knowledge. At most two students can do this project as a group.

ETB Project 3; EBT-method for cells

Objective:
In the EBT-method the population give birth to individuals of a certain size, which is the minimal size of individuals in the population. Can the EBT-method be altered to deal with cells, which by nature divide (we will probably assume binary fission)?
Task:
Investigate if we can use the EBT-method for cells? Implement the solver in MATLAB.
Difficulty: (7.5hp)
Hard/impossible? The student will need advanced knowledge in biology, numerical methods, mathematics and implementation.

More OpenMath symbols

OpenMath (`www.openmath.org`) is a standard for encoding mathematical content in computer-readable form (related to MathML, but less cluttered). The standard as such codifies what a fully formalised mathematical formula is, but the codification of a detailed mathematical vocabulary (concepts such as collinearity, continuity, Legendre polynomials, Petersen graph, and even such fundamental things as logical 'and' and addition of numbers) is delegated to a separate kind of document called a Content Dictionary. For example, ordinary addition is the symbol 'plus' in the official content dictionary 'arith1', whereas collinearity is the symbol 'are_on_line' in the experimental content dictionary 'plangeo1'.
Elementary mathematics is pretty well covered by the existing content dictionaries, but as one approaches the advanced/master level, coverage gets quite spotty. In principle this is not a problem, since anyone may author a new content dictionary, but in practice it presents a considerable obstacle to a researcher if one has to create half a dozen documents defining perfectly well-known things (simply because no-one else has bothered to formally define those yet) before it becomes possible to state something about the problems one is really working on. As things currently stand, every little piece helps.

In this project, you would: First pick an area of mathematics with which you have some familiarity and create a content dictionary for some piece of it. (Obviously, you should not repeat what someone else has already defined. If you pick an area which already has partial cover, then you must go further than previous authors have.) And second demonstrate the viability of your dictionary by using it to state, as an OpenMath object, some nontrivial mathematical theorems or definitions. (Note that this does not ask you to give proofs of them; doing also that would be serious overkill for a project this size.)
If the result turns out well, the prospective student is recommended to publish the content dictionary by submitting it to `www.openmath.org` (where it would appear as a contributed content dictionary).
Caveat:
This project will require you to exercise some passing knowledge of XML. If you already have that, then good. If you don't, you would have to read up on that subject (in addition to OpenMath itself). For some mathematicians, this is a straightforward exercise of basic logic, whereas for others it may remain a mysterious skill belonging to the foreign field of computer science; success with previous studies in mathematics is not a good indicator as to where between those two extremes you would fall. Therefore take time to carefully consider where your aptitude lies before committing to this project.
Literature/resources:
Obviously, some textbook on the area of mathematics you pick as well as the OpenMath standard [16] and maybe a chapter or two of [17].

Network order, polynomial inequalities, and genetic algorithms

A certain problem of some research interest (constructing a "network order" with certain properties, using a particular construction) leads to the problem of finding a nontrivial solution to a certain system of polynomial inequalities: there are 16 variables (one for each element in a 4x4 matrix) and a couple of dozen inequalities. There are exact algorithms which in principle could solve the system, but unfortunately this problem has too many variables for those to be practical. Instead this project is to use genetic algorithms to try to numerically evolve solutions: an individual would be an assignment of values (nonnegative real values) to the variables, and the fitness of an individual would be to how well it satisfies all the inequalities.
A good starting point would be [18].

References

[1] *Yield curve estimation and prediction with Vasiček model*, Bayazit, D., Middle East Technical University, 2004

[2] *On the Vandermonde matrix and its role in mathematical finance*, Norberg, R., University of Copenhagen, Laboratory of Actuarial Mathematics, 1999

[3] *Arnold tongues in human cardiorespiratory systems*, McGuinness, M. Hong, Y., Galletly, D., Larsen, P., *Chaos* March 2004, Vol. 14 Issue 1 page 1-6

[4] *The array CGH and its clinical applications*, Shinawi, M. , Cheung, S.W., Drug Discovery Today, September 2008, Vol. 13 Issue 17-18 page 760-770

[5] *A statistical approach for array CGH data analysis*, Picard, F. *et al*, BMC Bioinformatics, February 2005, Vol. 6

[6] *Algorithms for the optimal identification of segment neighborhoods*, Auger, I.E., Lawrence C.E., Bulletin of Mathematical Biology 1989, Vol. 51 Issue 1 page 39-54

[7] *Numbers*, H.-D. Ebbinghaus *et al.*, Graduate Texts in Mathematics, Vol. 123, Springer-Verlag, New York, 1991.

[8] *Octonions, Jordan Algebras and Exceptional Groups*, T. A. Springer and F. D. Veldkamp, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000.

[9] *An introduction to central simple algebras and their applications to wireless communication*, G. Berhuy, F. Oggier, Mathematical Surveys and Monographs, 191. American Mathematical Society, Providence, RI, 2013.

[10] *Space-time block codes from nonassociative division algebras*, S. Pumplün and T. Unger, Adv. Math. Commun. 5 (2011), no. 3, 449–471.

[11] *Division algebras and wireless communication*, B. A. Sethuraman, Notices Amer. Math. Soc. 57 (2010), no. 11, 1432–1439.

[12] *An introduction to nonassociative algebras*, R. D. Schafer, Corrected reprint of the 1966 original, Dover Publications, Inc., New York, 1995. x+166 pp.

[13] *Numerical methods for structured population models: the escalator boxcar train*, A. M. de Roos, Numer. Methods Partial Differential Equations, 4(3):173-195, 1988.

[14] Brännström *et al*, arXiv preprint 1210.1444.

[15] Gwiazda *et al*, arXiv preprint 1309.2408.

[16] *The OpenMath Standard 2.0:*
`http://www.openmath.org/standard/om20-2004-06-30/omstd20.pdf`

[17] *Beginning XML:*
`http://ep.bib.mdh.se/login?url=http://site.ebrary.com/lib/malardalen/Doc?id=10278318`

[18] *Genetic algorithm*
`http://en.wikipedia.org/wiki/Genetic_algorithm`