SAM GRADUATE PROJECTS (2008-2009)

Abstract. This note provides description of graduation projects for SAM graduates in Computational Mathematics and Statistics direction at HCMC's US and UT- VNU during 2008-2009.

1. Inventory Control in Logistics

Current need for Logistics Optimization. Inventory management is one of the oldest and most studied areas of *Operation Research*, in particular, it is useful for *Logistics Industry* at developed nations but less natural resources. Logistics plays a key role in Import and Export activities, as witnessed by the cargo stagnancy in ports of HCMC and Southern Vietnam recently.

Steps of the investigation.

A. Pre-graduate research period:

a/ Learn basic concepts of Economic Order Quantity (ECQ) and Economic Production Quantity (EPQ) of the problem. Take the course *Stochastic Processes* at AM section, HCMUT.

b/ **Algebraically model the simple case**: Predict uncertain demand in a single period At the same time, visit a manufacturing or sale firm in HCMC, for industrial internship, to see how inventory matter in those sectors, in 2-4 weeks.

c/ Design an application package that allows manuauver inventory events. Study a computing machinary such as `Singular`, `Matlab`, `R` or `OpenModelica`.

B. Graduate research period:

d/ Investigate more complex cases in Inventory Control

e/ Learn approriate algorithms and implement them

f/ Implement a pilot application package with basic functionality allowing prediction of small cases.

Date: **August 29, 2008**

e-mail: `mnguyen@cse.hcmut.edu.vn` .

2. Blind channel Identification with Algebraic Statistics: A case study

Graduate: **Tran Vinh Tan**.

Keywords. computational algebraic geometry, rational univariate representation (RUR), discrete-time stochastic processes, regression, digital communications, signal processing

Introduction & Aim. We employ Symbolic computation and Statistics methods in some problems of advanced *filter design* for signal processing, more precisely the construction of wavelet filters for which the usual *spectral factorization* approach is not applicable.

Our aim is to show how the design equations can be written as multivariate polynomial systems of equations and accordingly how Groebner algorithms offer an effective way to obtain solutions of practical interest in many of these cases.
(Read few reference articles I enclose with this).

Basic terms. These examples of multiwavelet bases and wavelet frames could not have been obtained without the use of these new tools from Algebraic Geometry. Although their high computational and memory costs, Groebner bases are indeed effective tools for the theoretical study and practical design of filter banks.

- Single-input single-output =SISO; Single Input Multiple Output = SIMO
- Frequence response function = FRF; Finite impulse response = FIR
- The Rational Fraction Polynomial Method (RFP) is a SISO method that first appeared in 1982 [Richardson and Formenti]. This method was extended to analyze globally a set of FRFs, using one single input reference and multiple outputs making it a SIMO method. This extension is called the Global Rational Fraction Polynomial Method (GRFP).
- Blind channel identification means identification without the exact knowledge of any input symbols. Blind identification or equalization is not a new subject, for it has been addressed as early as 1980. However, most of the algorithms are adaptive, that is, recursive in time, and converge quite slowly, sometimes even to local minima.

Problem Formulation. We consider one important issue in digital communications and signal processing, the mitigation of the effects of the propagation channel. This is the role of the so-called equalizer. Reliable equalizing methods have been developed, but usually need prior knowledge of the channel. A good estimation of the channel, also referred to as channel identification, is thus necessary and quite critical.

We will consider here the case of a linear and time-invariant (LTI) scalar single-input single-output (SISO) communication channel. Such a channel can be described as the convolutive filtering of the input signal $x[n]$ by a filter $h[n]$. We will assume furthermore here that $h[n]$ has a finite number of coefficients and so finite impulse response (FIR). The goal is to *build the equations that will allow to compute the channel coefficients* $h[m]$ from the sole statistics of the observation $y[n]$.

3. Algebraic Statistics for Nomlinear System Identification: A case study

Graduate: **Bui Hoang Giang**.

Keywords. Structural health monitoring, damage prognosis, statistical pattern recognition

Introduction & Aim. Structural health monitoring (SHM), an upcoming technology, is a process of implementing a damage detection strategy for infrastructures in civil, mechanical, and aerospace engineering. The SHM process involves the observation of a system over time using periodically sampled dynamic response measurements from an array of sensors. The extraction of damage-sensitive features from these measurements and the statistical analysis of these features is then used to determine the current state of system health.
SHM is a new approach to collect data about critical structural elements using sensors to provide indicators when some anomalies are detected in a structure.
(Read few reference articles I enclose with this).

Basic terms.
Random eigenvalue problems arise in the dynamic analysis of linear stochastic systems – for example in the problem of dynamic response prediction of cars rolling out from a production chain... The study of probabilistic characterization of the eigensolutions of random matrix and differential operators is now an important research topic in the field of *Stochastic structural mechanics*.

Problem Formulation. We intent to study several vibration-based damage detection methods. The feasibility of continuously monitoring such a structure for the onset of damage was also going to be studied. However, the restrictions that the damage must be relatively benign or repairable made it difficult to take the damage identification portion of the study to completion. Subsequently, this study focused on quantifying the variability in identified **modal parameters** caused by sources other than damage. These sources include: a) variability in testing procedures, b) variability in test conditions, and c) environmental variability.

These variabilities must be understood and their influence on identified modal properties quantified before *vibration-based damage detection* can be applied with unambiguous results. Quantifying the variability in the identified modal parameters led to the development of statistical analysis procedures that can be applied to the experimental modal analysis results. These statistical analysis procedures represent one of the major contributions of these studies to the vibration-based damage detection field.

We aim to identify modal properties using Algebraic Statistics methods such as Rational Fraction Polynomial and/or Eigensystem Realization Algorithm (ERA), among others.

4. Constructions of Hadamard matrices and designs

Introduction. Hadamard matrices are specific matrices with all (-1,1)-entries such that the rows and columns are mutually pairwise orthogonal. That is the inner product of any pair of rows (or any pair of columns) equals to 0.

Research motivation. Hadamard matrices have been found importantly useful in commodity manufacturing as well as in signal processing, not mention in theoretic studies. We investigate the constructions of Hadamard matrices and their corresponding balanced designs by using tools of numeric theory, combinatorics and computing.

The project aim. To study specific constructions of Hadamard matrices and if time allows, balanced ternary designs using discrete mathematics such as finite geometries, numeric theory, combinatorics and computing as well as algebra.

Scope of our investigation. Restrict to the following specific constructions with the aim of constructing all Hadamard matrices and designs with runsize at most 1000. Furthermore, the research is aimed at studying some specific mathematical techniques for computing ternary designs with runsize at most 500, if time allows. These binary and ternary designs are used intensively in **Industrial Manufacturing**, in **Pharmaceutics, Agriculture** and **Software Engineering**

Working plan. We will conduct the followings.

Learn: the theoretical material and practical contexts at the same time

Survey/Propose: making a survey well known constructions, decide which constructions are efficient, aimed at proposing mathematical methods/techniques to construct Hadamard matrices and ternary designs with small runsizes.

Design and Implement: a software package that allows us to generate or study major characteristics of Hadamard matrices with runsize at most 1000.

Remarks. The CAS used for the research is G.A.P (*Group*, *Algorithms* and *Programming*), an open source computing soft of St. Andrew Univ., UK.

5. Statistical Models and Simulation in Project Risk Management

Introductory. We aim at investigating Statistical Methods for a specific risk management problem using Probabilistic techniques combined with Computer Algebra and Numerical Simulation.

Basic terms. Modern project management requires forecasting techniques for *costs, duration (time)* and *performance* of a project, not only under normal circumtances, but also under external extremities (extreme events).

Problem Formulation. We study the project costs, specifically predict the probability and the bad/good impact of various discruptive/high-risky events on the project performance ultimately, by using a Bayesian framework, Computer Algebra and ideas from Experiment Designs. A few typical aspects of Modeling External Risks are:

(1) how to employ expert's opinions and past company performance in an effective way,
(2) what mathematical and statistical techniques could be used to predict and analyze unexpected-but-servere-impact events in project management?

Scope of our investigation. Aim to conduct a case study in the area of Business or Construction sector using Bayesian Statistics. In particular, we try to analyze, measure and assess mathematically the project global performance by studying concrete and coherence components of a project. Our proposal has few goals:

(1) to investigate and employ powerful techniques of computational algebra and Bayesian statistics in project management;
(2) if time allows, to investigate the problem when removing few limitations of the model.

Steps of the investigation

A. Pre-graduate research period:

a/ Learn the context of the problem
b/ **Algebraically model the problems**, in particular the concept of *gravities* of a project and the additive case of event effects. At the same time, visit a Finance, Construction or Banking firm in HCMC, for industrial internship, to see how unexpected-but-servere-impact events happen in those sectors, in 2-4 weeks.
c/ Design an application package that allows manuauver extremities/external events. Study a computing machinary such as `Singular`, `Matlab`, `R` or `OpenModelica`.

B. Graduate research period:

d/ Forecasting the effects of events in the maximum case
e/ Learn the *Metropolis-Hasting* algorithm and relevants
f/ Implement a pilot application package with basic functionality allowing prediction of small cases.

6. Optimization Methods for Portfolio Risk Management

Graduate: **Doan Thi Ngoc Trinh**.

Introductory. The Modern Portfolio theory (MPT) invented by Harry Markowitz in the 60s still plays an important role in many areas of Finance and Insuarance Management. MPT discusses on how risk-averse investors can construct portfolios to optimize or maximize expected return based on a given level of market risk, emphasizing that risk is an inherent part of higher reward.
Basic terms. A *Markowitz Efficient Portfolio* is one where no added diversification can lower the portfolio's risk for a given return expectation (alternately, no additional expected return can be gained without increasing the risk of the portfolio). The *Markowitz Efficient Frontier* is the set of all portfolios that will give the highest expected return for each given level of risk. These concepts of efficiency were essential to the development of the *Capital Asset Pricing Model* (CAPM).

Problem Formulation. Mathematically, MPT models an asset's return as a random variable, and models a portfolio as a weighted combination of assets so that the return of a portfolio is the weighted combination of the assets' returns. Moreover, a portfolio's return is a random variable, and consequently has an expected value and a variance. Risk, in this model, is the standard deviation of return.

Scope of our investigation. Uitility Functions in Risk Theory and the portfolio investment problem have a very close relationship. We aim at investigating this relationship, using Optimization Methods for a specific risk problem in Portfolio Management, and combining algebraic technique with quadratic optimization. Our proposal has few goals:

(1) Formulate the expected utility maximization problem and then
(2) employ Computational algebra and quadratic optimization techniques together with appropriate software to solve.

Steps of the investigation

A. Pre-graduate research period:
- Study Uitility Functions in Risk Theory
- Study The Modern Portfolio theory
- Design a pilot package

B. Graduate research period:
- Solve the expected utility maximization problem
- Implement the pilot package

Contents

Implementation the software for industrial use. Discuss later, based mainly on practical request from industries and services, if any.

- All SAM graduate has no right to release/commercialize the proposed model in the graduate work and relevant product (as software product).
- HCMUT, through the Graduate Advisor keeps the **Intellectual Property Right** (IPR) of the work. In case of commercializing the software, the UT and the graduate should discuss a compromise solution.

References

[1] Industrially mathematical modeling worshop, SAMSI (2006)
[2] Mathematical Modeling and Computer Simulation, **Wiley**
[3] A.K. Basu, *Introduction to Stochastic Processes*, Alpha Science 2005
[4] L. Kleinrock, *Queueing Systems*, vol 2, John Wiley & Sons, 1976
[5] L. Kleinrock, *Time-shared systems: A theoretical treatment*, Journal of the ACM 14 (2), 1967, 242-261.

[6] Bernd Sturmfels, *Solving Polynomial Systems*, AMS, 2002
[7] Lior Pachter and Bernd Sturmfels, *Algebraic Statistics for Computational Biology*, `http://bio.math.berkeley.edu/ascb/`
[8] Zhang, Ma and Wu, *A compartmental model for analysis of SARS transmission patterns and outbreak control measures in China*, Applied Mathematics and Computation, 162, 2005