Gilbert Saporta, Conservatoire National des Arts et Métiers, Paris, France, gilbert.saporta@cnam.fr

Some Topics On Linear Functional Data Analysis.

### Abstract

In a regression problem where we have a single response variable Y, the case of an infinite (not countable) set of predictors occurs when the predictors are functions or curves ie  $X_t$  with  $t \in [0; T]$ . Supervised classification into two categories by means of Fishers linear discriminant function may be viewed as a special case where the dependent variable has 2 values.

Linear methods look for a predictor which may be expressed as an integral sum (Fisher, 1924):

$$\hat{Y} = \int_0^T X_t \beta(t) dt$$

When t takes continuously its values in an interval [0; T], multicollinearity leads to inconsistent estimation of the parameters. In order to solve this problem, we focus here on linear methods based on an orthogonal decomposition of the predictors leading to dimensionality reduction. Components derived from the Karhunen-Loeve decomposition give, for functional data, the equivalent of principal components regression (PCR). However PLS (partial least squares) performs better than PCR, since principal components are obtained irrespective of the response (Preda, Saporta, 2005 and Preda & al., 2007).

It is often of practical interest to anticipate the prediction of the response earlier than the end of the observed process. A method based on the use of ROC curves and AUC is presented: For real time applications we determine an optimal time s < T giving a prediction based on [0; s] almost as good as the prediction using all data (Costanzo & al, 2006) by using a bootstrap test for the AUC criterion.

### References

1- Costanzo D., Preda C., Saporta G. (2006) Anticipated prediction in discriminant analysis on functional data for binary response . *In COMPSTAT2006*, p. 821-828, Physica-Verlag

2- Fisher R.A. (1924) The Influence of Rainfall on the Yield of Wheat at Rothamsted. *Philosophical Transactions of the Royal Society*, B: 213: 89-142

3- Preda C., Saporta G. (2005) PLS regression on a stochastic process, *Computational Statistics and Data Analysis*, 48, 149-158.

4- Preda C., Saporta G., Lvder C., (2007) PLS classification of functional data, *Computational Statistics*, 22(2), 223-235.

Abdallah Lyzzaik, American University of Beirut, Lebanon, lyzzaik@aub.edu.lb

The Riemann Mapping Theorem For Harmonic Mappings.

# Abstract

Let  $\Omega$  be a bounded simply connected domain containing a point  $w_0$  and having a locally connected boundary, and let  $\omega$  be an analytic function of the open unit disc  $\mathbf{D} = \{\mathbf{z} : |\mathbf{z}| < \mathbf{1}\}$  satisfying  $|\omega| < 1$ . It is known that there exists a one-to-one planar harmonic map of  $\mathbf{D}$  normalized by  $f(0) = w_0$  and  $f_z(0) > 0$  and which maps  $\mathbf{D}$  into  $\Omega$  such that (i) the unrestricted limit  $f^*(e^{it}) =$  $\lim_{z \to e^{it}} f(z)$  exists and belongs to  $\partial\Omega$  for all but a countable subset E of the unit circle  $\mathbf{T} = \partial \mathbf{D}$ , (ii)  $f^*$  is a continuous function on  $\mathbf{T} \setminus \mathbf{E}$  such that for every  $e^{is} \in E$  the one-sided limits  $\lim_{t \to s^+} f^*(e^{it})$  and  $\lim_{t \to s^-} f^*(e^{it})$  exist, belong to  $\partial\Omega$ , and are distinct, and (iii) the cluster set of f at  $e^{is} \in E$  is the straight line segment joining the one-sided limits  $\lim_{t \to s^+} f^*(e^{it})$  and  $\lim_{t \to s^-} f^*(e^{it})$ . It has been a major question, since the thirties of the twentieth century, to establish the uniqueness of f. The purpose of this talk is to survey the main results obtained in the course of resolving this question.

Fouad Chedid, Notre Dame University-Louaize, Lebanon, fchedid@ndu.edu.lb

What Makes An Object Complex?.

### Abstract

This paper surveys three notions that can be used to measure the complexity of an object as envisaged by Grassberger: complexity is something that is between total regularity and total randomness. We propose the amount of useful or nonrandom information in a string as a measure of its complexity. At the heart of our discussion are Kolmogorov's minimal sufficient statistic, Koppel's sophistication, and Bennett's logical depth.

Fozi M. Dannan, Arab International University, Damascus, Syria, fmdan@scsnet.org

# On Kantorovich Inequality.

#### Abstract

The well-known arithmetic-harmonic inequality

$$\left(\sum_{i=1}^{n} \frac{\alpha_i}{x_i}\right)^{-1} \leq \sum_{i=1}^{n} \alpha_i x_i \quad (\sum \alpha = 1, \ x > 0)$$

and it's reversal-Kantorovich inequality has been generalized for scalar case and for the matrix case using the Loewner partial order on the set of  $n \times n$  nonnegative definitive matrices. Some open problems will be given. Fatemah Ayatollah Zadeh Shirazi, University of Tehran, Iran, fatemah@khayam.ut.ac.ir.

On Orbit Spaces In Transformation Groups With A Connected Phase Space.

### Abstract

By a transformation group  $(X, G, \mu)$  (or simply  $(X, G, \mu)$ ) we mean a topological space X (phase space), a topological group G (phase group) with identity e and a continuous map  $\mu : X \times G \to X(\mu(x,g) = xg \ (\forall x \in X, \forall g \in G))$  such that for all  $x \in X$  and  $g, h \in G$  we have xe = x and x(gh) = (xg)h.

In this talk transformation groups with a connected, path connected, locally connected, locally path connected, hyperconnected, ultraconnected, scattered, continua, totally separated, extermally disconnected, or zero dimensional phase space with a look to their orbit space are considered. E.g. it will be verified that if G is hyperconnected, then "X is hyperconnected if and only if X/G is hyperconnected if and only if X/G is ultraconnected if and only if X/G is ultraconnected".

<u>Fadi Awawdeh</u> and Wasfi Shatanawi, The Hashemite University- Jordan, awawdeh@hu.edu.jo

Homotopy Analysis Method For Solving System Of Nonlinear Algebraic Equations.

## Abstract

Solutions of system of nonlinear equations is a relatively complicated problem for which a number of different approaches have been proposed. In this talk, we employ the Homotopy Analysis Method (HAM) to point out an iterative method for solving system of nonlinear algebraic equations. Numerical tests show that the method provide highly accurate results in less number of iterations as compared to some well-known existing methods.

Wassim Daher, Gulf University for Sciences & Technology, Kuwait, Leonard J. Mirman, University of Virginia, U.S.A and Marc Santugini, HEC, Montreal, Canada, *daher.wassim@gmail.com* 

Asymmetric Information And Learning In I.O. And Finance Models.

## Abstract

In this paper, we apply the concept of rational expectations to the study of information flows among agents in a noiseless and imperfectly competitive environment. Specifically, we consider a single market in which several firms supply a homogeneous good to price-taking buyers. The good's characteristics are known to all firms, but only to some buyers. The remaining buyers are uninformed, but anticipate that the market-clearing price conveys information about the characteristics of the good. Our focus is on a rational expectations equilibrium (REE) that is fully revealing of the information held by the informed agents in the context of three classes of models: a monopoly, a Cournot oligopoly, and a market microstructure model in which both the choice and allocation of a risky asset are determined.

**Ray Solomonoff**, Oxbridge Research, Boston, U.S.A. and **Elias G. Saleeby**, Notre Dame University-Louaize, Lebanon, *esaleeby@ndu.edu.lb* 

Application Of Algorithmic Probability To Linear And Nonlinear Prediction: Comparisons With AIC and MDL.

## Abstract

The role that information theory plays in model selection is very well known. This is examplified by the use of the popular Akaike Information Criteria (AIC) and Rissanen's principle of minimum description length (MDL) for model selection. Algorithmic probability is an inductive model that uses all possible models in parallel for prediction with weights dependent upon the figure of merit of each model. We consider a nested class of autoregressive models and derive ALP weights. We apply these weights to model selection and compare them with MDL and AIC criteria.

Nabil Nassif, American University of Beirut, Lebanon, nn12@aub.edu.lb

Rescalling And Parallel Time Integration For Systemns Of Order Differential Equations.

## Abstract

We consider a system of time-dependent Ordinary Differential Equations (ODE's), where one seeks  $Y : [0, \infty) \to \mathbf{R}^k$ :

$$\frac{dY}{dt} = F(Y), \ 0 < t \le T \le \infty, \ Y(0) = Y_0.$$
(1)

Numerical (time) integration procedures to obtain approximate solutions to (1) are inherently sequential. All ODE's solvers are based on algorithms that "advance" a numerical process from a time  $t_k$  to time  $t_{k+1}$ , where the sequence  $\{t_k\}$  is usually "refined" uniform or adaptive subdivision of [0, T]. Time- Parallel methods become interesting when one seeks to reduce the excessively high computational time of (1), where the time of existence is of very large order  $T \approx \infty$ . Time Parallelization is based on iterative Predictor-Corrector Schemes. It success depends on appropriate predictions of the solution at the beginning of each "time slice" of some coarse grid  $\{[T_{n-1}, T_n) | n = 1, 2, \dots, N \leq \infty\}$  where  $[0, T] = \bigcup_{n=1}^{N} [T_{n-1}, T_n]$ . Good predictions would automatically reduce the total number of iterations  $N_{it}(N_{it} \leq N)$ . A key performance indicator as to the high

performance of the method should verify the criterion  $N_{it}(N_{it} \ll N)$ , i.e. the number of iterations should be much less than the number of slices.

On the basis of a multi-scaling technique that rescales both the variables Y and t on every slice, we introduce a concept of similarity which leads to a "ratio property" in the case when

$$(F(Y))_i = \sum_j a_{ij} Y_i^{k_{ij}} Y_j^{l_i j}$$

Such situations occur when solving diffusion reaction problems an also Lotka-Voltera predators-preys logistics models.

In this paper, these concepts are applied to a second order initial value model:

$$y'' - |y'|^{q-1} + |y|^{p-1}y = 0, \ t > 0, \ y(0) = y_{1,0}, \ y'(0) = y_{2,0}$$
(2)

This ODE describes the motion of a membrane. Specifically, when

$$p \le q \le \frac{2p}{p+1} \quad \text{and} \ p < 1, \tag{3}$$

the existence of the solution is global with a non-ascillatory blow-up at  $\infty$ , i.e. :

$$\lim_{t \to \infty} |y(t)| = \lim_{t \to \infty} |y'(t)| = \infty, \tag{4}$$

the roots of y and y' being both countably infinite sets.

We show through slices rescaling, that similarity can be obtained with a ratio property leading to an efficient parallel time solver for (2).

**Nassar H.S. Haidar**, Suffolk University, USA ; Dakar Campus, Senegal, *nhaidar@suffolk.edu* 

Finite Integral Transformation On Momentum Hilbert Spaces.

# Abstract

Recent consistency findings about Heisenberg's Uncertainty Principle are to be used to motivate momentum Hilbert spaces  $\mathfrak{L}L^2(R,A,\mu_s)$ . We shall report on solving the inverse problem of finding the kernel of an asymptotic singular integral operator under which monomial signals, that are compactly supported over the closed unit interval  $[0,1^-]$ , are asymptotically generalized fix-points over the semi-closed interval (0,1]. This operator defines a certain Even-Hilbert Riemann-Lebesgue transformation with a kernel that is double parameterized over a certain momentum Hilbert space. Inversion of this ill-posed singular transformation is to be shown to be in the form of a singular integral over some acceleration Hilbert space  $\mathfrak{F}LL^2(R,A,\mu_s)$ .

Friedemann Brock, American University of Beirut, Lebanon, fb13@aub.edu.lb

Sharp Constants For Hardy-Type Inequalities In A Half-Space.

## Abstract

Hardy-type inequalities on Euclidean space can be used to obtain a-priori estimates and symmetry results for solutions of elliptic PDE in general domains, and they have received considerable interest during the last decade. The following result is well known: Let a < (N-2)/2. Then

$$\int_{\mathbf{R}^N} |x|^{-2a} |\nabla u|^2 \, dx \ge C(N,a) \int_{\mathbf{R}^N} u^2 |x|^{-2a-2} \, dx,\tag{1}$$

for every  $u \in C_0^{\infty}(\mathbf{R}^N)$ , where  $C(N, a) = (N - 2(a+1))^2/4$ . The constant C(N, a) in (1) is sharp.

We show the following result: Let H be the halfspace  $\{x \in \mathbf{R}^N : x_1 > 0\}$ . Then

$$\int_{H} |x|^{-2a} |\nabla u|^2 \, dx \ge B(N,a) \int_{H} u^2 |x|^{-2a-2} \, dx, \tag{2}$$

for every function  $u \in C_0^{\infty}(H)$ , with  $B(N, a) = (N - 2a)^2/4 + 2a$ . The constant B(N, a) in (2) this is sharp. Notice that  $B(N, a) \ge C(N, a)$ .

John Haddad, American University of Beirut, Lebanon, jhaddad@aub.edu.lb

Preliminary Estimation In Time Series ARMA Models.

## Abstract

In this talk, Preliminary estimation of the  $k^{th}$ -Lag autocorrelation function in the Gaussian stationary processes will be considered. An estimation procedure is derived from the ratio of the sum filter and the difference filter. The performance of the derived estimator is compared to the sample estimator, the usual initial estimator, through a Monte Carlo study. Finally, comments on the asymptotic distribution are discussed.

Mourad Mahmoud, Lebanese University, Lebanon, fmmourad@yahoo.fr

An Econometric Survey Of The Money Demand In Lebanon Using The Vector Autoregressive Model.

#### Abstract

This paper analyses the relationship between the aggregate monetary M3 in Lebanon and the BDL coincident indicator and the spread between the shortterm market interest rate and the own rate of M3. Four variables in monthly data will be studied for the period January 1995 through December 2005. All the data are in indexes and the year 1995 is used as the reference base period. The analysis of the stationarity will de made using the Augmented Dickey-Fuller (ADF) procedure. For every equation of the vector autoregressive model (VAR), the identification of the autoregressive order of every variable will be made using the FPE criterion following the strategy proposed by Hsiao. The forecasting analysis are generated from (possibly cointegrated) vector autoregressions for the corresponding time series. The ex-post forecasts will be calculated for twelve horizons and for one horizon (one-step ahead forecast). The quality of forecasts will be measured using the mean absolute percentage error (MAPE) criterion. Finally we are going to propose recommendations for the pursuit of this research.

Dania Sheaib, American University of Beirut, Lebanon, dws02@aub.edu.lb

Numerical Solutions Of Scalar Population Models.

# Abstract

Numerical solutions of scalar population models and SIR epidemic models are investigated through variational formulations. Petrov-Galerkin semi- discrete approximations are then derived using Legendre polynomials. This leads to systems of Ordinary Differential Equations whose computation is done using Runge-Kutta type solvers.Implementation is done first on linear scalar population models and are extended to non-linear SIR epidemic models.

Ziad Abdallah, Ndèye Niang-Keita, Genane Youness, ISAE, CEDRIC CNAM, Beirut, Lebanon and Gilbert Saporta, CEDRIC CNAM, Paris, France, *zabdallah@cas.gov.lb* 

Data Stream Mining: Modeling, Classification And Concept-Drifting.

## Abstract

In many areas such as telecommunications, web services, marketing CRM, environment, finance, health, and biology and so on, massive amount of data are continuously generated at a very fast rhythm which requires new methodologies and techniques to capture dynamically and analyze data streams. Studies in recent years on the theme of data flows have brought a full-fledged field at the intersection of databases, information and statistics. There are two major directions of research: the design of Data Stream Management System (DSMS) and the methods and algorithms for data stream. We work d mainly in the Data Stream Mining methods.

The main difficult for these methods is linked to "infinite" nature of data flows. Two techniques are being considered to address this problem: on the one hand be limited to a portion of the stream and on the other hand try to summarize the data stream. Incremental automatic classification methods, research rules associations on the flow of data have been proposed and are the subject of interesting developments.

Another fundamental difficult is the dynamic aspect (real time) of data stream which raises the question of evolution, and detection of changes in the flow of data: the distribution of data flows may indeed change over time which could impact the consistency of the used models. The purpose of our work is to contribute to the development of methodological tools relevant to the consideration of changing distributions of data stream.

The work consists to achieve state of the art methods of data stream mining in general and more particularly dynamic methods for the processing of data flow, then to present the research done on the detection of changes "concept drifting" in the flow of data using nonparametric methods.

# References

1- C. Aggarwal, J. Han, J. Wang, P. S. Yu, A *Framework for Clustering Evolving Data Streams*, Proc. 003 Int. Conf. on Very Large Data Bases, Berlin, Germany, Sept. 2003.

2- C. Aggarwal, J. Han, J. Wang, and P. S. Yu, *On Demand Classification of Data Streams*, Proc. 004 Int. Conf. on Knowledge Discovery and Data Mining, Seattle, WA, Aug. 2004.

3- T. Dasu, S. Krishnan, S. Venkatasubramanian, and K. Yi. An informationtheoretic approach to detecting changes in multi-dimensional data streams. Interface, 2006.

4- H. Wang, W. Fan, P. Yu and J. Han, *Mining Concept-Drifting Data Streams using Ensemble Classifiers*, in the 9th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD), Aug. 2003, Washington DC, USA.

5- N. Jiang and L. Gruenwald, "Research issues in Data Stream Association Rule Mining," ACM SIGMOD RECORD, March 2006

6-S. Muthukrishnan. Data Streams: Algorithms and Applications. Now Publishers, 2005.

7- La Revue MODULAD- Special issue on Data Stream Analysis: 36, 2007

8- J. Gao, W. Fan, J. Han Philip, S. Yu:A General Framework for Mining Concept-Drifting Data Streams with Skewed Distributions 2007

9- Ivan Koychev (2007) Approaches for Learning Classifiers of Drifting Concept

Holem M. Saliba, Notre Dame University-Louaize, Lebanon, hsaliba@ndu.edu.lb

On One Variant of H. Weyl-van der Corput Inequality.

## Abstract

Estimating sums is a major study in complex analysis and analytic number theory. One fundamental result in this field, known as H.Weyl - van der Corput inequality, is obtained in the first quarter of the 20th century. This result is studied. As a consequence, a one variant of it is obtained.

Hicham Gebran, Notre Dame University-Shouf, Lebanon, hg03@aub.edu.lb

Fredholm And Properness Properties Of Quasilinear Second Order Elliptic Systems.

# Abstract

We consider a large class of quasilinear second order elliptic systems of the form

$$-\sum_{\alpha,\beta=1}^{N} a_{\alpha\beta}(x,u(x),\nabla u(x))\partial_{\alpha\beta}^{2}u(x) + b(x,u(x),\nabla u(x)) = 0$$

where x varies in an unbounded domain  $\Omega$  of the Euclidian space  $\mathbf{R}^N$  and  $u = (u^1, \dots, u^m)$  is a vector of functions. These systems generate operators acting between the Sobolev spaces  $W^{2,p}(\Omega, \mathbf{R}^m)$  and  $L^p(\Omega, \mathbf{R}^m)$  for p > N. We investigate then the Fredholm and properness properties of these operators and some connections between them. These functional properties play important roles in the existence theory of nonlinear differential equations, and they are related to two recent topological degrees

Khaldoun El-Khaldi, Ziad Rached and <u>Elias G. Saleeby</u>, Notre Dame University-Louaize, Lebanon, *esaleeby@ndu.edu.lb* 

On The Positivity Of Solutions Of A System Of Integrodifferential Equations.

### Abstract

In industrial crystallization, having a narrow crystal size distribution (CSD) is a desirable property for most products. In general, agglomeration of particles gives a longer tail CSD. Both staging and recycle are known to influence the CSD. In this article, we develop a mathematical model for a two-stage crystallizer that includes both particle agglomeration and slurry recycle. We prove that our model — a system of integrodifferential equations, has a positive unique solution under conditions that are physically very reasonable. Moreover, we solve the model numerically and examine the effect of recycle on the CSD. Our results show that slurry recycle attenuated the tail of the CSD.