Proceedings of Workshops on Inverse Problems, Data, Mathematical Statistics and Ecology

LiTH-MAT-R–2011/11–SE

Editors: Vladimir Kozlov, Martin Ohlson and Dietrich von Rosen

Department of Mathematics
Linköping University
SE–581 83 Linköping, Sweden
Linköping 2011
I Introduction

II Contributions

A Nest Spectrum of Symmetric Semi-Bounded Operator and Reconstruction of Manifolds - M. I. Belishev

1 Basic Objects .................................................. 10
   1.1 Operator $L_0$ ........................................... 10
   1.2 Dynamical System ....................................... 10
   1.3 Extension $E$ ............................................ 10
   1.4 Collection $\mathcal{U}$ .................................. 11
2 Nest Spectrum ................................................... 11
   2.1 Nests ...................................................... 11
   2.2 Lattice $\mathcal{L}$ ....................................... 11
   2.3 Space $(\Omega_{L_0}, \tau)$ ................................ 12
3 Reconstruction of Manifolds ................................. 12
   3.1 Simple manifolds ........................................ 12
   3.2 Inverse Problems ........................................ 13
   3.3 Reconstruction .......................................... 14
References ....................................................... 14

B Solving Ill-Posed Cauchy Problems in Three Space Dimensions Using Krylov Methods - L. Eldén

1 A Cauchy Problem on a Cylindrical Domain .................. 18
2 Numerical Implementation .................................... 19
Contents

References ........................................................................ 19

C Dispersion Exponential Models - A. Hassairi ........................................... 21
1 Natural Exponential Families .............................................................. 24
  1.1 Preliminaries ........................................................................ 24
  1.2 Influence of an Affine Transformation on a NEF ....................... 25
  1.3 Influence of Taking Powers of Convolution in NEFs ................. 25
  1.4 Action of $G = GL(R^{d+1})$ on NEFs on $R^d$ ......................... 25
2 Classifications of NEFs ................................................................ 26
  2.1 Morris Class of Quadratic Real NEFs ....................................... 26
  2.2 Letac Mora Class of Cubic Real NEFs ..................................... 26
  2.3 Simple Quadratic NEFs. (Casalis) ......................................... 26
  2.4 Simple Cubic NEFs. (Hassairi) ............................................. 27
References ........................................................................ 27

D An Inverse Problem in Glaciology - V. Kozlov .................................... 29
References ........................................................................ 33

E Key Issues of Modelling Extreme Floods - V. Kuzmin et al. ............... 35
References ........................................................................ 40

F The Inverse Problem for the Wave Equation and Two Acquisition Geometries - M. Lassas and L. Oksanen ................................................................. 43
1 Introduction ........................................................................ 46
2 The Inverse Problem for the Wave Equation .................................. 46
3 Two Acquisition Geometries with Disjoint Sources and Receivers .... 47
References ........................................................................ 48

References ........................................................................ 56

H Testing of Exponentiality with Application to Historic Data - Y. Nikitin and I. Piskun ................................................................. 59
1 A Review of Exponentiality Tests .................................................. 62
2 Efficiency of Tests .................................................................... 62
3 Treatment of Historical Data ...................................................... 63
References ........................................................................ 65

I More on the Kronecker Structured Covariance Matrix - M. Ohlson et al. 67
1 Introduction ........................................................................ 70
2 Model .................................................................................. 70
3 Estimation ............................................................................ 71
References ........................................................................ 72
<table>
<thead>
<tr>
<th>J</th>
<th>The Inverse Problem, EEG Data and Multivariate Statistics - M. Ohlson and D. von Rosen</th>
<th>75</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>78</td>
</tr>
<tr>
<td>2</td>
<td>Assessing Local Complexity of EEG Signals</td>
<td>79</td>
</tr>
<tr>
<td>3</td>
<td>Multivariate Statistical Analysis</td>
<td>80</td>
</tr>
<tr>
<td>References</td>
<td>80</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>K</th>
<th>Comparison of Two Measurement Models of Convolution Type - P. Piiroinen</th>
<th>83</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>86</td>
</tr>
<tr>
<td>2</td>
<td>LeCam Theory and Comparison of Measurements</td>
<td>87</td>
</tr>
<tr>
<td>3</td>
<td>Notes on the Continuous Model and Related Technical Issues</td>
<td>88</td>
</tr>
<tr>
<td>References</td>
<td>88</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>L</th>
<th>Inverse Scattering Problems that have Applications in 2D Non-linear Optics - V. Serov</th>
<th>91</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Formulation of the Problem</td>
<td>94</td>
</tr>
<tr>
<td>2</td>
<td>Born Approximation</td>
<td>96</td>
</tr>
<tr>
<td>3</td>
<td>Fixed Energy Problem</td>
<td>96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>M</th>
<th>Robust Minimax Bias Estimation of the Correlation Coefficient - G. L. Shevlyakov</th>
<th>99</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>102</td>
</tr>
<tr>
<td>2</td>
<td>Main Result</td>
<td>102</td>
</tr>
<tr>
<td>3</td>
<td>Conclusion</td>
<td>104</td>
</tr>
<tr>
<td>References</td>
<td>104</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>Partial Balayage and the Inverse Problem of Potential Theory - T. Sjödin</th>
<th>105</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>108</td>
</tr>
<tr>
<td>2</td>
<td>Partial Balayage</td>
<td>108</td>
</tr>
<tr>
<td>3</td>
<td>Application to the Exterior Inverse Problem</td>
<td>109</td>
</tr>
<tr>
<td>References</td>
<td>111</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>O</th>
<th>Methane Armageddon: Is it Possible? - I. Sudakov and S. Vakulenko</th>
<th>113</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>116</td>
</tr>
<tr>
<td>2</td>
<td>Phase Transition Approach</td>
<td>116</td>
</tr>
<tr>
<td>3</td>
<td>Lake Area Growth and Methane Emission</td>
<td>117</td>
</tr>
<tr>
<td>References</td>
<td>118</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P</th>
<th>Bifurcations in Large Free-Scale Networks - S. Vakulenko and M. Zimin</th>
<th>121</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>124</td>
</tr>
<tr>
<td>2</td>
<td>Main Result</td>
<td>125</td>
</tr>
<tr>
<td>References</td>
<td>126</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q</th>
<th>Goodness-of-Fit Tests Based on Distribution Characterizations, and their Efficiencies - K. Volkova</th>
<th>129</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tests of Normality Based on Shepp Property, and their Efficiencies.</td>
<td>132</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>2 Tests of Exponentiality Based on Rossberg’s Characterization, and their Efficiencies</td>
<td>133</td>
<td></td>
</tr>
<tr>
<td>References</td>
<td>134</td>
<td></td>
</tr>
</tbody>
</table>
# Contributors

<table>
<thead>
<tr>
<th>Author</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmad, M. R.</td>
<td>67</td>
</tr>
<tr>
<td>Belishev, M. I.</td>
<td>7</td>
</tr>
<tr>
<td>Eldén, L.</td>
<td>15</td>
</tr>
<tr>
<td>Hassairi, A.</td>
<td>21</td>
</tr>
<tr>
<td>Korotygina, U.</td>
<td>35</td>
</tr>
<tr>
<td>Kozlov, V.</td>
<td>29</td>
</tr>
<tr>
<td>Kuzmin, V.</td>
<td>35</td>
</tr>
<tr>
<td>Lassas, M.</td>
<td>43</td>
</tr>
<tr>
<td>Makin, I.</td>
<td>35</td>
</tr>
<tr>
<td>Nazarov, A. I.</td>
<td>51</td>
</tr>
<tr>
<td>Nikitin, Y.</td>
<td>59</td>
</tr>
<tr>
<td>Ohlson, M.</td>
<td>67, 75</td>
</tr>
<tr>
<td>Oksanen, L.</td>
<td>43</td>
</tr>
<tr>
<td>Piiroinen, P.</td>
<td>83</td>
</tr>
<tr>
<td>Piskun, I.</td>
<td>59</td>
</tr>
<tr>
<td>von Rosen, D.</td>
<td>67, 75</td>
</tr>
<tr>
<td>Rumyantsev, D.</td>
<td>35</td>
</tr>
<tr>
<td>Serov, V.</td>
<td>91</td>
</tr>
<tr>
<td>Shevlyakov, G. L.</td>
<td>99</td>
</tr>
<tr>
<td>Sjödin, T.</td>
<td>105</td>
</tr>
<tr>
<td>Sudakov, I.</td>
<td>113</td>
</tr>
<tr>
<td>Surkov, A.</td>
<td>35</td>
</tr>
<tr>
<td>Vakulenko, S.</td>
<td>113, 121</td>
</tr>
<tr>
<td>Volkova, K.</td>
<td>129</td>
</tr>
<tr>
<td>Zimin, M.</td>
<td>121</td>
</tr>
</tbody>
</table>
Part I

Introduction
Preface

Processes in Nature may be considered as deterministic or/and random. We are observing global problems such as climate changes (e.g. warming and extreme weather conditions), pollutions (e.g. acidification, fertilization, the spread of many types of pollutants through air and water) and whole ecosystems that are under pressure (e.g. the Baltic sea and the Arctic region). To understand the processes in Nature and (predict) understand what might occur it is not enough with empirical studies. One needs theoretical fundaments including models and theories to perform correct actions against different threats or at least to carry out appropriate simulation studies. For example, extreme value theory can explain some of the observed phenomena, classical risk analysis may be of help, different types of multivariate and high-dimensional analysis can explain data, time series analysis is essential, for forthcoming studies the theory of experimental designs is of interest, data assimilation together with inverse problem technique is useful for adjustment of data into mathematical models and the list can be made much longer. Behind all these approaches mathematics is hidden, sometimes at a very advanced level. Chemical and physical processes influence all observations but the challenge is to do appropriate approximations so that mathematical/statistical models can be applied. The main aim of this project is to present state of the art knowledge concerning the modelling of Nature with focus on mathematical modelling, in particular “inverse and ill-posed problems”, as well as spatio-temporal models. Inverse and ill-posed problems are characterized by the property that the solutions are extremely sensitive to measurement and modelling errors. There are established connections between inverse problems and Bayesian inference but very little has been carried out with focus on parametric inference such as the likelihood approach. Concerning spatio-temporal models these are usually extensions of classical time series models or/and classical multivariate analysis models.

From the Nordic Council of Ministers, within the program Nordic - Russian Cooperation in Education and Research we asked for funding of 3 preparatory meetings where the plan was to create a series of events taking place during 2011-2013. Partner organizations were

- Institute of Problems of Mechanical Engineering, St. Petersburg
- St. Petersburg State University
- Helsinki University
- Swedish Agricultural University
- Stockholm University
- Linköping University

However, there were also some other participants from other universities.

The planned events should be connected to the following fields: applied mathematics, biophysics and mathematical statistics. Within applied mathematics: mathematical modelling and partial differential equations, inverse and ill-posed problems, data assimilation, dynamical systems, linear algebra, matrix theory; within biophysics; neural networks and
inverse modelling of objects; within mathematical statistical; analyses of stochastic processes, spatio-temporal modelling, experimental design, where considered. There exists a wide overlap between these areas and it is challenging to systemize this overlap and transmit this knowledge to students and stakeholders. However, due to unsure funding it was decided to discuss what can be presented during a one-year program. Moreover, due to practical reasons only 2 meetings/workshops were held:


The output from the above events can be summarized as follows:

- We have identified a number of different areas which can be taught on from different perspectives depending on students background of mathematics.
- We have learned to know many interesting researchers who are willing to share there experiences when for example creating a summer school.
- There is no doubt that we can organize cross-disciplinary summer/winter schools with focus on either the Baltic or Arcthic regions.

This booklet is also part of the deliverables. It comprizes extended abstracts of the majority of the talks of the participants showing their great interest. It is in some way a unique cross-disciplinary document which has joined researchers from different areas from Russia, Finland and Sweden.

We are extremely grateful for the support given by the Nordic Council of Ministers (NCM-RU-PA-2009/10382) and all the enthusiastic contributions by the participants, including our host in Helsinki, professor Lassi Päivärinta.
Part II

Contributions
Paper A

Nest Spectrum of Symmetric Semi-Bounded Operator and Reconstruction of Manifolds

Author: M. I. Belishev

Supported by RFBR grants 08-01-00511 and NSh-4210.2010.1.
Nest Spectrum of Symmetric Semi-Bounded Operator and Reconstruction of Manifolds

Mikhail I. Belishev

St. Petersburg Department of the Steklov Mathematical Institute,
St. Petersburg, Russia,
E-mail: belishev@pdmi.ras.ru
1 Basic Objects

1.1 Operator $L_0$

Let $L_0$ be a densely defined positive definite symmetric operator in a separable Hilbert space $\mathcal{H}$ with nonzero defect indexes:

$$(L_0y, y) \geq \gamma \|y\|^2 \quad (\gamma \geq 0), \quad 1 \leq \dim \ker L_0^* \leq \infty.$$ 

Let $L$ be its extension by Friedrichs, so that $L_0 \subset L \subset L_0^*$ and one has

$$L = L^* = \int_0^\infty \lambda dQ_\lambda, \quad (Ly, y) \geq \gamma \|y\|^2,$$

where $dQ_\lambda$ is the spectral measure of $L$.

1.2 Dynamical System

With $L_0$ one associates an evolutionary dynamical system

$$v_{tt} + Lv = h, \quad t > 0 \quad (2)$$

$$v|_{t=0} = v_t|_{t=0} = 0 \quad (3)$$

where $h \in L^2_{loc}((0, \infty); \mathcal{H})$ is a $\mathcal{H}$-valued function of time (control). Its finite energy class solution $v = v^h(t)$ is represented by the Duhamel formula

$$v^h(t) = \int_0^t L^{-\frac{1}{2}} \sin \left( (t - s)L^{\frac{1}{2}} \right) h(s) \, ds = \langle \text{see (1)} \rangle$$

$$= \int_0^t ds \int_0^\infty \frac{\sin \sqrt{\lambda} (t - s)}{\sqrt{\lambda}} dQ_\lambda h(s), \quad t \geq 0 \quad (4)$$

(see, e.g., [3]). Fix a subspace $A \subset \mathcal{H}$; the set

$$V^t_A := \{ v^h(t) \mid h \in L^2_{loc}((0, \infty); A) \}, \quad t > 0$$

of all states produced by $A$-valued controls is called reachable (at the moment $t$, from the subspace $A$). Reachable sets increase as $A$ increases and/or $t$ grows.

1.3 Extension $E$

Let $\text{Lat}\mathcal{H}$ be the lattice of the (closed) subspaces in $\mathcal{H}$ \footnote{so that $A, B \in \text{Lat}\mathcal{H}$ implies $A \cap B, A \vee B, \mathcal{H} \ominus A \in \text{Lat}\mathcal{H}$}. Define a family $E = \{E^t\}_{t \geq 0}$ of the maps $E^t : \text{Lat}\mathcal{H} \to \text{Lat}\mathcal{H}$ by

$$E^0 A := A, \quad E^t A := \text{clos} V^t_A, \quad t > 0.$$ 

As one can show (see [2]), $t \geq t'$ and $A \subseteq A'$ imply $E^t A \subseteq E^{t'} A'$, i.e., $E$ extends subspaces. By this we call it a space extension.
1.4 Collection $\mathcal{U}$

Define a class

$$\mathcal{M} := \{ h \in C^\infty ([0, \infty); \ker L_0^*) \mid \supp h \subset (0, \infty) \}$$

of smooth $\ker L_0^*$-valued controls vanishing near $t = 0$. The sets

$$\mathcal{U}_t := \{ h(t) - v^h(t) \mid h \in \mathcal{M} \} = \{ \text{see (4)} \} = \{ h(t) - t \int_0^t L^{-\frac{1}{2}} \sin \left( (t-s) L^\frac{1}{2} \right) h''(s) ds \mid h \in \mathcal{M} \}, \quad t \geq 0 \quad (5)$$

(here $( \cdot )' := \frac{d}{dt}$) are said to be reachable from boundary $^2$. The sets $\mathcal{U}_t$ increase as $t$ grows. We put $\mathcal{U} := \{ \text{clos} \mathcal{U}_t \}_{t \geq 0} \subset \text{Lat} \mathcal{H}$.

2 Nest Spectrum

2.1 Nests

A family of subspaces $n \subset \text{Lat} \mathcal{H}$ is a nest if for any $A, B \in n$ one has $A \subseteq B$ or $A \supseteq B$ (see [4]). Here we deal with the parametrized nests $n = \{ N_t \}_{t \geq 0}$: $N_t \subseteq N_{t'}$ as $t \geq t'$. These nests are partially subordinated by

$$\{ n \preceq n' \} \Leftrightarrow \{ N_t \subseteq N_{t'}, \quad t \geq 0 \}.$$ 

For a linear set $A \subset \mathcal{H}$, by $P_A$ we denote the (orthogonal) projection in $\mathcal{H}$ onto $\text{clos} A$. A nest $n = \{ N_t \}_{t \geq 0}$ is said to be the limit of a nest sequence $n_1 = \{ N_t^1 \}_{t \geq 0}$, $n_2 = \{ N_t^2 \}_{t \geq 0}, \ldots$ if $P_{N_t} = s\lim_{j \to \infty} P_{N_t^j}$; in such a case we write $n_j \to n$.

2.2 Lattice $\mathcal{L}$

For each subspace $A \in \text{Lat} \mathcal{H}$, the space extension $E$ determines a parametrized nest $n_A := \{ E_t A \}_{t \geq 0}$. The operator $L_0$ determines a (sub)lattice $\mathcal{L} \subset \text{Lat} \mathcal{H}$ such that

1. $\mathcal{U} \subset \mathcal{L}$

2. $E_t \mathcal{L} \subset \mathcal{L}$ for all $t \geq 0$

3. $\mathcal{L}$ is closed w.r.t. the nest convergence introduced above

4. $\mathcal{L}$ is the minimal lattice obeying 1 – 3.

The lattice $\mathcal{L}$ is well defined and determined by the operator $L_0$.

$^2$This term is motivated by applications considered below
2.3 Space \((\Omega_{L_0}, \tau)\)

A nest \(m \in \mathcal{L}\) is said to be \textit{minimal} if the relations \(n \in \mathcal{L}\) and \(n \preceq m\) imply \(n = m\). By \(\Omega_{L_0}\) we denote the set of all minimal nests of the lattice \(\mathcal{L}\) and call it a \textit{nest spectrum} of the operator \(L_0\).

The collection \(\Omega\) is also a parametrized nest. The set \(\partial \Omega_{L_0} := \{ m \in \Omega_{L_0} | m \preceq \Omega\}\) is said to be the \textit{boundary} of \(\Omega_{L_0}\).

The nest spectrum can be endowed with an intrinsic topology. At first, define a function \((\text{quasi-distance})\) \(\rho : \Omega_{L_0} \times \Omega_{L_0} \to [0, \infty) \cup \{\infty\}\) by the following rule. For \(m, m' \in \Omega_{L_0} : m = \{M_t\}_{t \geq 0}, m' = \{M'_t\}_{t \geq 0}\) we put

\[
\rho(m, m') := \begin{cases} 
\infty & \text{if } P_{M_t} P_{M'_t} = \emptyset \text{ for all } t \\
2 \inf \{ t \geq 0 | P_{M_t} P_{M'_t} \neq \emptyset \} & \text{otherwise}.
\end{cases}
\]

Also, introduce the \textit{quasi-balls}

\[
B^r[m] := \{ m' \in \Omega_{L_0} | \rho(m, m') < r \}, \quad r > 0.
\]

Let \(\tau\) be the (weakest) topology on \(\Omega_{L_0}\) generated by the system of quasi-balls \(\{B^r[n] | n \in \Omega_{L_0}, r > 0\}\).

So, \((\Omega_{L_0}, \tau)\) is a topological space associated with the operator \(L_0\) in a canonical way. A fact, which easily follows from the definitions, is that the unitarily equivalent operators have the homeomorphic nest spectra, i.e., if \(\mathcal{H} = U\mathcal{H}\) and \(\tilde{L}_0 = U L_0 U^*\) with a unitary \(U : \mathcal{H} \to \tilde{\mathcal{H}}\) then \((\Omega_{\tilde{L}_0}, \tilde{\tau}) \sim (\Omega_{L_0}, \tau)\). In other words, a nest spectrum is a unitary invariant of the operator class under consideration.

3 Reconstruction of Manifolds

3.1 Simple manifolds

Let \(\Omega\) be a \(C^\infty\)-smooth (possibly noncompact) Riemannian manifold with the boundary \(\partial \Omega\), \(\Delta\) the (scalar) Laplace operator. The minimal Laplacian on \(\Omega\)

\[
L_0 = -\Delta|_{C^\infty_c(\Omega \setminus \partial \Omega)}
\]

is a densely defined symmetric positive definite operator in \(\mathcal{H} := L_2(\Omega)\); its Friedrichs extension is the Dirichlet Laplacian

\[
L = -\Delta|_{H^2(\Omega) \cap H^1_0(\Omega)}
\]

(\(H^k\) are the Sobolev classes), whereas

\[
L^*_0 = -\Delta|_{\{y \in \mathcal{H} | \Delta y \in \mathcal{H}\}}
\]

is the maximal Laplacian.
A class $S$ of *simple manifolds* is introduced in [2]. Roughly speaking, a simplicity means that the symmetry group of $\Omega \in S$ is trivial. The class $S$ is generic: any $\Omega$ can be made into a simple manifold by arbitrarily small smooth variations of its boundary. As is shown in [2],

- for any $\Omega$, the quasi-distance $\rho$ on the nest spectrum of its minimal Laplacian is a metric (distance), so that $(\Omega_{L_0}, \rho)$ is a metric space
- for $\Omega \in S$, there is an isometry (of metric spaces) $i: (\Omega_{L_0}, \rho) \rightarrow \Omega$ such that $i(\partial \Omega_{L_0}) = \partial \Omega$.

Therefore, a simple $\Omega$ is determined by any unitary copy $\tilde{L}_0$ of its minimal Laplacian up to isometry.

### 3.2 Inverse Problems

With the manifold $\Omega$ one associates an evolutionary dynamical system

\begin{align*}
  u_{tt} - \Delta u &= 0 & \text{in } \Omega \times (0, \infty) \tag{6} \\
  u|_{t=0} &= u_t|_{t=0} = 0 & \text{in } \Omega \tag{7} \\
  u|_{\partial \Omega} &= f(t) & \text{for } 0 \leq t < \infty \tag{8}
\end{align*}

with a boundary control $f \in F := L_{2, loc}^\infty \left((0, \infty); L_2(\partial \Omega)\right)$; the solution $u = u^f(x,t)$ describes a wave, which is initiated by boundary sources and propagates from the boundary into the manifold. The collection $U$ defined in sec 1.4 consists of the reachable sets $U_t := \{ u^f(\cdot,t) \mid f \in F \}$ (see (5)). The input/output correspondence is realized by a response operator $R: f \mapsto \partial_\nu u^f|_{\partial \Omega \times [0,\infty)}$ defined on smooth controls vanishing near $t = 0$ (here $\nu$ is the outward normal to $\partial \Omega$, $\partial_\nu$ the normal derivative). The time-domain (dynamical) inverse problem is

**IP 1:** given the response operator $R$ of the system (6)–(8), to recover the manifold $\Omega$.

With the manifold $\Omega$ one associates a stationary dynamical system

\begin{align*}
  (-\Delta + z) w &= 0 & \text{in } \Omega \tag{9} \\
  w &= \varphi & \text{on } \partial \Omega, \tag{10}
\end{align*}

where $z \in \mathbb{C} \setminus \text{spec } L$ is a complex parameter (frequency), $w = w^\varphi(x)$ is the solution. The *Weyl-Titchmarsh function* of the system is

\[ M(z) \varphi = \partial_\nu w^\varphi|_{\partial \Omega}, \]

which is a $L_2(\partial \Omega) \rightarrow L_2(\partial \Omega)$-operator valued function. The frequency-domain inverse problem is

**IP 2:** given the W-T function $M$ of the system (9)–(10), to recover the manifold $\Omega$.

---

3. There are easily checkable sufficient conditions of geometric character, which provide $\Omega \in S$: see [2] for detail.
Assume that the manifold $\Omega$ is compact. Let $\{\lambda_k\}_{k=1}^{\infty}$ be the spectrum of the Dirichlet Laplacian $L$, $\{\phi_k\}_{k=1}^{\infty}$ its eigen basis in $\mathcal{H}$ normalized by $(\phi_k, \phi_l) = \delta_{kl}$. The set of pairs 
\[ \Sigma := \{\lambda_k; \partial^\nu \phi_k |_{\partial \Omega}\}_{k=1}^{\infty} \]
is called the (Dirichlet) spectral data of the manifold $\Omega$. The spectral inverse problem is 
**IP 3:** given the spectral data $\Sigma$, to recover the manifold $\Omega$.

### 3.3 Reconstruction

Setting the goal to determine an unknown manifold from its boundary inverse data, one has to keep in mind the evident nonuniqueness of such a determination: all isometric manifolds with the mutual boundary have the same data. Therefore, the only reasonable understanding of "to recover" is to construct a manifold, which possesses the prescribed data [1].

An affirmative common feature of the problems IP 1–3 is that each kind of their data (i.e., $R$, $M$, and $\Sigma$) determines the minimal Laplacian $L_0$ up to unitary equivalence: see [1], [2]. By this, each kind of data determines the wave spectrum $\Omega_{L_0}$ up to isometry.

Assume that $\Omega$ is a simple manifold. Given its inverse data, one can

- determine a relevant unitary copy $\tilde{L}_0$ of the minimal Laplacian
- find its nest spectrum $\Omega_{\tilde{L}_0}$ and endow it with the distance $\tilde{\rho}$

and hence get an isometric copy $(\Omega_{\tilde{L}_0}, \tilde{\rho})$ of the original manifold $\Omega$. By isometry, the copy has the same boundary data as the original. Thus, the inverse problems are solved.

The above described scheme of reconstruction elucidates the operator background of the boundary control method, which is an approach to inverse problems based upon their relations to system and control theory [1].

### References


Paper B

Solving Ill-Posed Cauchy Problems in Three Space Dimensions Using Krylov Methods

Author: L. Eldén
Solving Ill-Posed Cauchy Problems in Three Space Dimensions Using Krylov Methods

Lars Eldén

Department of Mathematics,
Linköping University,
Linköping, Sweden.
E-mail: laeld@mai.liu.se
1 A Cauchy Problem on a Cylindrical Domain

Let $\Omega$ be a connected domain in $\mathbb{R}^2$ with smooth boundary $\partial \Omega$, and assume that $L$ is a linear, self-adjoint, and positive definite elliptic operator defined in $\Omega$. We consider the ill-posed Cauchy problem,

\[
\begin{align*}
    u_{zz} - L u &= 0, & (x, y, z) &\in \Omega \times [0, z_1], \\
    u(x, y, z) &= 0, & (x, y, z) &\in \partial \Omega \times [0, z_1], \\
    u(x, y, 0) &= g(x, y), & (x, y) &\in \Omega, \\
    u_z(x, y, 0) &= 0, & (x, y) &\in \Omega.
\end{align*}
\]

The problem is to determine the values of $u$ on the upper boundary, $f(x, y) = u(x, y, z_1)$, $(x, y) \in \Omega$.

In this talk we describe the main ideas behind the paper [6].

The problem (1) is ill-posed in the sense that the solution (if it exists), does not depend continuously on the data. It is a variant of a classical problem considered originally by Hadamard, and it is straightforward to analyze it using an eigenfunction expansion.

Since the domain is cylindrical with respect to $z$, we can use a separation of variables approach, and write the solution of (1) formally as

\[
    u(x, y, z) = \cosh(z \sqrt{L}) g.
\]

The operator $\cosh(z \sqrt{L})$ can be expressed in terms of the eigenvalue expansion of $L$. Due to the fact that $L$ is unbounded, the computation of $\cosh(z \sqrt{L})$ is unstable and any data errors or rounding errors would be blown up, leading to a meaningless approximation of the solution.

The problem can be stabilized (regularized) if the operator $L$ is replaced by a bounded approximation, see [3, 4, 5, 9, 10, 2], where similar problems were treated. Since it is the large eigenvalues of $L$ (those that tend to infinity) that are associated with the ill-posedness, it would be natural to devise the following regularization method:

- Compute approximations of the smallest eigenvalues of $L$ and the corresponding eigenfunctions, and discard the components of the solution (2) that correspond to large eigenvalues.

It is straightforward to prove that such a method is a regularization method in the sense that the solution depends continuously on the data. However, in the direct implementation of such a method one would use unnecessarily much work to compute eigenvalue-eigenfunction approximations that are not needed for the particular data function $g$. Thus the main contribution of the paper [6] is a numerical method for approximating the regularized solution that has the following characteristics:

- The solution (2) is approximated by a projection onto a subspace computed by means of a Krylov sequence generated using the operator $L^{-1}$. The hyperbolic cosine of the restriction of the operator $L^{-1}$ to that subspace is computed.

- At each step of the Krylov recursion, dealing with $L^{-1}$ corresponds to solving a well-posed two-dimensional elliptic problem involving $L$. Any standard (black box) elliptic solver, derived from the discretization of $L$, can be used.
The method takes advantage of the fact that the regularized solution operator is applied to the particular data function $g_m$.

It is demonstrated in [6] that the proposed method requires considerably fewer solutions of two-dimensional elliptic problems, than the approach based on the eigenvalue expansion.

There are many engineering applications of ill-posed Cauchy problems, see [7, 8, 11] and the references therein. A standard approach for solving Cauchy problems of this type is to apply an iterative procedure, where a certain energy functional is minimized; a recent example is given in [1]. Very general (non-cylindrical) problems can be handled, but if the procedure from [1] were to be applied to our problem, then at each iteration four well-posed elliptic equations would have to be solved over the whole three-dimensional domain. In contrast, our approach for the cylindrical case requires the solution of only one two-dimensional problem at each iteration\(^1\). A few details of the method are described below.

2 Numerical Implementation

Assume that the problem has been discretized with respect to $(x, y)$ so that now the elliptic, self-adjoint operator $L$ is represented by a symmetric matrix, for simplicity also denoted $L$. Solving the 2D elliptic problem $Lv = w$ is equivalent to applying the the inverse $v = L^{-1}w$. Using the Lanczos procedure we compute an orthonormal basis $(q_1, q_2, \ldots, q_k)$ for the Krylov subspace

$$K_k(L^{-1}, g_m) = \text{span}(g_m, L^{-1}g_m, L^{-2}g_m, \ldots, L^{-(k-1)}g_m),$$

where $g_m$ is a discrete version of the data function in (1). This gives a low-rank approximation of $L^{-1}$,

$$L^{-1}Q_k \approx Q_kT_k,$$

where $T_k$ is a tridiagonal matrix. It is very likely that some of the large eigenvalues of $L$ are rather well approximated by $T_k$. Therefore we compute the eigenvalue decomposition of $T_k$, and replace those eigenvalues by zero, giving a “regularized” matrix $\hat{T}_k$. We then compute an approximation of (2) by projection to the $k$-dimensional Krylov subspace,

$$u_k(z) = Q_k \cosh(\hat{T}_k^\dagger)Q_k^T g_m,$$

where $\hat{T}_k^\dagger$ is the Moore-Penrose pseudoinverse of $\hat{T}_k$.

The procedure described can be generalized in a straightforward way to problems in more than three space dimensions.

References


\(^{1}\)In one of our examples the two-dimensional problem had 8065 degrees of freedom and the three-dimensional one had 250015.


Paper C

Dispersion Exponential Models

Author: A. Hassairi
Dispersion Exponential Models

Abdelhamid Hassairi

Laboratory of Probability and Statistics,
Sfax Faculty of Sciences,
Sfax, Tunisia.
E-mail: abdelhamid.hassairi@fss.rnu.tn
1 Natural Exponential Families

1.1 Preliminaries

Let $E$ be a linear space dimension $d$, let $E^*$ be its dual $E^* \times E \to \mathbb{R} : (\theta, x) \mapsto \langle \theta, x \rangle$ be the duality bracket. We denote by $\mathcal{L}_s(E^*, E)$ the space of the symmetric linear maps from $E^*$ to $E$. If $\mu$ is a positive Radon measure on $E$, we denote $\Theta(\mu) = \text{interior} \{ \theta \in E^*; L_\mu(\theta) < +\infty \}$, $k_\mu = \log L_\mu$.

$\mathcal{M}(E)$ is the set of measures $\mu$ such that $\Theta(\mu)$ is not empty and $\mu$ is not concentrated on an affine hyperplane of $E$. To each $\mu$ in $\mathcal{M}(E)$ and $\theta$ in $\Theta(\mu)$, we associate the probability distribution on $E$

$$P(\theta, \mu)(dx) = \exp\{\langle \theta, x \rangle - k_\mu(\theta)\} \mu(dx).$$

The set

$$F = F(\mu) = \{P(\theta, \mu); \theta \in \Theta(\mu)\}$$

is called the natural exponential family generated by $\mu$. If $\mu$ and $\mu'$ are in $\mathcal{M}(E)$, then $F(\mu) = F(\mu')$ if and only if there exists $(a, b)$ in $E \times \mathbb{R}$ such that $\mu'(dx) = \exp\{\langle a, x \rangle + b\} \mu(dx)$. Therefore, if $\mu$ is in $\mathcal{M}(E)$ and $F = F(\mu)$,

$$\mathcal{B}_F = \{\mu' \in \mathcal{M}(E); F(\mu') = F\}$$

is the set of basis of $F$.

The function $k_\mu$ is strictly convex and real analytic. $k'_\mu$ is a diffeomorphism between $\Theta(\mu)$ and its image $M_F$ called the domain of the means of $F$.

The inverse function of $k'_\mu$ is denoted by $\psi_\mu$ and setting $P(m, F) = P(\psi(m), \mu)$ the probability of $F$ with mean $m$, we have

$$F = \{P(m, F); m \in M_F\}.$$

The density of $P(m, F)$ with respect to $\mu$ is

$$f_\mu(x, m) = \exp\{\langle \psi_\mu(m), x \rangle - k_\mu(\psi_\mu(m))\},$$

Now the covariance operator of $P(m, F)$ is denoted by $V_F(m)$. Clearly

$$V_F(m) = k''_\mu(\psi_\mu(m)) = (\psi'_\mu(m))^{-1} \in \mathcal{L}_s(E^*, E).$$

The variance function $m \mapsto V_F(m)$ characterizes the family $F$ in the following sense: If $F$ and $F'$ are two NEFs such that $V_F(m)$ and $V_F'(m)$ coincide on a nonempty open subset of $M_F \cap M_{F'}$, then $F = F'$. 


1.2 Influence of an Affine Transformation on a NEF.

Let $\varphi(x) = \delta(x) + \gamma$ where $\delta$ is in $\text{GL}(E)$ and $\gamma$ is in $E$, and let $F = F(\mu)$ be a NEF on $E$.

$$
\begin{align*}
\varphi(F) &= F(\varphi(\mu)), \\
M_{\varphi(F)} &= \varphi(M_F), \\
V_{\varphi(F)}(m) &= \delta V_F(\varphi^{-1}(m))^t \delta.
\end{align*}
$$

1.3 Influence of Taking Powers of Convolution in NEFs

Let $\alpha$ be a positive number, not necessarily an integer. If $\mu$ is in $\mathcal{M}(E)$, let us introduce the Jorgensen set $\Lambda = \{\alpha > 0; \text{there exists } \mu_\alpha \text{ in } \mathcal{M}(E) \mid \Theta(\mu_\alpha) = \Theta(\mu) \text{ and } k_{\mu_\alpha}(\theta) = \alpha k_\mu(\theta)\}$.

Denoting $F_\alpha = F(\mu_\alpha)$ where $\alpha$ is in $\Lambda$, one has

$$
\begin{align*}
M_{F_\alpha} &= \alpha M_F, \\
V_{F_\alpha}(m) &= \alpha V_F \left(\frac{m}{\alpha}\right).
\end{align*}
$$

1.4 Action of $G = GL(R^{d+1})$ on NEFs on $R^d$

An element $g = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$ of $G = GL(R^{d+1})$ is defined by its blocs $(\alpha, \beta, \gamma, \delta)$ in $R \times R^d \times R^d \times L(R^d)$. We denote

$$d_g(m) = \alpha + \langle \beta, m \rangle \text{ et } h_g(m) = (d_g(m))^{-1} (\gamma + \delta(m)).$$

We define the action of the group $G$ on a NEF $F$ by,

$$(T_g V_F) (m) = (d_g(m))^{-1} (h_g'(m))^{-1} V_F (h_g(m)) (h_g'(m))^* -1.$$

If $\alpha = 1$ et $\beta = 0$, the image $F_1$ of $F$ by the affinity $x \mapsto \delta(x) + \gamma$ satisfies

$$V_{F_1} = T_g V_F.$$

Also if $\alpha$ is in the Jorgensen set of $F$, then for $g = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix}$, $T_g$ corresponds to the Jorgensen transformation of parameter $\alpha$. In particular, when $d = 1$, the action of an element of $GL(R^2)$ on a real NEF $F$ is given by

$$(T_g V_F) (m) = \frac{(\alpha + \beta m)^3}{(\alpha \delta - \beta \gamma)} V_F \left(\frac{\gamma + \delta m}{\alpha + \beta m}\right).$$

Let $G_0$ the subgroup of $G$ such that $\beta = 0$ and $\alpha > 0$.

All the classifications of NEFs are done up to affine transformations and Jorgensen transformations, that is up to $G_0$ orbits.


## 2 Classifications of NEFs

### 2.1 Morris Class of Quadratic Real NEFs

The class of NEFs with quadratic variance functions contains exactly six $G_0$ orbits.

<table>
<thead>
<tr>
<th>Name</th>
<th>$M_F$</th>
<th>$V_F(m)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>$\mathbb{R}$</td>
<td>1</td>
</tr>
<tr>
<td>Poisson</td>
<td>$[0; +\infty]$</td>
<td>$m$</td>
</tr>
<tr>
<td>Binomial</td>
<td>$[0; n]$</td>
<td>$m - m^2$</td>
</tr>
<tr>
<td>Negative binomiale</td>
<td>$[0; +\infty]$</td>
<td>$m + m^2$</td>
</tr>
<tr>
<td>Gamma</td>
<td>$[0; +\infty]$</td>
<td>$m^2$</td>
</tr>
<tr>
<td>Hyperbolic-cosine</td>
<td>$\mathbb{R}$</td>
<td>$1 + m^2$</td>
</tr>
</tbody>
</table>

### 2.2 Letac Mora Class of Cubic Real NEFs

The class of NEFs with variance functions of degree three contains six $G_0$ orbits.

<table>
<thead>
<tr>
<th>Type</th>
<th>$\mu$</th>
<th>$V_F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abel</td>
<td>$\sum_{k=0}^{\infty} p(p + k)^{k-1} \frac{\delta_k}{k!}$</td>
<td>$m(1 + \frac{m}{p})^2$</td>
</tr>
<tr>
<td>Takàcs</td>
<td>$\sum_{k=0}^{\infty} a p \prod_{j=1}^{k-1} (a(p + k) + j) \frac{\delta_k}{k!}$</td>
<td>$(m + \frac{m^2}{p}) (1 + \frac{a+1 + m}{a + m})$</td>
</tr>
<tr>
<td>Strict arcsine</td>
<td>$\sum_{k=0}^{\infty} p_k(p) \frac{\delta_k}{k!}$</td>
<td>$m(1 + \frac{m^2}{p^2})$</td>
</tr>
<tr>
<td>Large arcsine</td>
<td>$\sum_{k=0}^{\infty} \frac{p}{p+k} p_k^{(*)} (a(p + k)) \frac{\delta_k}{k!}$</td>
<td>$(m + \frac{2m^2}{p} + \frac{a^2+1 + m^3}{a^2 + m^2})$</td>
</tr>
<tr>
<td>Ressel</td>
<td>$\frac{p x^{x+p-1} e^{-x}}{(x+p+1)} I_{[0, +\infty]}$</td>
<td>$\frac{m^3}{p} (1 + \frac{m}{p})$</td>
</tr>
<tr>
<td>Inverse Gaussian</td>
<td>$x^{-\frac{3}{2}} \exp \left( - \frac{p}{2x} \right) \frac{p}{\sqrt{2\pi}} I_{[0, +\infty]}$</td>
<td>$\frac{m^3}{p^2}$</td>
</tr>
</tbody>
</table>

(*) $p_{2n}(a) = \prod_{k=0}^{n} (a^2 + 4k^2)$ et $p_{2n+1}(a) = \prod_{k=0}^{n} (a^2 + (2k + 1)^2)$

### 2.3 Simple Quadratic NEFs. (Casalis)

The multivariate version of the Morris class is the class of NEFs on $\mathbb{R}^d$ with variance functions of the form

$$V_F(m) = a m \otimes m + B(m) + C,$$

where $m \otimes m(\theta) = \langle \theta, m \rangle$, $m > m$ and $B$.

The class of simple quadratic NEFs on $E$, contains $(2d + 4) G_0$--orbits:

- $(d + 1)$ Poisson-Gaussian $G_0$--orbits,
- $(d + 1)$ negative multinomial-gamma $G_0$--orbits,
- a multinomial $G_0$--orbit,
- an hyperbolic $G_0$--orbit built from particular mixtures of families of normal, Poisson, gamma, hyperbolic on $\mathbb{R}$ and (negative) multinomial distributions.
2.4 Simple Cubic NEFs. (Hassairi)

The twelve $G_0$—orbits of the Letac-Mora are divided into 4 $G$-orbits

<table>
<thead>
<tr>
<th>Gaussian 1</th>
<th>Inverse Gaussian $m^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson $m$</td>
<td>$\text{Gamma } m^2$</td>
</tr>
<tr>
<td>Binomial $m(1 - m)$</td>
<td>Negative-Binomial $m(1 + m)$</td>
</tr>
<tr>
<td>Hyperbolic $1 + m^2$</td>
<td>$\text{Large arcsine } m \left(1 + 2m + \frac{1+a^2}{a^2}m^2\right)$</td>
</tr>
</tbody>
</table>

Let $F$ be simple quadratic NEF and $g = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$ in the group $G$. Then $T_gV_F$ is a polynomial in $m$ of degree $\leq 3$

$$V_{F_1} = \left\{ a m \otimes m + [I + m \otimes \beta][B(m) + (< \beta, m > + 1)C] \right\} [I + \beta \otimes m].$$

The obtained NEF $F_1 = T_gF$ is said simple cubic.

The class of simple cubic NEFs on $\mathbb{R}^d$ is distributed into $3d + 1$ $G$-orbits

References


Paper D

An Inverse Problem in Glaciology

Authors: V. Kozlov

This talk is based on joint work with S. Avdonin, D. Maxwell, and M. Truffer [1].
An Inverse Problem in Glaciology

Vladimir Kozlov

Department of Mathematics,
Linköping University,
Linköping, Sweden.
E-mail: vladimir.kozlov@liu.se
Many problems in geophysics are ill-posed. For example, there is no known method to measure basal velocities, but surface velocities can be measured directly on the ground or by a variety of remote sensing methods. Basal velocities must then be found through inverse methods. We will consider an ice flow model suggested in [2]. They treated a first order model of planar ice flow along a longitudinal cross section of a glacier, and showed that the longitudinal velocity component obeys a non-linear Poisson equation.

Let $\Omega$ be the domain in the $xy$-plane with Lipschitz boundary, which consists of the upper boundary $S = (0, l)$, a surface of the ice sheet, and $B$, a bottom, connecting end points of $S$ and lying below $S$. We consider the following system of equations in $\Omega$:

\begin{align*}
-\nabla \cdot (G(|\nabla u|)\nabla u) &= f \quad \text{in } \Omega \\
\partial_y u |_S &= 0
\end{align*}

where $f$ is a given function. We suppose that

$$G(t) = \frac{F'(t)}{t},$$

where $F$ is a convex function for $t \geq 0$, $F'(0) = 0$ and $F'$ is Lipschitz continuous, i.e. $|F'(t) - F'(\tau)| \leq C|t - \tau|$. For such $F$, $F''(t)$ exists almost everywhere and we suppose that

$$\nu(1 + t)^{\frac{1}{n} - 1} \leq F''(t) \leq \mu(1 + t)^{\frac{1}{n} - 1}$$

for some positive $\nu, \mu$ and $n \geq 1$.

A finite viscosity version of Glen’s flow law, which is often used in glaciology (e.g. [3]), gives an example of the function $G(t)$ as the solution of the equation

$$\frac{1}{G(s)} = \left[T_0^2 + G^2(s) s^2\right]^{\frac{1}{n} - 1}$$

with $T_0 \neq 0$. Then

$$F(t) = \int_0^t G(\tau) \tau d\tau$$

is strictly convex and satisfies (3), see [2].

Conditions (1)–(2) do not distinguish the unique solution of equation (1), so we add an additional condition on $S$ which can be interpreted as a surface measurements:

$$u |_S = \varphi$$

(4)

Problem (1)–(2), (4) is ill-posed: its solution exists not for every (even smooth) functions $\varphi$ and one cannot expect continuous dependence of the solution on $\varphi$.

The alternative method for solving ill-posed linear problem was introduced in [4] and [5]. For our nonlinear problem it runs as follows. The approximation $u_0$ is found by solving the boundary value problem (1), (2) supplied with the Dirichlet boundary condition on $B$: $u |_B = \psi$, where $\psi$ is an arbitrary function from an appropriate Sobolev space on $B$. In order to find the $(k + 1)$th iteration one must solve the nonlinear equation (1) for $u = u_{k+1}$ supplied with the Dirichlet-Neumann (or Neumann-Dirichlet, depending on $k$
is odd or even) boundary conditions on $S$ and $B$. The boundary condition on $B$ is taken from the given Cauchy data and the boundary condition on $S$ is taken from the previous iteration step. It is possible some variations in this procedure, for example instead of the nonlinear equation (1) we can solve the (linear) equation

$$-\nabla \cdot (G(|\nabla u_k|)\nabla u_{k+1}) = f$$

with the same boundary conditions.

Our numerical experiments with the nonlinear problem (1), (2), (4) indicate that this procedure appears to have similar convergence properties as for the linear case.

A Landweber-type iterative procedure for solving the above non-linear problem is also discussed.

References


Paper E

Key Issues of Modelling Extreme Floods

Authors: V. Kuzmin, U. Korotygina, I. Makin, D. Rumyantsev and A. Surkov
Key Issues of Modelling Extreme Floods

Vadim Kuzmin*, Ulyana Korotygina, Ivan Makin, Denis Rumyantsev and
Aleksandr Surkov

Russian State Hydrometeorological University,
St. Petersburg, Russia.
*E-mail: vknoaa@hotmail.com
In extreme floods modelling and forecasting, a large number of mathematical issues usually arise. Most of them are associated with solution of the incorrect inverse problem, which should be more or less efficiently solved for identifying the applied model parameters. Efficient parameters are those providing acceptable accuracy of floods forecast that allow undertaking various preventive measures to decrease or even avoid significant losses.

Let us consider in details such issues as

- high uncertainty of the model input data and data used for evaluation of the model output;
- selection of the effective and efficient forecasting model;
- selection of an objective function;
- developing objective methods of smoothing a response surface;
- selection of a simple yet efficient optimization procedure, which can be applied in conjunction with methods of handling the above issues.

The first issue in our list is high uncertainty of the observational data obtained from the national hydrometeorological networks. Often, such data are quite inconsistent or have irregular spatial and temporal resolution; stochastic structure of errors associated with data collection and processing also can vary in time and space; even in ”successful” countries, available data sets are incomplete in terms of their ability to satisfy our understanding of the modeled hydrometeorological processes. This issue generally means that hydrologists cannot use physically and mathematically rigor models that consist of some number of differential equations, because these equations cannot be supplied with necessary data. Alternatively, hydrologists have to use so-called conceptual model, which describe the modeled process less rigorously, however, they can be forced by using available data. Thus, selection of an appropriate model, in fact, is searching for a compromise between our understanding of the modeled process and ability to obtain something useful from available observations.

That is why correct selection of the best forecasting model, which can guarantee that issued forecasts are accurate enough, becomes a serious problem. Formally, the most efficient model should be that providing the highest success rate (say, in terms of Nash-Sutcliff criteria [2]) estimated for an independent validation period. However, nobody can guarantee that such a model will perform well in case of natural disasters (even if some disasters were involved in the model calibration). Our experience shows that the most reliable and robust model should be efficient for the entire spectrum of possible scenarios of the modeled process. In this case, the model robustness means that structure of the considered model and its parameters should not be too sensitive to the forcing data.

Selection of an appropriate objective function represents another problem. In general, a hypothetically ”perfect” objective function should provide the absolute fit of the modeled and actual (observed) processes, what never happens due to model and data uncertainties. Thus, we do not pretend to get the perfect fit of two or more plots and agree to get some compromise. It’s quite interesting, that matrices of losses caused by incorrect forecasts are, perhaps, the worst objective function for the model calibration and, certainly, the best criterion to evaluate the model performance.
Uncertainty in the model and data and various disadvantages of the selected objective function make the response function’s bottom extremely irregular. The forecasting model calibration becomes a quite rough attempt to find out a unique solution of an essentially incorrect inverse problem. Often, hydrologists try to smooth the response surface by applying various ways of regularization, penalty functions etc. After all, the modified surface becomes more or less smooth, yet quite useless, because any physical sense of the response surface is lost.

Finally, regardless of our success in making the response surface more smooth and regular, we still need to find out a physically correct optimum parameter set, and now we encounter another problem: most of hydrological models have approximately 8-20 parameters, hence, all the said above was about multidimensional parameter space. If so, any, even quite simple procedure becomes more complicated, because even simple algorithms now require tremendous processor resources. In addition, local minima are not stable in time (to imagine the response surface behaviour, we often use a visual model of a "floating udder").

In our opinion, all the described problems cannot be resolved separately. Let us describe several innovations which allow us to reach certain success in handling them.

First, we select a model or an optimum set of \( N \) parameters by using indices of \( F \)-robustness, which reflect relative stability of the local optimum's neighbourhood. In general, \( F \)-index is defined as a \((N + 1)\)-dimensional integral of the response surface in the interval \([P_i - r; P_i + r]\), (where \( P_i \) denotes an examined parameter set). In the simplest case, \( N \)-dimensional \( F \)-index \( F_N \) is the average value of an objective function in radius \( r \) around \( P_i \). \( F_n \) can be found by using the following equation:

\[
F^{n,r} = \int_{P-r}^{P+r} J dP,
\]

where \( J \) denotes the objective function and \( r \) denotes radius of the objective function averaging (F-radius). In practical tasks, when function \( J(P) \) is discrete, index \( F^n \) can be interpreted as the average value of \( J(P) \) within a given distance from the inspected parameter set (for example, 1, 2, 3 or more steps in each direction):

\[
F^n = \frac{1}{n} \sum_{i=1}^{n} \left[ \sum_{j=1}^{s+1} \frac{J_{i,j}}{(2s + 1)} \right],
\]

where \( s \) denotes a number of steps made in each direction, and \( J_{i,j} \) is the objective function for \( n \) parameters and \( s \)-step radius around each of them. In this case, any small shifts of "pits” do not cause significant worsening of corresponding objective function values. As for the model selection, the best model should have the smallest F-index for the effective region of used parameters. If an alternative or potentially suitable model or parameter sets have a "deeper" optimum, but greater F-index, they should be declined (Figure 1).

Second, we perform "natural” smoothing of the response surface by using multi-scale objective functions (MSOF). The particular objective function used in this work has the
following form:

\[
J = \sqrt{\sum_{k=1}^{n} \left( \frac{\sigma_1}{\sigma_k} \right)^2 \sum_{i=1}^{m_k} \left( q_{o,k,i} - q_{s,k,i}(X) \right)^2},
\]

(2)

where \(q_{o,k,i}\) and \(q_{s,k,i}\) denote the observed and simulated flows averaged over time interval \(k\) (i.e. the \(k\)-th aggregation scale), \(\sigma_k\) denotes the standard deviation of discharge at that scale, \(n\) denotes the total number of scales used, and \(m_k\) is the number of ordinates at the scale \(k\). In this work, we used hourly, daily, weekly and monthly scales corresponding to \(k = 1, 2, 3\) and 4, respectively. A value of \(k\) may not be restricted and can be assigned any number of scales (in this case, we get so-called all-scales objective function (ASOF). Note in Eq. (2) that the weight associated with each term is given by the inverse of the standard deviation of the flow at the respective scales. This weighting scheme assumes that the uncertainty in modeled streamflow at each scale is proportional to the variability of the observed flow at that scale. Another important motivation for using the multi-scale objective function (MSOF) is that it smoothes the objective function surface, and hence reduces the likelihood of the search getting stuck in tiny ’pits’ [1].

Finally, after all the described procedures, we apply a quasi-local optimization in a physically predetermined (or predefined) region of the parameter space by using a step-wise line search algorithm (SLS [1]). If the model input data are very uncertain, an enhanced configuration of the SLS algorithm is used: we apply a random generator to produce a number of ensembles of the input time series and perform model calibration for all the scenarios. The best (unbiased) parameters are those corresponding the smallest MSOF value.

The described measures allow us significant reducing of the mentioned issues impact and successful implementing the described approach for extreme floods forecasting in data scarce regions. The performed research was supported by the Federal Purpose-Oriented Program of the Russian Federation (Project P1103 ”Developing technologies of the catastrophic river floods forecasting in NW Russia and making decisions”).

References


Figure 1: Illustration of robust and non-robust phases: if the added parameter provides a "deeper" optimum, but greater $F$-index, it should be declined, because the model becomes more sensitive to the model input data.
Paper F

The Inverse Problem for the Wave Equation and Two Acquisition Geometries

Authors: M. Lassas and L. Oksanen
The Inverse Problem for the Wave Equation and Two Acquisition Geometries

Matti Lassas and Lauri Oksanen

Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland.
E-mail: {matti.lassas, lauri.oksanen}@helsinki.fi
1 Introduction

We consider the inverse problem of determining the spatially varying wave speed inside a domain from measurements of acoustic pressure wave fields on the boundary of the domain. The measurements are modelled as an operator on the wave fields between various source-receiver pairs. The arrangement of the source-receiver pairs is called the acquisition geometry.

In seismic imaging, the wave fields are often stimulated by explosions, [25, 23]. In such a case receivers need to be far away from the sources. For acquisition geometries commonly employed in seismic imaging see [22].

We study two acquisition geometries with disjoint sources and receivers, and show that the inverse problem for the wave equation with these acquisition geometries has unique solution.

Inverse problems where data is given on a part of the boundary are widely studied, the Calderón’s inverse problem for the conductivity equation and the closely related inverse problem for the Schrödinger equation being the paradigm problems, [6, 24, 21, 1], [4, 9].

The inverse problem for Schrödinger equation on a bounded domain of $\mathbb{R}^n$, $n \geq 3$, with receivers on an open subset $\Gamma$ of the boundary is solved in [17]. The inverse problem for Schrödinger equation on a bounded domain of $\mathbb{R}^n$, $n = 2$, with sources and receivers on the same open subset $\Gamma$ of the boundary is solved in [10]. The corresponding problem on a compact Riemannian surface is solved in [8]. For related results with measurements on a part of the boundary, see [5, 7, 11].

The inverse problems for the wave equation and for more general hyperbolic equations are studied in [2, 3, 20, 12, 15]. The inverse problem for the wave equation with sources and receivers on the same open subset $\Gamma$ of the boundary is solved in [13], see also [18, 16].

2 The Inverse Problem for the Wave Equation

Let $M$ be a bounded, connected and open set in $\mathbb{R}^n$, $n \geq 2$, with $C^\infty$-smooth boundary $\partial M$, and let $c$ be a strictly positive $C^\infty$-smooth function on $M$. Let us consider an inverse problem for the wave equation

$$
\partial_t^2 v - c(x)^2 \Delta v = 0 \quad \text{in } M \times (0, \infty),
$$

$$
v|_{\partial M \times (0, \infty)} = f,
$$

$$
v|_{t=0} = \partial_t v|_{t=0} = 0.
$$

Denote by $v^f(x, t) = v(x, t)$ the solution of (1) for $f \in C_0^\infty(\partial M \times (0, \infty))$, and define the hyperbolic Dirichlet-to-Neumann operator

$$
\Lambda : C_0^\infty(\partial M \times (0, \infty)) \to C^\infty(\partial M \times (0, \infty)), \quad \Lambda f := -\partial_\nu v^f|_{\partial M \times \mathbb{R}_+},
$$

where $\partial_\nu$ is the Euclidean normal derivative on $\partial M$. Denote by $\Lambda_{T_1, T_2}^\Gamma$ the restriction of the Dirichlet-to-Neumann operator

$$
\Lambda_{T_1, T_2}^\Gamma : C_0^\infty(\Gamma_1 \times (0, T)) \to C^\infty(\Gamma_2 \times (0, T)), \quad \Lambda_{T_1, T_2}^\Gamma f := \Lambda f|_{\Gamma_2 \times (0, T)},
$$
where $\Gamma_1, \Gamma_2 \subset \partial M$ are open. Furthermore, denote $\Lambda_{\Gamma_1, \Gamma_2} := \Lambda_{\Gamma_1, \Gamma_2}^\infty$.

Here $\Lambda_{\Gamma_1, \Gamma_2}^T$ for a certain collection $G$ of pairs of nonempty open sets $(\Gamma_1, \Gamma_2)$ can be interpreted as a continuum idealization of measurements of acoustic waves. The collection $G$ defines the acquisition geometry. The inverse problem for the wave equation (1) with the acquisition geometry $G$ is to determine the wave speed $c(x), x \in M$, given the domain $M$ and the data

$$\Lambda_{\Gamma_1, \Gamma_2}^T, \quad (\Gamma_1, \Gamma_2) \in G. \quad (2)$$

The operator $\Lambda$ determines the wave speed $c$, [2], that is, if the acquisition geometry puts no restrictions on the inverse problem, the problem has a unique solution. Even if the wave speed is anisotropic, the Dirichlet-to-Neumann operator determines it up to a change of coordinates, [3]. Moreover, if there are sources and receivers operating simultaneously in the same place, that is if there is $\Gamma \subset \partial M$ such that $(\Gamma, \Gamma) \in G$, then the wave speed $c$ is uniquely determined, [13].

In [19] we considered the inverse problem for a related geometric problem with two acquisition geometries where no sources and receivers operate simultaneously in the same places. There, the inverse problem is studied on a Riemannian manifold $(M, g)$ and the operator $c^2 \Delta$ in equation (1) is replaced by the operator $\Delta_g + g$, where $\Delta_g$ is the Laplace-Beltrami operator on $(M, g)$ and the potential $g$ is a $C^\infty$-function on $M$. It was shown that the boundary data related to $\Delta_g + g$ determine $(M, g)$ up to a change of coordinates.

Using the above result, we will here prove the uniqueness for the inverse problem for the equation (1). This is done by showing that the data (2) determines the related data for a gauge equivalent operator of the type $\Delta_g + g$ and by observing that the obtained Riemannian manifold $(M, g)$ can be embedded in a unique way in the domain $M \subset \mathbb{R}^n$.

3 Two Acquisition Geometries with Disjoint Sources and Receivers

We consider here only the case with $T = \infty$. The finite time case studied in [19] can be adapted for the equation (1) in a similar manner. In the two theorems below the set $M$ is fixed and known.

**Theorem F.1**

If $(\Gamma_1, \Gamma_2) \in G$ and $\Gamma_1 \cap \Gamma_2 \neq \emptyset$, then the data (2) and $c|_{\partial M}$ determine the wave speed $c(x), x \in M$, uniquely.

**Theorem F.2**

Let $(\Gamma_1, \Gamma_2), (\Gamma_1, \Gamma_3), (\Gamma_2, \Gamma_3) \in G$. Then the data (2) and $c|_{\partial M}$ determine the wave speed $c(x), x \in M$, uniquely.

Let us describe the proof Theorem F.1. We may write the operator $c^2 \Delta$ in equation (1) as a weighted Laplace-Beltrami operator

$$c^2 \Delta v = \mu^{-1} |g|^{-1/2} \sum_{j,k=1}^n \frac{\partial}{\partial x^j} \left( \mu |g|^{1/2} g^{jk} \frac{\partial}{\partial x^k} v \right),$$
where the Riemannian metric
\[ g := (g_{jk})_{j,k=1}^{n} = (c^{-2} \delta_{jk})_{j,k=1}^{n} \]

in the coordinates of \( \mathbb{R}^n \), the weight \( \mu = c^{n-2} \), \(|g|\) denotes the determinant of \( g \) and \((g^{jk})_{j,k=1}^{n}\) is the inverse of \( g \).

Let us denote \( \kappa := c^{(n-2)/2} \). The gauge transformation of \( c^2 \Delta \),

\[ Av := \kappa c^2 \Delta (\kappa^{-1} v), \]

is of form \( \Delta_g + q \) by [14] [KKL, pp. 257-258]. Let us denote the Dirichlet-to-Neumann operator corresponding to \( A \) by \( \widetilde{\Lambda} \). By [KKL, pp. 202-203],

\[ \widetilde{\Lambda} f = \kappa \Lambda (\kappa^{-1} f) + (\partial_{\nu} \kappa) \kappa^{-1} f. \]

As \( c|_{\partial M} \) and the restriction \( \Lambda_{\Gamma_1, \Gamma_2} \) are known, so are \( \kappa|_{\partial M} \) and the corresponding restriction of \( \widetilde{\Lambda} \).

Hence \((M, g)\) is determined as an abstract Riemannian manifold by Theorem 1 of [19]. This means that we know the metric \( g \) up to a change of coordinates. Using [KKL, Lemma 4.46 and the proof of Lemma 4.45] we obtain \( c \) in the original coordinates of \( \mathbb{R}^n \).

Here it is crucial to assume that \( c|_{\partial M} \) is known.

The proof of Theorem F.2 is similar with the above proof.

References


The Multiplicity of Positive Solutions for Boundary Value Problems to Quasilinear Equations: A Small Survey

Author: A. I. Nazarov
The Multiplicity of Positive Solutions for Boundary Value Problems to Quasilinear Equations: A Small Survey

Alexander I. Nazarov

Department of Mathematics and Mechanics,
St. Petersburg State University,
St. Petersburg, Russia.
E-mail: an@AN4751.spb.edu
We begin with the Dirichlet problem to a simple quasilinear equation in an annulus \( \Omega = B_{R_2} \setminus B_{R_1} \subset \mathbb{R}^n \):

\[
\begin{cases}
-\Delta_p u = u^{q-1} & \text{in } \Omega, \\
u = 0 & \text{on } \partial \Omega, \\
u > 0 & \text{in } \Omega,
\end{cases}
\]

(here \( \Delta_p u \equiv \text{div} (|\nabla u|^{p-2} \nabla u) \) is \( p \)-Laplacian; we always assume that \( n \geq 2 \) and \( 1 < p < \infty \)). This equation describes, e.g., reaction-diffusion processes in chemistry and biology.

The classical case \( p = 2 \) corresponds to conventional linear diffusion.

The non-uniqueness of solutions to semilinear BVPs was investigated by many authors beginning from the classical paper [1]. The phenomenon of multiplicity was discovered by Coffman [2]. Namely, it turned out that, for \( n = 2 \), \( p = 2 \) and \( q > 2 \), sufficiently thin annulus admits arbitrary many different (rotationally nonequivalent) solutions of the problem (1). This result was generalized for \( n \geq 4 \) in [3] and for \( n = 3 \) (the most difficult case!) in [4]. We underline that for \( \Omega = B_R \) this phenomenon does not occur, the solution to (1) is unique, as it was established in a remarkable paper [5].

For arbitrary \( 1 < p < \infty \) the multiplicity in (1) was established in [6] for \( n = 2 \), in [7] for \( n \geq 4 \) and in a recent paper [8] for \( n = 3 \).

**Remark G.1 (Remark 1).** For the slow reaction \( (1 < q \leq p) \) there exists a unique solution to (1) for any bounded domain \( \Omega \!\!\!.\!\!\!

We set \( \epsilon := \frac{R_2}{R_1} - 1 \) and denote by \( p_n^* \) the Sobolev critical exponent in \( \mathbb{R}^n \):

\[
\frac{1}{p_n^*} = \frac{1}{p} - \frac{1}{n}.
\]

We introduce the so-called \((m, k)\)-decomposition of the space \( \mathbb{R}^n = (\mathbb{R}^m)^\ell \oplus \mathbb{R}^k \), where \( m \geq 2 \) and either \( k = 0 \) or \( k \geq m \). Denote by \( x = (y_1, \ldots, y_\ell; z) \) points in \( \mathbb{R}^n \); here \( ys \) belong to \( \mathbb{R}^m \)'s while \( z \in \mathbb{R}^k \). Note that a nontrivial decomposition \((m < n)\) is admissible only if \( n \geq 4 \).

**Theorem G.1 ([7])**

Let \( 1 < p < q < \infty \). Then there exists \( \hat{\epsilon} \) depending only on \( p \) and \( q \) such that for any \( \epsilon < \hat{\epsilon} \) and any \((m, k)\)-decomposition satisfying \( q < p_{n-m+1}^* \) there exists a weak solution of the problem (1) depending only on \( |y_1|, \ldots, |y_\ell|, |z| \). Moreover, for different pairs \((m, k)\) corresponding solutions are different (rotationally nonequivalent).

**Remark G.2 (Remark 2).** Note that \( p_{n-m+1}^* \geq p_n^* \), and strict inequality holds if \( p < n \). In particular, a radial solution to (1) exists for all \( q < \infty \). We recall that for \( \Omega = B_R \) the problem (1) has no solution if \( q \geq p_n^* \), see [9] for \( p = 2 \).

Theorem 1 does not give multiplicity since the number of different \((m, k)\)-decompositions of \( \mathbb{R}^n \) is limited. The next theorem provides multiplicity in the thin layer for \( n \neq 3 \).

**Theorem G.2 ([7])**

Let \( 1 < p < \infty, p < q < p_n^* \). For any \( t_0 \in \mathbb{N} \) there exists \( \hat{\epsilon} \), depending only on \( p, q \) and \( t_0 \), such that for any \( \epsilon < \hat{\epsilon} \), for all \((2, k)\)-decompositions and for all \( 2 \leq t \leq t_0 \), the problem...
(1) has a weak solution \( \hat{u}_{2,k,t}(y_1, \ldots, y_\ell; |z|) \) symmetrical with respect to arguments \( y_j \) and \( 2\pi/t \)-periodic with respect to polar angle of \( y_j \in \mathbb{R}^2 \). Moreover, for different pairs \((k,t)\) corresponding solutions are different. The only exception may be \( t = 2, t' = 4, 2k' - k = n \).

Corresponding result for \( n = 3 \) is more difficult to prove and reads as follows.

**Theorem G.3 ([8])**

Let \( n = 3, 1 < p < \infty, p < q < p^*_n \). For any \( t_0 \in \mathbb{N} \) there exists \( \hat{\varepsilon} \), depending only on \( p, q \) and \( t_0 \), such that for any \( \varepsilon < \hat{\varepsilon} \) and for all \( 2 \leq t \leq t_0 \), the problem (1) has a weak solution \( \hat{u}_{2,k,t}(x_1, x_2; |x_3|) \), \( 2\pi/t \)-periodic with respect to polar angle of \((x_1, x_2)\). Moreover, for different \( t \) corresponding solutions are different.

**Remark G.3 (Remark 3).** For even \( n \geq 4 \) S. Kolonitskii recently proved multiplicity of solutions to (1) in thin layer under assumption \( p < q < p^*_n \) only. These solutions have more complicated symmetries.

Now we discuss the ”complementary” effect of multiplicity for fixed layer thickness and large \( q \).

**Theorem G.4 ([7])**

Let \( n \) be even, \( n \leq p < \infty, t_0 \in \mathbb{N} \). There exists \( \hat{q} \) depending only on \( p, \varepsilon \) and \( t_0 \) such that for any \( \hat{q} < q < \infty \) and \( t \leq t_0 \) the problem (1) has a weak positive solution \( \hat{u}_{2,0,t}(y_1, \ldots, y_{n/2}) \) symmetrical with respect to arguments \( y_j \) and \( 2\pi/t \)-periodic with respect to polar angle of \( y_j \). If \( t \neq t' \) then \( \hat{u}_{2,0,t} \) is non-equivalent to \( \hat{u}_{2,0,t'} \).

The results of Theorems 1, 2, 4 are generalized for the problem

\[
- \text{div } a(\nabla u) = f(u) \quad \text{in } \Omega, \quad u > 0 \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega,
\]

where \( a(\sigma) \equiv \nabla A(\sigma) \) and, roughly speaking, function \( A(\sigma) \) is ”not more convex” then \( |\sigma|^p \), and function \( F(s) = \int_0^s f(u)du \) is ”more convex” then \( s^p \).

Let us explain briefly the basic ideas of proofs. The multiplicity in the thin layer is proved by \textit{a priori} estimates. Namely, we consider the energy functional

\[
J(u) = \frac{\|\nabla u\|_{L_p(\Omega)}}{\|u\|_{L_q(\Omega)}}
\]

and show that energy levels split for sufficiently small \( \varepsilon \) (for \( n = 3 \) the method is more complicated). In Theorem 4 we proceed by the method of local bifurcations. Namely, we show that the second differential of energy functional is not positive for sufficiently large \( q \).

A. Shcheglova [10] transferred these results to nonlinear Neumann boundary value problem of the type

\[
\begin{align*}
-\Delta_p u + u^{p-1} &= 0 \quad \text{in } B_R \\
|\nabla u|^{p-2}\langle \nabla u; n \rangle &= u^{q-1} \quad \text{on } \partial B_1 \\
u > 0 &= u^{q-1} \quad \text{in } B_R.
\end{align*}
\]
describing the non-linear diffusion process with boundary reaction.

The multiplicity of solutions to (2) was proved:
a) \( n \neq 3, 1 < p < q < p^*_n \frac{n-1}{n} \), – for sufficiently large \( R \);
b) \( n \) even, \( n \leq p < \infty \), – for sufficiently large \( q \).

Now we consider the Dirichlet problem for the weighted equation

\[
\begin{align*}
-\Delta_p u &= |x|^{\alpha q} u^{q-1} \quad \text{in } B_1, \\
 u &= 0 \quad \text{on } \partial B_1, \\
 u &> 0 \quad \text{in } B_1.
\end{align*}
\]

(3)

This equation describes the process with reaction near the boundary. For \( p = 2 \) it is called the Hénon equation, see [11].

The non-uniqueness of solutions to (3) was established in [12] and independently (for \( p = 2 \)) in [13]. The multiplicity effect was proved in [14]:
a) \( n \neq 3, 1 < p < q < p^*_n \), – for sufficiently large \( \alpha \);
b) \( n \) even, \( \alpha > 0, n \leq p < \infty \), – for sufficiently large \( q \).

Thus, we can say that the multiplicity phenomenon can occur if the diffusion cannot equalize the density inhomogeneity generated by the reaction. It is the case if either the layer of reaction is sufficiently thin with respect to a typical size of the domain or the reaction is ”too fast” with respect to diffusion.

References


Paper H

Testing of Exponentiality with Application to Historic Data

Authors: Y. Nikitin and I. Piskun

Yakov Nikitin was supported by the grants RFBR 10-01-0154a and FZP 2010-1.1-111-128-033.
Testing of Exponentiality with Application to Historic Data

Yakov Nikitin* and Irina Piskun

Department of Mathematics and Mechanics,
St. Petersburg State University,
St. Petersburg, Russia.
*E-mail: yanikit47@gmail.com
1 A Review of Exponentiality Tests

Testing of exponentiality is very important in Probability and Statistics, especially in reliability theory, queuing theory and survival analysis. First tests of exponentiality appeared in the papers by Greenwood, Sherman, Moran and Epstein in the period from 1946 till 1953. Later appeared hundreds of papers and numerous surveys on the subject, and even some handbooks and reference books on exponential distribution.

Let $X_1, \ldots, X_n$ be a sample with df $F$. We are interested in testing the composite hypothesis $H_0 : F \in \mathcal{E}$, where

$$\mathcal{E} = \{ F : F(x) = 1 - e^{-\theta x}, \ \theta > 0, \ x \geq 0 \}$$

is the class of exponential distributions with unknown scale parameter $\theta$. The alternative $H_1$ consists in that $F \notin \mathcal{E}$. Often the class of alternatives is supposed to be narrower (IFR, DFR, NBU, NBUE, and other classes arising in statistical reliability theory.) Often the alternatives are particular parametric families like Gamma, Weibull, Makeham, LFR (linear failure rate), etc.

Consider some groups of tests for exponentiality.

I Tests using characterizations. Exponential law has more characterizations than any other law. Many tests are based on the "loss-of-memory" characterization. Some tests use other characterizations based on order statistics and spacings, on entropy properties, on empirical Laplace transform, on empirical characteristic function, etc.

II Normalized $L-$statistics. Consider the test statistics of the form

$$T_n = \frac{1}{nX} \sum_{i=1}^{n} w_{i,n} X_{(i)},$$

where $X_{(i)}$ are the order statistics, while $w_{i,n}, \ i = 1, \ldots, n,$ are appropriate weights. Examples are the Gini statistic (1914), the Jackson statistic (1967) and the Fortiana-Grané statistic (2002) which correspond to the special choice of weights.

III Scale-free tests based on ratios. Denote $U_i := X_i / \bar{X}, \ i = 1 \ldots n,$ and consider the statistics $V_n = n^{-1} \sum_{i=1}^{n} r(U_i),$ where $r$ is some real function on $R^+$. Particular cases are the Greenwood statistic (1946) for $r(x) = x^2$; the Moran statistic (1951) for $r(x) = \ln x$; the Cox-Oakes statistic (1984) for $r(x) = (1 - x) \ln(1 - x)$, etc.

IV Kolmogorov-type statistics. We mean the test statistics of the form $J_n = \sup_{x \geq 0} \left| \frac{1}{n} \sum_{i=1}^{n} r(U_i, x) \right|,$ where $r(s, x)$ is some function on $[0, 1]^2$. Examples are the tests of Lilliefors (1969), of Baringhaus-Henze (2000) and of Khmaladze (1979, 2008).

How to compare all these tests? How can we choose the best and most efficient test? How do these tests work for real data? In this talk we use the approach using the asymptotic efficiency of tests.

2 Efficiency of Tests

Suppose we have two competitive sequences of test statistics $\{T_n\}$ and $\{V_n\}$ corresponding to the level $\alpha \in (0, 1)$. Denote $N_{T}(\alpha, \beta, \theta)$ and $N_{V}(\alpha, \beta, \theta)$ the sample sizes necessary to achieve the power $\beta \in (0, 1)$ under the alternative $\theta$. Define the relative
Table 1: Values of Bahadur efficiency.

<table>
<thead>
<tr>
<th>Test</th>
<th>Makeham</th>
<th>LFR</th>
<th>Weibull</th>
<th>Gamma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moran</td>
<td>0.694</td>
<td>0.388</td>
<td>0.943</td>
<td>1</td>
</tr>
<tr>
<td>Greenwood</td>
<td>0.75</td>
<td>1</td>
<td>0.607</td>
<td>0.388</td>
</tr>
<tr>
<td>Gini</td>
<td>1</td>
<td>0.75</td>
<td>0.876</td>
<td>0.694</td>
</tr>
<tr>
<td>Jackson</td>
<td>0.75</td>
<td>1</td>
<td>0.607</td>
<td>0.388</td>
</tr>
<tr>
<td>Lilliefors</td>
<td>0.538</td>
<td>0.356</td>
<td>0.607</td>
<td>0.503</td>
</tr>
<tr>
<td>Khmaladze</td>
<td>0.75</td>
<td>1</td>
<td>0.607</td>
<td>0.388</td>
</tr>
</tbody>
</table>

(finite-sample) efficiency as $ef_{TV}(\alpha, \beta, \theta) = N_V(\alpha, \beta, \theta)/N_T(\alpha, \beta, \theta)$. This quantity is too complicated to be found. Therefore one finds Pitman efficiency, Hodges-Lehmann efficiency and Bahadur efficiency as the limits of the finite-sample efficiency as $\theta \rightarrow \theta_0, \beta \rightarrow 1$ and $\alpha \rightarrow 0$.

The calculation of efficiencies is a hard task requiring the theorems on asymptotic normality and large deviations. Most of efficiencies for tests of exponentiality were found in 1995 - 2005 by Henze, Klar, Litvinova, Nikitin, Tchirina. Look at the numerical values of local Bahadur efficiency for 6 statistics and 4 alternatives to exponentiality. This type of efficiency is most adequate here as the asymptotic normality of test statistics is not required.

3 Treatment of Historical Data

Let apply the exponentiality tests to historical data. It is well-known that the life length of human beings cannot be exponentially distributed. Humans are aging, hence human lifetimes have an increasing failure rate, or force of mortality. As people age, they are more likely to die. In contrast, exponentially distributed variables have constant failure rates.

It seems that the situation with historical rulers, kings, governors, etc. would be similar: they would have stopped ruling or reigning as a consequence of accumulated controversies and tensions plus damages of economic, social, politic or personal nature, like it happens with human lives.

However, in the recent paper [1] the authors proclaimed that the lengths of rules of Roman Emperors are in agreement with exponential distribution. Hence we are led to the surprising conclusion that their reigns ceased suddenly, "purely in random" in unexpected and unpredictable ways.

According to the chronology by Kienast (1990), the reigns of 76 Roman Emperors are
as follows (in years): Augustus – 41,5; Tiberius – 22,6; Caligula – 3,8; Claudius – 13,6; Nero – 13,7, and so on till the Fall of the Roman empire in 476 AD.

Using 53 lengths for the "decline period" from Nerva till Theodosius (395 AD) and applying the Khmaladze test, the authors of [1] concluded that the hypothesis of exponentiality is accepted. They also found the same for 161 Chinese emperors’ from 221 BC to 1911 AD, 23 Spanish monarchs, starting with Isabella I in 1474, 24 Russian tsars, starting with Ivan IV the Terrible in 1547, and 22 British kings, from Henry VII in 1483 to George VI (1936-1952).

The authors of [1] propose the following explanation of this: "There are many challenges and threats for a reign, but there are also strong means to protect the ruler. Hence only exceptionally strong challenges could stop the reign. But in Probability, such occurrences form a Poisson process. Hence the inter-occurrences times would have an exponential distribution."

We have inspected the conclusions of [1] using numerous tests of exponentiality.

First we considered the data of the decline period (96 - 395 AD) from Nerva till Theodosius they have studied and have used exclusively the Khmaladze test. We found that some tests support the hypothesis of exponentiality at the level 0.05 while certain tests reject it. Tests which accept the hypothesis: Greenwood test, Gini test, Epps-Pulley test. Accept but scarcely: Lilliefors test. Reject: Moran test, Fisher test, chi-square test.

It is enough to reject the hypothesis of exponentiality. According to A. Einstein [2]: "No amount of experimentation can ever prove me right; a single experiment can prove me wrong."

In case we have tested the hypothesis of exponentiality for all 76 Roman emperors, we also have rejected the null hypothesis by means of Moran test, Gini test, Epps-Pulley test, Lilliefors test, Fisher test, chi-square test, but the Greenwood test accept this hypothesis.

Hence we disagree with [1]. The reason is that possibly Khmaladze test is not sensitive and efficient enough to discover the deviation from exponentiality. Another reason is that possibly the sample size is not large enough to apply the asymptotic formulas.

What happens with other dynasties of kings and emperors? It is an inexhaustible topic. We have explored the reign lengths of Israel kings. There exists the famous hypothesis ascending to Nikolai Morozov, Russian theoretician of terrorism, that Israel kings were in fact the emperors of Western Roman empire, that the Judean kings were the emperors of Byzantian empire and that the same sequence of rulers was counted twice by historians and chronologists. Note that the division of Israel in two kingdoms: Judah and Israel after the death of king Solomon is in agreement with the irrevocable division of Roman Empire in the Western Roman Empire and the Byzantine Empire after the emperor Diocletian.

From the Bible, we know the lengths of rules for the Kings of (undivided) Israel in order of reign (in years): Saul – 24 , David – 42, Solomon – 37. Kings of Israel in order of reign after division have the following lengths of rules: Jeroboam I – 21, Nadab – 1, Baasha – 23, ... till Hoshea – 11 before Assyrian captivity of Israel in 724 BC by Sargon II.

Here the data is not very reliable as it corresponds to the events occurring 3000 years ago, and the only source of it is the Bible. However, in all cases exponentiality is rejected, and the homogeneity hypothesis of two samples (Western Roman emperors - Israel kings) is not supported.

The whole topic needs the detailed study and the use of additional tests of exponen-
tiality.

References


More on the Kronecker Structured Covariance Matrix

Author: M. Ohlson, M. R. Ahmad and D. von Rosen
More on the Kronecker Structured Covariance Matrix

Martin Ohlson*, M. Rauf Ahmad* and Dietrich von Rosen†*

*Department of Mathematics,
Linköping University,
Linköping, Sweden.
E-mail: {martin.ohlson, dietrich.von.rosen}@liu.se, muahm@mai.liu.se

†Department of Energy and Technology,
Swedish University of Agricultural Sciences,
Uppsala, Sweden.
E-mail: dietrich.von.rosen@slu.se
1 Introduction

In this paper we consider estimation of a Kronecker structured covariance matrix of order three. The main goal is to extend the estimation procedure, suggested by [13], for the matrix normal distribution $\text{vec}X \sim N_{pq}(\text{vec}M, \Psi \otimes \Sigma)$ to the case where $\text{vec}\mathcal{X} \sim N_{pqr}(\text{vec}\mathcal{M}, \Theta \otimes \Psi \otimes \Sigma)$, with some vectorization $\text{vec}\mathcal{X}$ of the third order tensor $\mathcal{X} = (x_{ijk}) : p \times q \times r$, which will be defined in Section 2 and where $\otimes$ denotes the Kronecker product. We will say that the covariance matrix $D(\text{vec}X) = \Theta \otimes \Psi \otimes \Sigma$ is double separable (or three-factor separable) compared to the separable covariance matrix $D(\text{vec}X) = \Psi \otimes \Sigma$. The Kronecker product restrictions make the family of densities to be curved, i.e., it belongs to the curved exponential family.

Recently [12] have studied doubly exchangeable linear models, which are suitable for three-level multivariate data, and closely related to double separability. Doubly exchangeable covariance structure assumes a block circulant covariance structure consisting of three unstructured covariance matrices for three multivariate levels.

Several authors, see for example [9, 10, 5, 7, 8, 13], considered estimation and testing under the separability assumption. [13] discussed estimability of the parameters under the separability assumption. From the likelihood function, constructed of independent observation matrices, [13] proved that the maximum likelihood estimates under the restriction $\psi_{qq} = 1$, where $\Psi = (\psi_{ij}) : q \times q$ are found by an iterative flip-flop algorithm. [13] also showed that the likelihood equations provide unique estimators. A similar algorithm has been suggested by [6, 2, 1] but without the restriction $\psi_{qq} = 1$.

In many applications, different structures of the covariance matrices have been discussed. In [10, 11, 13] the intraclass covariance structure was considered and in [11] an autoregressive structure hold.

Also a structure on the mean has been considered. In [14] the growth curve model for the mean $M = ABC$, where $A : p \times s$ and $C : t \times q$ are known design matrices and $B : s \times t$ is the parameter matrix, was assumed. Under the restriction $\psi_{qq} = 1$ and some full rank assumption unique estimators for $B, \Sigma, \Psi$ were derived.

2 Model

Let $\mathcal{X}$ be a tensor of order three, with the dimension $p, q$ and $r$ in the $x$, $y$ and $z$ direction, respectively, see Figure 1. If $r = 1$ we have a special case with the tensor equal to a $p \times q$ matrix. For such a matrix $X = (x_1, \ldots, x_q) : p \times q$ the standard way to vectorize is as

$$\text{vec}X = (x'_1, \ldots, x'_q)'.$$

Vectorization of the three dimensional tensor $\mathcal{X}$ can be done in several ways. Let us use the following definition.

**Definition I.1.** Let $\mathcal{X} = (x_{ijk}) : p \times q \times r$ be a three dimensional tensor. Define the vectorization of $\mathcal{X}$ as

$$\text{vec}\mathcal{X} = \sum_{i=1}^{p} \sum_{j=1}^{q} \sum_{k=1}^{r} x_{ijk} e^3_k \otimes e^2_j \otimes e^1_i,$$

where $e^3_k$, $e^2_j$ and $e^1_i$ are the unit basis vectors of size $r$, $q$ and $p$, respectively.
\[ X = \begin{pmatrix} x_{11k} & \ldots & x_{1qk} \\ \vdots & \ddots & \vdots \\ x_{p1k} & \ldots & x_{pqk} \end{pmatrix} \]

\( k = 1, \ldots, r \)

**Figure 1:** The box visualizes a three dimensional data set as a third order tensor.

We will assume that the vectorization of \( X \) follows a multivariate normal distribution with a double separable covariance matrix

\[ D(\text{vec} X) = \Theta \otimes \Psi \otimes \Sigma, \]

with \( \Sigma : p \times p, \Psi : q \times q \) and \( \Theta : r \times r \), assumed to be positive definite. This structure is a generalization of the separable covariance matrix discussed by, e.g., [2, 5, 13].

If we use Definition 2.2.3 in [4] we can write the double separable model for \( X \) (or \( \text{vec} X \)) as

\[ \text{vec} X = \sum_{ijk} \mu_{ijk} e^3_k \otimes e^2_j \otimes e^1_i + \sum_{ijk} \sum_{i'j'k'} \varsigma_{ii'} \tau_{jj'} \upsilon_{kk'} u_{i'j'k'} e^3_k \otimes e^2_j \otimes e^1_i, \]

where \( M = (\mu_{ijk}) : p \times q \times r, \Sigma = \varsigma \varsigma' , \Psi = \tau \tau' \) and \( \Theta = \vartheta \vartheta' \) and \( u_{i'j'k'} \sim N(0,1) \), iid (independent and identically distributed). The density of \( X \) can now be written

\[ (2\pi)^{-pqr/2} |\Theta|^{-pq/2} |\Psi|^{-pr/2} |\Sigma|^{-qr/2} \exp \left\{ -\frac{1}{2} \text{vec}'(X - M)(\Theta \otimes \Psi \otimes \Sigma)^{-1} \text{vec}(X - M) \right\} \]

and is denoted

\[ X \sim N_{p,q,r}(M, \Sigma, \Psi, \Theta). \]

For more details about the multilinear normal distribution (1) see [4, p. 215].

## 3 Estimation

All the parameters of \( \Sigma, \Psi \) and \( \Theta \) in the covariance matrix \( D(\text{vec} X) = \Theta \otimes \Psi \otimes \Sigma \) are not uniquely defined. Several authors have discussed this for a separable covariance matrix \( D(\text{vec} X) = \Psi \otimes \Sigma \), e.g., [3, 9]. The parametrization problem is related to the fact that

\[ \Psi \otimes \Sigma = (e^t \Psi) \otimes \left( \frac{1}{c} \Sigma \right) \]
and this leads to estimability problems. Recently, [13] also considered the problem and suggested, without any loss of generality, to set $\psi_{qq} = 1$. For a double separable covariance matrix we have a similar problem since

$$
\Theta \otimes \Psi \otimes \Sigma = \left( \frac{1}{ab} \right) \otimes (aPsi) \otimes (bSigma).
$$

In this case to get an unique parametrization, without any loss of generality and similar to [13], we suppose $\Sigma: p \times p$ to be unstructured, $\Psi = (\psi_{ij}): q \times q$ with $\psi_{qq} = 1$ and $\Theta = (\theta_{ij}): r \times r$ with $\theta_{rr} = 1$.

Now, assume that we have $n$ independent observations $X_{jl}: p \times q \times r, j = 1, \ldots, n$, from (1). One can easily see that $M: p \times q \times r$ will be estimated by averaging. Hence, in the subsequent without any loss of generality, we may put $M = 0$.

Furthermore, with $M = 0$ the likelihood function for $\Sigma, \Psi$ and $\Theta$ is proportional to

$$
|\Theta|^{-pqn/2} |\Psi|^{-prn/2} |\Sigma|^{-qrn/2} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n} \text{vec}' X_{jl} (\Theta \otimes \Psi \otimes \Sigma)^{-1} \text{vec} X_{jl} \right\}.
$$

(2)

The following theorem can now be stated.

**Theorem I.1**

The likelihood equations that are maximizing the likelihood function (2) under the conditions $\psi_{qq} = 1$ and $\theta_{rr} = 1$ are given by

$$
\hat{\Sigma} = \frac{1}{qrn} \sum_{l=1}^{n} X_{l}' \left( \hat{\Theta} \otimes \hat{\Psi} \right)^{-1} X_{l},
$$

(3)

$$
\hat{\Psi} = \frac{1}{prn} \sum_{l=1}^{n} Y_{l}' \left( \hat{\Sigma} \otimes \hat{\Theta} \right)^{-1} Y_{l},
$$

(4)

$$
\hat{\Theta} = \frac{1}{pqn} \sum_{l=1}^{n} Z_{l}' \left( \hat{\Psi} \otimes \hat{\Sigma} \right)^{-1} Z_{l},
$$

(5)

where

$$
X_{l} = \sum_{ijk} x_{ijk}^{(l)} (e_{k}^{3} \otimes e_{j}^{2}) (e_{i}^{1})' : (qr) \times p,
$$

$$
Y_{l} = \sum_{ijk} x_{ijk}^{(l)} (e_{i}^{1} \otimes e_{k}^{3}) (e_{j}^{2})' : (pr) \times q,
$$

$$
Z_{l} = \sum_{ijk} x_{ijk}^{(l)} (e_{j}^{2} \otimes e_{i}^{1}) (e_{k}^{3})' : (pq) \times r.
$$

Note that the likelihood equations (3)-(5) are nested and there exist no explicit solution. Thus, we can solve (3)-(5) using the so called flip-flop algorithm.
References


Paper J

The Inverse Problem, EEG Data and Multivariate Statistics

Authors: M. Ohlson and D. von Rosen
The Inverse Problem, EEG Data and Multivariate Statistics

Martin Ohlson* and Dietrich von Rosen†*

*Department of Mathematics,
Linköping University,
Linköping, Sweden.
E-mail: (martin.ohlson, dietrich.von.rosen)@liu.se

†Department of Energy and Technology,
Swedish University of Agricultural Sciences,
Uppsala, Sweden.
E-mail: dietrich.von.rosen@et.slu.se
1 Introduction

In applied mathematical modeling the object is to obtain information via data about a model reflecting "Nature"/"Society". It is not clear that it is possible but in this article our starting point is a belief that this is a fruitful path to walk along, i.e., once stated a model data will tune the model and increase the information. For example, via some mapping of past observations into a model we can increase the prediction accuracy of the model, i.e., increase the information. The aim of the article is to show how to merge ideas from applied mathematics with ideas from statistical inference. To illustrate the ideas EEG signals recorded on the scalp from depressed patients will be used, showing the advantage of combining applied mathematics knowledge and tools with advanced statistical techniques. However, other types of time series, in particular series connected to environmental problems could very well have been considered. Different schools think differently about data and their applications. Often applied mathematicians think about data as a representation of some infinite set of ideal noise free data which for some reasons has been "polluted" with noise (mostly deterministic but also stochastic noise exists). Statisticians on the other side look upon data as representations of random variables which follow a distribution. Indeed a statistical model is nothing but a statement about a distribution. In the statistical world there is usually a finite number of observations and usually analysis is heavily dependent on the number of observations.

![Figure 1: Showing the variability and complexity in an EEG serie obtained after treating depressed patients with ECT](image)

It is very difficult to find distributions governing stochastic (statistical) models which describe complex dynamic processes such as for example EEG signals. An example is presented in Figure 1. Biological systems and processes are often neither pure deterministic nor pure stochastic. Applied mathematicians may focus on existence, uniqueness, construction of solutions, and their stability when there are disturbances in the data. Statisticians, according to the statistical paradigm, are more interested in designing experiments, discussing various estimation approaches, studying concepts such as unbiasedness, smallest variance, consistency, significance tests and distributions of suitable statistics (often only asymptotically) and last but not least develop methods for validating models. The concept of independent/non-independent observations is also crucial to statisticians. It is
clear that the two schools are complementary and we believe that because of increased complexity in measurements it is beneficial to jointly develop a new research area.

Compared to today's "engineering practise" our goal is that in this project the statistical part will be stronger and at a higher level whereas at the same time the mathematics in the analysis of data is much more advanced than usually is applied in statistical data analysis.

2 Assessing Local Complexity of EEG Signals

Nonlinear dynamics includes the theory of deterministic chaos. Typical to chaotic systems are that they show predictability within a short-time-period but nonpredictability in a long-time-period perspective because of senisitivy to initial conditions. The EEG series we are studying share this property, i.e. predictivity exist within a short time range but not for longer time periods. Moreover, EEG series are non-stationary and on a "micro level" errors seem to be non-random. The only error which really matters is a model error which then, unfortunately, technically introduces strong dependencies (e.g. autocorrelations). Hence, to use classical statistical time series analysis in order to investigate EEG signals may be questioned. Indeed, since it is difficult to formulate a statistical model any probability statement can be questioned. All this leads to that deterministic chaos approaches may be adopted to summarize the information in the EEG series. In this article we are going to use Higuchi’s fractal dimension, $D_f$, see e.g. Klonowski (2007), [1]. It is the fractal dimension of the curve, $L(k)$, representing amplitude of the signal under consideration in a plane as a function of time. If the ‘length’ of this curve (see below) scales as

$$L(k) \propto k^{-D_f}$$

the curve is said to have fractal dimension $D_f$. Fractal dimension will be taken as a basis for the subsequent statistical evaluation. Since a simple curve is of dimension 1 and a plane is of dimension 2 $D_f$ is always between 1 and 2. The Higuchi’s fractal dimension measures the filling of a plane over time which also may be viewed as a measure of complexity. In general, a huge number of data points are summarized in an interpretable way, i.e. as estimators of the curve’s fractal dimension. The procedure of estimating the fractal dimension may be done in a sliding time window of as few as 100 data points, giving a running fractal dimension $D_f(t)$ that shows changes of the signal’s complexity. Higher values of $D_f(t)$ correspond to the presence of higher frequencies in the signal’s Fourier spectrum. Consider the time sequence $x(1), x(2), \ldots, x(N)$ from which we construct $k$ new time series

$$x_m^k = \{x(m), x(m+k), \ldots, x(m + [(N - m)/k]k)\}, \quad m = 1, 2, \ldots, k,$$

where $k$ also is the time between consecutive observations and $\lfloor \cdot \rfloor$ is the integer function. For each of the constructed time series, the length of the curve is estimated via an average length

$$\hat{L}_m(k) = \frac{(N - 1) / k \sum_{i=1}^{\lfloor(N-m)/k\rfloor} (x(m + ik) - x(m + (i - 1)k))}{[(N - m)/k]k}$$
and then the total length is estimated by

\[ \hat{L}(k) = \frac{1}{k} \sum_{m=1}^{k} \hat{L}_m(k). \]

Using (1) we get a log-log relation to determine the unknown \( D_f \)

\[ \ln \hat{L}(k) = \alpha - D_f \ln(k) + \text{error}(k), \quad k = k_{\text{min}}, \ldots, k_{\text{max}}. \]

The values \( k_{\text{min}} \) and \( k_{\text{max}} \) have to be chosen in advance, e.g. 2 and 8. As most others we will apply the least squares approach. This means that the error is assumed to be the same, with mean 0, for all \( k \) and that its distribution is symmetric. This can easily be checked by inspection of residuals and if it does not hold alternative estimation approaches can be utilized. However, one of the main advantages with the least squares approach is that it is computationally straight forward.

### 3 Multivariate Statistical Analysis

Our starting point is the MANOVA model

\[ X = BC + E, \quad (2) \]

where \( X: p \times n \), \( B: p \times k \), \( C: k \times n \) \( E \sim N_{p,n}(0, \Sigma, I) \) and \( \Sigma \), supposed to be p.d., are unknown parameters and \( N_{p,n}(0, \Sigma, I) \) denotes the matrix normal distribution where the columns of \( E \) are independent and the dispersion of each column equals \( \Sigma \). Moreover, mostly one is interested in testing the hypothesis \( BF = 0 \), where \( F \) is known, but also hypotheses of the form \( BF = 0 \) may be considered. The well known growth curve model extends (2) and is given by

\[ X = ABC + E, \quad (3) \]

where the matrices are the same as in (2) but \( A: p \times q \) is a new known within individuals design matrix. For references and techniques how to analyze (3) see Kollo & von Rosen (2005), [2].

The data which is to be analyzed consists of 28 individuals from 3 diagnostic groups which all have received ECT (electroconvulsive therapy). For each individual two EEG series (FP1 and FP2) from the scalp are obtained. Each series comprises about 9000 observations. However, as also can be seen from Figure 1, the series comprises 4 phases. For each phase and within each series \( \hat{D}_f \) is calculated. Thus, for each individual there are 8 correlated observations which will be analyzed by either (2) or (3). Hypothesis which are of interest are for example

- \( H_0^1 \): The 3 diagnostic groups differ
- \( H_0^2 \): The FP1 and FP2 series differ
- \( H_0^3 \): It is only for one diagnostic group where FP1 and FP2 series differ
References


Comparison of Two Measurement Models of Convolution Type

Author: P. Piironen
Comparison of Two Measurement Models of Convolution Type

Petteri Piirinen

Department of Mathematics and Statistics,
University of Helsinki,
Helsinki, Finland.
E-mail: petteri.piirinen@helsinki.fi
1 Introduction

We are interested in determining how much information can be inferred from a pair of linear models

\[
\begin{align*}
  m_1 &= \varepsilon_1 \ast \sigma + \xi_1, \\
  m_2 &= \varepsilon_2 \ast \sigma + \xi_2
\end{align*}
\] (1)

and how much information is lost if we just use one the models. The equation of form

\[
m = \varepsilon \ast \sigma + \xi
\] (2)

represents a generic measurement model of convolution type. The symbol \(m\) stands for the measurement, the symbol \(\sigma\) represents the unknown and \(\xi\) denotes the noise. In this model, we assume that the object of interest \(\sigma\) and the outcome \(m\) of the measurement experiment are related in the ideal conditions via convolution with a function \(\varepsilon\). However, in real measurement experiments the measurements are corrupted by noise. The noise can originate from the measurement device, the unknown itself or it can be coming from an external source. We assume that the noise \(\xi\) and the unknown \(\sigma\) are statistically independent. This assumption dictates that we have a statistical interpretation of the model (2).

The main question in relation to the pair of models (1) is

**Question** Can we characterise when \(m_1\) contains more information of \(\sigma\) than \(m_2\)?

The choice of analysing convolution type models is not arbitrary. The motivation for studying this type of measurement models comes from the radar measurement experiments. Ideally, the radar transmits coded signals which we may assume to be beam-like objects which are representable by one-dimensional functions. The transmitted signal scatters from the object of interest and the receiver antenna captures the scattered response. The physical model can be simplified to a convolution type equation (2), where \(\varepsilon\) is obtained from the code with a simple deterministic transformation. We used this model as the underlying physical model in [4]. We wanted to know what kind of coded signals we should transmit in order to get most realistic reconstructions of the unknown \(\sigma\). We note that this problem can be reduced to the main question.

Before such a question can be studied we have to formalise the meaning of the measurement model. In [4] we approached the question within the statistical inverse theory framework i.e. we used the infinite-dimensional Bayesian approach. We assumed that the measurement \(m\), the unknown \(\sigma\) and the noise \(\xi\) are all random variables with infinite-dimensional state spaces and the solution is the conditional distribution of \(\sigma\) given the measurements \(m\). In this framework we answered the question using the theory of comparison of statistical measurements (c.f. [7]).

To this end, we can state the results concerning the main question. Let \(m_1 \succeq_\sigma m_2\) denote the concept that \(m_1\) contains more information of \(\sigma\) than \(m_2\). Let us assume that \(\xi_1, \xi_2\) are complex Gaussian white noises with the same variance. Then we have

**Theorem K.1 (Sufficiency)**

If \(|\hat{\varepsilon}_1(t)| \geq |\hat{\varepsilon}_2(t)|\) for almost every \(t \in \mathbb{R}\), then \(m_1 \succeq_\sigma m_2\).

**Theorem K.2 (Necessity)**

If \(\hat{\sigma}\) is a complex Gaussian measure with strictly positive variance \(v\) and \(m_1 \succeq_\sigma m_2\), then \(|\hat{\varepsilon}_1(t)| \geq |\hat{\varepsilon}_2(t)|\) for almost every \(t \in \mathbb{R}\).
2 LeCam Theory and Comparison of Measurements

In the introduction, we formalised the measurement model with random variables and Bayesian framework. Another general formalisation is the statistical experiment formalism. We will introduce both these frameworks side by side.

Definition K.1. A statistical experiment \( E \) with a parameter set \( \Theta \neq \emptyset \) is a triple \( (X, \mathcal{X}, P_\Theta) \), where \( P_\Theta = \{P_\theta \mid \theta \in \Theta\} \) is a family of pr-measures on the measurable space \( (X, \mathcal{X}) \). The set \( \Theta \) represents the possible states of the unknown and the set \( P_\Theta \) form the explaining theories.

Definition K.2 (Definition ([7])). A statistical measurement with the unknown \( x \) is a pair \( (m, x) \) of random variables with a state space \( (M \times X, M \otimes \mathcal{X}) \).

The former definition is from [9]. The latter definition is abstracted from the definition used in the unpublished manuscript (c.f. [5]). The preliminary versions of Theorems K.1 and K.2 were already in this manuscript with more restrictive assumption (see also [6]).

Lucien LeCam introduced the theory of approximate sufficiency (the deficiency) in 1960’s. This theory is nowadays usually known as the LeCam theory (c.f. [2, 3, 1, 9]). We do not, however, need the general frame of deficiencies here, so we just introduce the \( \succ \) relation in the case of statistical experiments.

Definition K.3. If \( E \) and \( F \) are two experiments with the same \( \Theta \), then \( E \succ_\Theta F \) iff for every \( \Theta_0 \subset \Theta \) finite, for every finite decision space \( T \), for every bounded loss function \( L \in L_s(T) \) it holds that

\[
\forall \sigma \in \mathcal{S}_T(F) \exists \rho \in \mathcal{S}_T(E) \forall \theta \in \Theta_0: R_E^L(L, \theta, \rho) \leq R_F^L(L, \theta, \sigma)
\]

where \( R_E^L(L, \theta, \rho) \) is the risk of the player’s \( E \) strategy (or decision) \( \rho \) with respect to loss \( L \) if the true state happens to be \( \theta \).

In the framework of statistical measurements the corresponding definition is

Definition K.4. We say that \( m_1 \succ_x m_2 \), if the original pr-space \( \Omega \) can be refined so that the triple \( (m_2, m_1, x) \) becomes weakly Markov. The triple \( (m_2, m_1, x) \) is weakly Markov, if there exists a r.v. \( \hat{m}_2 \) s.t. \( (\hat{m}_2, m_1, x) \) is Markov and \( (m_2, x) \sim (\hat{m}_2, x) \) are identically distributed.

The statistical measurement \( (m, x) \) gives rise to a statistical experiment in many ways but the following definition appears to be most fruitful.

Definition K.5 (Definition ([7])). Let \( (m, x) \) be a statistical measurement with state space \( (M \times X, M \otimes \mathcal{X}) \). We denote

\[ E_x(m) := \{M, M, \hat{m}^x(\Theta)\}, \text{ where } \hat{m}^x(f) : U \mapsto E(f(x) \mid m \in U)\}
\]

and \( \Theta = \{f \in L^1(X, \mathcal{X}, Pr_x) \mid \|f\|_1 = 1, f \geq 0\} \). We call the experiment \( E_x(m) \) the statistical experiment generated by the statistical measurement \( (m, x) \).

The \( \succ \) relation has a more analytic characterisation which can be used for analysing models with functional analytic methods.
**Theorem K.3 ([7])**

Suppose the state space of \((m_2, m_1, x)\) is, say, Suslin measurable space. Then \(m_1 \succ_x m_2\) iff there exists a Markov operator \(\Phi: M(m_2) \to M(m_1)\) so that \(\Lambda(x \rightsquigarrow m_2) = \Lambda(x \rightsquigarrow m_1) \circ \Phi\), where \(\Lambda(x \rightsquigarrow m_j): M(m_j) \to M(x)\) is

\[
(\Lambda(x \rightsquigarrow m_j)f) \circ x = E(f(m_j)|x) \quad \text{a.s.}
\]

and \(M(m_j)\) is the \(M\)-space of the experiment \(\mathcal{E}_x(m_j)\).

**3 Notes on the Continuous Model and Related Technical Issues**

In [4] the model was discrete convolution. This simplifies the technical issues a lot, but after this we wanted to know what happens in the continuous time. In that case, the noises are no longer processes but generalised processes. This part is still work in progress ([8]).

The generality we use here needs introduction of tempered complex Gaussian generalised random variables over the real line. This enables the effective use of Fourier transform techniques and the models in question may be reduced to models of multiplication type \(\hat{m} = \hat{\varepsilon} \hat{\sigma} + \hat{\xi}\). We analyse these models by studying the eigenfunctions of the conditional expectation operators \(\Lambda(\hat{\sigma} \rightsquigarrow \hat{m}_j)\). This is where the Theorem K.3 may be used effectively. We also note that the main difficulties arise from the zeros of the multiplier functions and the degeneracy of the prior beliefs.

In one direction (Necessity) the zeros of the multipliers have only rôle of complicating techniques but in the other direction (Sufficiency) the degenerate parts can be incorporated to the zeros of multipliers and at these places the comparison of the moduli of the multipliers \(\hat{\varepsilon}_j\) breaks down.

However, the intuitive argument does not seem to reach the subtleties of the problem and therefore the sufficient condition is not the optimal one.

**References**


Inverse Scattering Problems that have Applications in 2D Non-linear Optics

Author: V. Serov
Inverse Scattering Problems that have Applications in 2D Non-linear Optics

Valeriy Serov

Department of Mathematical Sciences,
University of Oulu,
Oulu, Finland.
E-mail: valeri.serov@oulu.fi
1 Formulation of the Problem

The time-independent part $u$ of the electromagnetic field (for time-harmonic dependence) in the Kerr-like nonlinear media can be described generally by the nonlinear Schrödinger equation in $\mathbb{R}^2$ of the form

$$-\Delta u(x) + q(x)u(x) + \frac{\alpha(x)|u(x)|^2}{1 + r|u(x)|^2}u(x) = k^2u(x)$$

with real number $k$, complex-valued function $q(x)$ and real-valued function $\alpha(x)$ and with parameter $r \geq 0$. In the applications usually $q$ and $\alpha$ are the characteristic functions of some bounded regions. A particular nonlinearity in (1) of cubic type ($r = 0$) can be met in the context of a Kerr-like nonlinear dielectric film, while the case when $r > 0$ corresponds to the saturation model. In scattering theory one considers solutions of (1) of the form

$$u(x, k, \theta) = u_0(x, k, \theta) + u_{sc}(x, k, \theta),$$

where $\theta \in S^1$, $u_0(x, k, \theta) = \exp(ik(x, \theta))$ is the incident wave and $u_{sc}(x, k, \theta)$ is the scattered wave. These solutions are the unique solutions of the Lippmann-Schwinger equation

$$u = u_0 - \int_{\mathbb{R}^2} G_k^+(|x - y|) \left( q(y)u + \frac{\alpha(y)|u|^2}{1 + r|u|^2}u \right) dy,$$

where $G_k^+$ is the outgoing fundamental solution of the corresponding Helmholtz equation. Now we assume in addition that the functions $q(x)$ and $\alpha(x)$ satisfy the following condition:

$$c_0(k) := \sup_{x \in \mathbb{R}^2} \int_{\mathbb{R}^2} |G_k^+(|x - y|)| (|q(y)| + |\alpha(y)|) dy \to 0$$

as $|k| \to +\infty$. It is not so difficult to verify that if $q(x)$ and $\alpha(x)$ belong to $L^p(\mathbb{R}^2)$ for some $1 < p \leq \infty$ and have compact support then condition (3) is satisfied.

Lemma L.1

If condition (3) is fulfilled and $r \geq 0$ then there exists $k_0 > 0$ such that there is a solution $u(x, k, \theta) \in L^\infty(\mathbb{R}^2)$ of (2) for which uniformly in $x \in \mathbb{R}^2$, $|k| \geq k_0$ and $\theta \in S^1$

$$u(x, k, \theta) = \lim_{j \to \infty} u_j(x, k, \theta),$$

where

$$u_{j+1} = u_0 - \int_{\mathbb{R}^2} G_k^+(|x - y|) \left( q(y)u_j + \frac{\alpha(y)|u_j|^2}{1 + r|u_j|^2}u_j \right) dy$$

for $j = 0, 1, 2, \ldots$ Moreover, $u(x, k, \theta) = u_0(x, k, \theta) + u_{sc}(x, k, \theta)$ with

$$\|u_{sc}\|_\infty \to 0, \quad |k| \to +\infty.$$
Lemma 1 allows to conclude that for fixed \( k > 0 \) as \( |x| \to \infty \),

\[
u = e^{ik(x,\theta)} - \frac{1 + i}{4\sqrt{\pi}} \frac{e^{ik|x|}}{(k|x|)^{1/2}} A(k, \theta', \theta) + o \left( \frac{1}{|x|^{1/2}} \right),
\]

where \( \theta, \theta' = \frac{x}{|x|} \in S^1 \) and the function \( A(k, \theta', \theta) \) is called a scattering amplitude and defined by

\[
A(k, \theta', \theta) := \int_{\mathbb{R}^2} e^{-ik(\theta',y)} \left( q(y)u + \frac{\alpha(y)|u|^2}{1 + r|u|^2} \right) dy.
\]

Setting \( u(x, k, \theta) = u(x, -k, \theta) \) for \( k < 0 \) we can extend \( A \) to negative \( k \leq -k_0 \) by \( A(k, \theta', \theta) = A(-k, \theta', \theta) \). And finally we put \( A(k, \theta', \theta) := 0 \) for \( |k| \leq k_0 \). The inverse problems that are considered here is to recover the discontinuities of the potentials \( q(x) \) and \( \alpha(x) \) by the knowledge of the scattering amplitude \( A(k, \theta', \theta) \) for some values of \( k, \theta, \theta' \). There are four well-known inverse problems:

1) \( A(k, \theta', \theta) \) is known for \( k \geq k_0 \) and for \( \theta \) and \( \theta' \) - general scattering.
2) \( A(k, \theta', \theta) \) is known for \( k \geq k_0 \) and for \( \theta \) and \( \theta' = -\theta \) - backscattering.
3) \( A(k, \theta', \theta) \) is known for \( k \geq k_0 \) and for \( \theta' \) and \( \theta = \theta_0 \) - fixed angle scattering.
4) \( A(k, \theta', \theta) \) is known for \( k = k_0 \) and for \( \theta \) and \( \theta' \) - fixed energy scattering.

The first result concerns to the general scattering.

**Theorem L.1**

Let us assume that \( q(x) \) and \( \alpha(x) \) belongs to \( L^p_{\text{comp}}(\mathbb{R}^2) \) and \( r \geq 0 \). Then

\[
\lim_{k \to +\infty} k \int_{S^1 \times S^1} e^{-ik(\vartheta-\vartheta',x)} A(k, \vartheta', \vartheta) d\vartheta d\vartheta' = 4\pi \int_{\mathbb{R}^2} \frac{q(y) + \alpha(y)}{1 + r} dy,
\]

where the limit is valid in the sense of distributions for \( p = 2 \) and uniformly in \( x \) for \( 2 < p \leq \infty \).

As a corollary of this formula we have uniqueness result.

**Theorem L.2**

Assume that the potentials \( q_1(x), \alpha_1(x) \) and \( q_2(x), \alpha_2(x) \) satisfy the conditions of Theorem 1, parameter \( r \geq 0 \) and the corresponding scattering amplitudes coincide for some sequence \( k_j \to +\infty \) and for all \( \vartheta, \vartheta' \in S^1 \). Then

\[
q_1(x) + \frac{\alpha_1(x)}{1 + r} = q_2(x) + \frac{\alpha_2(x)}{1 + r},
\]

in the sense of distributions. Moreover, in the sense of tempered distributions

\[
q(x) + \frac{\alpha(x)}{1 + r} = \lim_{k \to +\infty} \frac{k^2}{8\pi} \int_{S^1 \times S^1} e^{-ik(\vartheta-\vartheta',x)} A(k, \vartheta', \vartheta)|\vartheta - \vartheta'| d\vartheta d\vartheta'.
\]
2 Born Approximation

Lemma 1 allows us to conclude that for $k \to +\infty$

$$A(k, \theta', \theta) = F \left( q(x) + \frac{\alpha(x)}{1 + r} \right) (k(\theta - \theta')) + o(1)$$

uniformly with respect to $\theta$ and $\theta'$ from $S^1$, where $F$ is the two-dimensional Fourier transform. This asymptotic justifies for the backscattering and fixed angle scattering problems the following definitions:

**Definition L.1.** The inverse backscattering Born approximation $q^b_B(x)$ of the potentials $q(x)$ and $\alpha(x)$ is defined by

$$q^b_B(x) := \frac{1}{2\pi^2} \int_{R \times S^1} e^{-ik(\theta, 2x)} A(k, -\theta, \theta)|k| \, dk \, d\theta.$$  \hspace{1cm} (6)

**Definition L.2.** The inverse fixed angle Born approximation $q^\theta_0(x)$ of the potentials $q(x)$ and $\alpha(x)$ is defined by

$$q^\theta_0(x) := \frac{1}{16\pi^2} \int_{R \times S^1} e^{-ik(\theta_0 - \theta', x)} A(k, \theta', \theta_0)|k||\theta_0 - \theta'|^2 \, dk \, d\theta'.$$  \hspace{1cm} (7)

The reconstruction of singularities in inverse scattering problems are presented in the following theorem.

**Theorem L.3**

Assume that the potentials $q$ and $\alpha$ are compactly supported and belong to $L^p(R^2)$, $2 \leq p \leq \infty$, their Fourier transform belong to $L^s(R^2)$ with $s \leq \frac{4p}{3p-1}$ and parameter $r \geq 0$. Then

$$q^b_B(x) - q(x) - \frac{\alpha(x)}{1 + r} \in H^t(R^2)(\text{mod } C^\infty(R^2))$$

$$q^\theta_0(x) - q(x) - \frac{\alpha(x)}{1 + r} \in H^t(R^2)(\text{mod } C^\infty(R^2)),$$

where $t < 1 - \frac{1}{p}$.

**Corollary L.1**

If a piecewise smooth compactly supported function $q + \frac{\alpha}{1 + r}$ contains a jump over smooth curve, then the curve and the height function of the jump are uniquely determined by the backscattering or by the fixed angle scattering data.

3 Fixed Energy Problem

For inverse problem at fixed energy the crucial role is played by the Green-Faddeev function

$$g_z(x) := \frac{1}{4\pi^2} \int_{R^2} \frac{e^{i(x, \xi)}}{\xi^2 + 2(z, \xi)} \, d\xi,$$
where \( z = (z_1, z_2) \in C^2 \) is a two-dimensional complex vector with \( z_1^2 + z_2^2 = 0 \), and its estimate for \(|z| > 1\) and with \( \gamma < 1 \)

\[
\|g_z \ast f\|_{L^\infty(R^2)} \leq \frac{C}{|z|^{\gamma}} \|f\|_{L^2(\Omega)}.
\] (8)

The estimate (8) allows us to prove that for parameter \( r > 0 \) and for \( q, \alpha \in L^p(\Omega), 2 \leq p \leq \infty \), there exists a special solution to the Schrödinger equation

\[
-\Delta u(x) + q(x)u(x) + \frac{\alpha(x)|u(x)|^2}{1 + r|u(x)|^2} u(x) = 0
\]
of the form (for \(|z|\) large enough)

\[
u(x, z) = e^{i(x, z)}(1 + R(x, z))
\]
with \( R \in L^\infty(R^2) \). Further more the following estimate holds

\[
\|R\|_{L^\infty(R^2)} \leq C|z|^{\gamma}, \quad \gamma < 1.
\] (9)

We can define now the scattering transform of the potentials \( q \) and \( \alpha \) as follows:

\[
T_{q,\alpha}(\xi) := \int_{R^2} e^{i(x, \xi)} \left( q(x) + \alpha(x) \frac{1 + R^2}{1 + r|1 + R^2|} (1 + R(x, z)) \right) dx,
\] (10)

where \( |\xi| \) is large, \( z = \frac{1}{2}(\xi - iJ\xi) \) and \( J = \{a_{ij}\} \) is the matrix with \( a_{11} = a_{22} = 0, a_{12} = -a_{21} = 1 \). For other values of \( \xi \) we put \( T_{q,\alpha}(\xi) \) be equal to 0.

**Definition L.3.** The inverse fixed energy Born approximation \( q^f_B(x) \) of the potentials \( q(x) \) and \( \alpha(x) \) is defined by

\[
q^f_B(x) := \frac{1}{4\pi^2} \int_{R^2} e^{-i(x, \xi)} T_{q,\alpha}(\xi) d\xi.
\] (11)

**Theorem L.4**
Assume that the potentials \( q \) and \( \alpha \) are compactly supported and belong to \( L^p(R^2), 2 \leq p \leq \infty \), and parameter \( r > 0 \). Then

\[
q^f_B(x) - q(x) - \frac{\alpha(x)}{1 + r} \in H^t(R^2)(mod \quad C^\infty(R^2)),
\]
where \( t < 1 \).

**Corollary L.2**
If a piecewise smooth compactly supported function \( q + \frac{\alpha}{1+r} \) contains a jump over smooth curve, then the curve and the height function of the jump are uniquely determined by the fixed energy scattering data.
Paper M

Robust Minimax Bias Estimation of the Correlation Coefficient

Author: G. L. Shevlyakov
Robust Minimax Bias Estimation of the Correlation Coefficient

Georgy L. Shevlyakov

St. Petersburg State Polytechnic University,
St. Petersburg, Russia.
E-mail: georgy.shevlyakov@gmail.com
1 Introduction

The simplest problem of correlation analysis is to estimate the correlation coefficient $\rho$ of a bivariate distribution density $f(x, y)$ with the observed sample $(x_1, y_1), \ldots, (x_n, y_n)$. Its classical estimator is given by the sample correlation coefficient

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\left(\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2\right)^{1/2}},$$

where $\bar{x} = n^{-1} \sum x_i$, and $\bar{y} = n^{-1} \sum y_i$ are the sample means.

It is well known that the sample correlation coefficient is extremely sensitive to the presence of outliers in the data: in particular, the sample correlation coefficient $r$ is strongly biased with respect to $\rho$ in the gross error model [5].

By Monte Carlo study of a variety of robust estimators of correlation in a variety of situations, in [3] it was observed that the robust estimator named as the median correlation coefficient $r_{med} = (\text{med}|u| - \text{med}|v|) / (\text{med}|u| + \text{med}|v|)$, (1)

performed better with respect to bias than the others. In formula (1), $\text{med}|z|$ is the sample median of the absolute values of $z$; $u$ and $v$ are the robust principal variables

$$u = \frac{x - \text{med}x}{\text{MAD} x} + \frac{y - \text{med} y}{\text{MAD} y}, \quad v = \frac{x - \text{med}x}{\text{MAD} x} - \frac{y - \text{med} y}{\text{MAD} y},$$

and $\text{MAD} z = \text{med}|z - \text{med} z|$ is the median absolute deviation.

Further, in [2], it was shown that the median correlation coefficient $r_{med}$ (1) is the limit case (as $\varepsilon \to 1$) of the asymptotically minimax variance estimator of correlation, namely, the trimmed correlation coefficient, for $\varepsilon$-contaminated bivariate normal distributions.

In this work, we state that the median correlation coefficient $r_{med}$ is also the minimax bias estimator of the correlation coefficient.

2 Main Result

Consider the class of bivariate independent component distribution densities allowing for the factorization in the principal coordinates $u$ and $v$ [2]

$$f(x, y) = \frac{1}{\sigma \sqrt{1 + \rho}} g\left(\frac{x}{\sigma \sqrt{1 + \rho}}\right) \frac{1}{\sigma \sqrt{1 - \rho}} g\left(\frac{y}{\sigma \sqrt{1 - \rho}}\right),$$

where $g(x)$ is a symmetric density, $u = (x + y)/\sqrt{2}$, $v = (x - y)/\sqrt{2}$ are the principal coordinates, $\sigma^2 = \text{var}(X) = \text{var}(Y)$, and $\rho$ is the correlation coefficient. In particular, the class (2) contains the standard normal density $f(x, y) = \mathcal{N}(x, y; 0, 0, 1, 1, \rho)$ when $g(x) = (2\pi)^{-1/2} \exp(-x^2/2)$.

Consider the following estimation procedure for the correlation coefficient $\rho$:

1) transform the initial data $u_i = (x_i + y_i)/\sqrt{2}, v_i = (x_i - y_i)/\sqrt{2}, i = 1, \ldots, n$;

2) compute the $M$-estimates of scale $\hat{\beta}_u$ and $\hat{\beta}_v$ of the principal variables as the solutions to the equations

$$\sum \chi(u_i/\hat{\beta}_u) = 0, \quad \sum \chi(v_i/\hat{\beta}_v) = 0,$$
where $\chi(\cdot)$ is a score function;
3) define the estimator of $\rho$ as
\[ \hat{\rho}_n = \frac{(\hat{\beta}_u^2 - \hat{\beta}_v^2)}{(\hat{\beta}_u^2 + \hat{\beta}_v^2)}. \] (4)

**Theorem M.1**
The asymptotic bias of estimator (4) has the form
\[ E(\hat{\rho}_n) - \rho = b(\chi, g)/n + o(1/n) \quad \text{with} \quad b(\chi, g) = -\rho(1 - \rho^2) V(\chi, g)/2, \] (5)

where
\[ V(\chi, g) = \int x^2(x)g(x) \, dx / \left( \int x\chi'(x)g(x) \, dx \right)^2 \]
is the asymptotic variance of $M$-estimators of scale (3).

**Proof:** Proof follows the ideas of the proof of Theorem 1 [2]. \qed

Consider the class $\mathcal{F}_\varepsilon$ of $\varepsilon$-contaminated bivariate independent component densities (2)
\[ \mathcal{F}_\varepsilon = \{ f : f(x, y) \geq (1 - \varepsilon)\mathcal{N}(x, y; 0, 0, 1, 1, \rho), \quad 0 \leq \varepsilon < 1 \}. \] (6)

with the correlation coefficient $\rho$.

Then the problem of minimax bias estimation of $\rho$ can be written as follows
\[ \min_{\chi} \max_{f \in \mathcal{F}_\varepsilon} n b_n(\chi, f) = \min_{\chi} \max_{g \in \mathcal{G}_{\gamma}} n b_n(\chi, g) = n b_n(\chi^*, g^*), \] (7)

where $(\chi^*, g^*)$ is an optimal saddle-point pair $(\chi^*, g^*)$ solving problem (7); $\mathcal{G}_{\gamma}$ is the class of $\gamma$-contaminated univariate standard normal distributions
\[ \mathcal{G}_{\gamma} = \{ g : g(x) \geq (1 - \gamma)\varphi(x), \quad 0 \leq \gamma < 1 \} \] (8)

with $\varepsilon = 2\gamma - \gamma^2$ (for details, see [2]). In other words, due to the structure of formula (5), the problem of minimax bias estimation of $\rho$ is equivalent to the problem of minimax variance estimation of a scale parameter. Since the latter problem is solved by Huber [4], we directly arrive at the following result.

**Theorem M.2**
Under the regularity conditions imposed on the score functions $\chi$ and the densities $g$ (for details, see [1]; pp. 125, 139), in the class $\mathcal{F}_\varepsilon$ of $\varepsilon$-contaminated bivariate independent component distribution densities, the minimax bias estimator of the correlation coefficient is given by estimator (4) with the optimal score $\chi_{MAD}(x) = \text{sgn}(|x| - 1)$ in equations (3), i.e., by the median correlation coefficient $r_{\text{med}}$.

**Proof:** Proof literally repeats the proof of Theorem 2 in [2]. \qed

Thus, the median correlation coefficient is simultaneously the asymptotically minimax bias and minimax variance estimator of the correlation coefficient.
3 Conclusion

The median correlation coefficient $r_{med}$ is a highly robust estimator of the correlation coefficient being a correlation analog of such classical robust estimators of location and scale, as the sample median and the median absolute deviation.

References


Paper N

Partial Balayage and the Inverse Problem of Potential Theory

Authors: T. Sjödin
Partial Balayage and the Inverse Problem of Potential Theory

Tomas Sjödin

Department of Mathematics,
Linköping University,
Linköping, Sweden.
E-mail: tosjo@mai.liu.se
1 Introduction

A bounded domain $\Omega$ in $\mathbb{R}^N$ is called solid if $\mathbb{R}^N \setminus \Omega$ is connected and $\Omega = \overline{\Omega}^\circ$. The exterior inverse problem of potential theory, which dates back to work of Novikov in the 1930’s, is as follows:

If $\Omega_1$ and $\Omega_2$ are solid domains in $\mathbb{R}^N$ such that $\lambda|_{\Omega_1}$ and $\lambda|_{\Omega_2}$ produce the same potential in the complement of $\Omega_1 \cup \Omega_2$, must $\Omega_1$ and $\Omega_2$ coincide?

Novikov himself proved that the answer is yes if both domains are assumed to be convex or, more generally, starshaped with respect to a common point. Although it is nowadays suspected that the answer to the general question may be negative, it has long been conjectured that convexity of one of the domains should be enough for a positive answer. That this is, in fact, the case was recently proved in [3]. The proof, which relies on partial balayage and the ‘moving plane’ method, will be outlined in the last section. We start with a section on partial balayage.

This summary is heavily based on the article [4].

2 Partial Balayage

We will be working in $\mathbb{R}^N$ ($N \geq 2$), and let $\lambda$ denote Lebesgue measure on $\mathbb{R}^N$. We also let $U_\mu$ denote the Newtonian (or logarithmic, if $N = 2$) potential of $\mu$, normalized so that $-\Delta U_\mu = \mu$ in the sense of distributions.

Here we recall some basic facts about the notion of partial balayage, which was originally developed by Gustafsson and Sakai [5]. A recent exposition of it may be found in [4]. For a positive measure $\mu$ with compact support in $\mathbb{R}^N$ we define

$$ V_\mu(x) = \sup \left\{ v(x) : v \text{ is subharmonic and } v \leq U_\mu + \frac{|\cdot|^2}{2N} \text{ on } \mathbb{R}^N \right\} - \frac{|x|^2}{2N} $$

and then put $B_\mu^\lambda = -\Delta V_\mu$. Then it turns out that

$$ B_\mu^\lambda = \lambda|_{\omega(\mu)} + \mu|_{\omega(\mu)^c} = \lambda|_{\Omega(\mu)} + \mu|_{\Omega(\mu)^c}, $$

where

$$ \omega(\mu) = \{ V_\mu < U_\mu \} $$

and

$$ \Omega(\mu) = \bigcup \{ U : U \subset \mathbb{R}^N \text{ open and } B_\mu^\lambda = \lambda \text{ in } U \}, $$

and these are bounded open subsets of $\mathbb{R}^N$. Further,

$$ B_\mu^\lambda \leq \lambda. $$

We note that $\omega(\mu) \subset \Omega(\mu)$ and that this inclusion may be strict, even when $\mu$ has compact support contained in $\Omega(\mu)$. It will be convenient to define

$$ W_\mu = U_\mu - V_\mu, $$

whence $W_\mu$ is lower semicontinuous, $-\Delta W_\mu \geq \mu - \lambda$ and $W_\mu \geq 0$ on $\mathbb{R}^N$. 


3 Application to the Exterior Inverse Problem

Theorem N.1
If $\Omega_1, \Omega_2$ are bounded solid domains in $\mathbb{R}^N$, where $\Omega_2$ is convex, and if
\[ U^{\chi_{\Omega_1}} = U^{\chi_{\Omega_2}} \text{ in } \mathbb{R}^N \setminus (\Omega_1 \cup \Omega_2), \]
then $\Omega_1 = \Omega_2$.

In order to describe the moving plane method, we need some notation. Points in $\mathbb{R}^N$ will be denoted by $(x', x_N) \in \mathbb{R}^{N-1} \times \mathbb{R}$, and we will write
\[ W_+ = \{(x', x_N) : x_N > 0\}, \quad W_- = \{(x', x_N) : x_N < 0\}, \quad H = \{(x', x_N) : x_N = 0\}. \]
Given a positive measure $\mu$ we denote by $\Omega(\mu)$ the largest open set $U$ for which $\left(\lambda - B_\mu^\lambda\right)(U) = 0$. (Thus $\Omega(\mu)$ will contain the set $\omega$.) We also know that
\[ B_\mu^\lambda = \lambda|_{\Omega(\mu)} + \mu|_{\Omega(\mu)^c}, \tag{3} \]
where $\mu|_{\Omega(\mu)^c} \leq \lambda$.

Lemma N.1
Let $\mu$ be a measure with compact support contained in $W_- \cup H$ and let $A = \{x' : (x', 0) \in \Omega(\mu) \cap H\}$. Then there is a continuous function $g : A \to (0, \infty)$, continuously vanishing on $\partial A$, such that
\[ \Omega(\mu) \cap W_+ = \{(x', x_N) : x' \in A \text{ and } 0 < x_N < g(x')\}. \tag{4} \]

The name ‘moving plane’ comes from the fact that we can apply the above lemma to all hyperplanes disjoint from the convex hull $K$ of the support of $\mu$ to see that $\Omega(\mu)$ cannot have any ‘holes’ outside $K$ (see the following figure).
Proof: Let $u = U\mu - UB^\lambda$. Thus $u \geq 0$. We may assume, by means of a limiting argument, that $\text{supp}\mu \subset W_+$, and so $u$ is continuously differentiable on an open set containing $W_+ \cup H$. Let $\overline{\pi}(x) = u(x', -x_N)$. We note that $U\mu - \overline{\pi} - \nu$ is subharmonic on $W_+$, and $UB^\lambda - \nu$ is subharmonic on all of $\mathbb{R}^N$. Since

$$U\mu - \overline{\pi} - \nu = U\mu - u - \nu = UB^\lambda - \nu \quad \text{on } H,$$

the function

$$w = \begin{cases} \max\{U\mu - \overline{\pi} - \nu, UB^\lambda - \nu\} & \text{on } W_+ \\ UB^\lambda - \nu & \text{on } W_- \cup H \end{cases}$$

is a subharmonic minorant of $U\mu - \nu$. Thus $w \leq UB^\lambda - \nu$ by the definition of $UB^\lambda$, whence $U\mu - \overline{\pi} \leq UB^\lambda$ on $W_+$ and so $u \leq \overline{\pi}$ there. It follows that $\partial u/\partial x_N \leq 0$ on $H$.

Let $\Omega_+ = \Omega(\mu) \cap W_+$. Since $u = 0$ on $\omega^c$ (where $\omega = \omega(\mu)$), and so on $\Omega(\mu)^c$, and since every point of $\partial \Omega(\mu)$ is the limit of some sequence of points of Lebesgue density of $\Omega(\mu)^c$, we see that $|\nabla u| = 0$ on $\partial \Omega_+ \cap W_+$. We note from (3) that $\Delta u$ is constant in $\Omega_+$, so the function $\partial u/\partial x_N$ is harmonic there, and hence $\partial u/\partial x_N \leq 0$ on $\Omega_+$, by the maximum principle. Further, since $u$ is nonconstant in each component of $\Omega_+$, and $u = 0$ on $W_+ \setminus \Omega_+$, we actually have $\partial u/\partial x_N < 0$ on $\Omega_+$. We now define

$$g(x') = \sup\{t > 0 : (x', t) \in \Omega_+\} \quad (x' \in A).$$

Clearly $\Omega(\mu) \cap W_+$ lies under the graph of $g$. Conversely, if $(x', x_N)$ lies under the graph of $g$ and $x_N > 0$, then $u(x', x_N) > 0$ and so $(x', x_N) \in \omega(\mu) \subset \Omega(\mu)$. Thus (4) holds.

It remains to check that $g$ is continuous and vanishes at $\partial A$. In fact, since $\Omega(\mu)$ is open and

$$\{x' : g(x') > c\} = \{x' : (x', c) \in \Omega(\mu)\} \quad (c > 0),$$

it is clear that $g$ is lower semicontinuous. On the other hand, if we apply the result of the previous paragraph with hyperplanes of varying orientation, we see that each point of $\partial \Omega_+ \cap W_+$ is the vertex of a vertical cone lying in $\Omega(\mu)^c$, and so $g$ is also upper semicontinuous. In fact $g$ vanishes continuously at $\partial A$, since we can apply the preceding reasoning with $H$ replaced by a slightly lower hyperplane. 

We now proceed to the proof of Theorem N.1. Let $\eta$ be the Riesz measure associated with $\min\{U^{\lambda_1}, U^{\lambda_2}\}$, and $\Omega_0 = \Omega(\eta)$. The first step is to prove a simple lemma about $\Omega_0$.

**Lemma N.2**

*With the above notation, $B^\lambda_\eta = \lambda|_{\Omega_0}$ and $\Omega_1 \cap \Omega_2 \subset \Omega_0$.***

**Proof:** The set of points

$$\{x : U^{\lambda_1}(x) = U^{\lambda_2}(x)\} \cap ((\Omega_1 \setminus \Omega_2) \cup (\Omega_2 \setminus \Omega_1))$$

has Lebesgue measure zero. Also, we can apply Kato’s inequality, also known as Grisin’s lemma (see [1, 2]), to $U^{\lambda_i} - U\eta$ ($i = 1, 2$) to see that $\eta \geq \lambda|_{\Omega_1 \cap \Omega_2}$. Hence $\eta$ is of the form $\eta|_D + \eta_1$, where $D$ is an open set, $\eta \geq \lambda|_D$ and $\eta_1$ is singular with respect to Lebesgue measure. Since $B^\lambda_\eta = \lambda|_{\Omega_0} + \eta|_{\Omega_0}$ we see that $D \subset \Omega_0$. It follows that $\eta|_{\Omega_0^c} = \eta_1|_{\Omega_0^c} = 0$, since $B^\lambda_\eta \leq \lambda$. 

\[\square\]
It is easy to see that $\lambda(\Omega_1) = \lambda(\Omega_2) = \lambda(\Omega_0)$ and, by construction, we have $U^{\chi_{\Omega_0}} \leq U^{\chi_{\Omega_1}} (i = 1, 2)$. Further, $\partial \Omega_i \subset \overline{\Omega}_0$ ($i = 1, 2$), because if there were an open ball $B$ in $\mathbb{R}^N \setminus \overline{\Omega}_0$ with center in $\partial \Omega_i$, then we would arrive at the impossible situation that $U^{\chi_{\Omega_1}} - U^{\chi_{\Omega_0}}$ is non-negative and superharmonic on $B$ and vanishes on $B \setminus \Omega_i$ but not on all of $B$.

Let $\mu = \lambda|_{\Omega_2} + \lambda|_E$, where $E = \Omega_2 \setminus \Omega_0$. Also let $D = \Omega_0 \cup E = \Omega_0 \cup \Omega_2$. The next step in the proof is to show that $B^\lambda = \lambda|_D$ and hence $D \subset \Omega(\mu)$. From this we will be able to obtain a contradiction if $\Omega_1 \neq \Omega_2$ by an application of the moving plane method. What follows is a sketch of the remaining proof; full details may be found in [3].

On the set $A_2 = \mathbb{R}^N \setminus (\Omega_0 \cup \Omega_1 \cup \Omega_2) = \mathbb{R}^N \setminus (D \cup \Omega_1)$ we have by assumption $U_\eta = U^{\chi_{\Omega_0}} = U^{\chi_{\Omega_1}} = U^{\chi_{\Omega_2}}$, so $U_\mu = U^{\chi_D}$ here. But on the set $A_1 = \Omega_1 \setminus \Omega_0 = \Omega_1 \setminus \overline{D}$ it is also clear that $U_\eta = U^{\chi_{\Omega_2}}$, since $-\Delta U^{\chi_{\Omega_1}} = \lambda$ on $A_1$. Hence $U_\mu = U^{\chi_D}$ on $\mathbb{R}^N \setminus (\partial \Omega_1 \cup \overline{D})$. It follows by continuity that $U_\mu = U^{\chi_D}$ on $\mathbb{R}^N \setminus \overline{D}$, and we also have $U_\mu \geq U^{\chi_D}$ everywhere. Hence $\Omega(\mu) \subset \overline{D}$. But $\lambda(\Omega(\mu)) = \lambda(D)$, so $D \subset \Omega(\mu)$, and these sets differ by at most a Lebesgue null set. Thus $B^\lambda = \lambda|_D$. Since $\lambda(\partial \Omega_2) = 0$ and the set $\Omega(\mu) \setminus (E \cup \partial \Omega_2)$ is open, the latter is a subset of $\Omega_0$ by construction (it differs from $D \setminus E = \Omega_0$ by at most a Lebesgue null set). Therefore $D \subset \Omega(\mu) \subset D \cup \partial \Omega_2$.

According to the moving plane method it is not possible for $\Omega(\mu)$ to have any holes outside the set $\overline{\Omega}_2$ (like $A_1$ or $A_2$ in the figure). Hence $\Omega_1 \subset \Omega(\mu)$, since $\partial \Omega_1 \subset \overline{\Omega}_0$, so $\Omega_1 \subset \Omega_0$, and using the fact that $\Omega_1$ and $\Omega_0$ have equal Lebesgue measure, and are both saturated with respect to Lebesgue measure, it follows that they are in fact identical.

But this implies that $U^{\chi_{\Omega_1}} \leq U^{\chi_{\Omega_2}}$ everywhere. Since $U^{\chi_{\Omega_2}} - U^{\chi_{\Omega_1}}$ is superharmonic on $\mathbb{R}^N \setminus \overline{\Omega}_1$ (which is connected) and attains the value 0 it follows that $U^{\chi_{\Omega_1}} = U^{\chi_{\Omega_2}}$ on this set. So $\Omega_2 \subset \Omega_1$, and since they are both solid with equal Lebesgue measure it follows that they are also identical.
References


Paper O

Methane Armageddon: Is it Possible?

Authors: I. Sudakov and S. Vakulenko
Methane Armageddon: Is it Possible?

Ivan Sudakov* and Sergey Vakulenko†

*Nansen International Environmental and Remote Sensing Centre, St. Petersburg, Russia.

†Institute for Mechanical Engineering Problems, Russian Academy of Sciences, St. Peterburg, Russia.
E-mail: vakulenfr@mail.ru
1 Introduction

In Siberia there is an ecological bomb, namely, about $80 \cdot 10^{15} kg$ of methane in a frozen state. As a result of warming, a number of small lakes form and extend, and through them methane enters for the atmosphere. In turn, this methane reinforces warming, thus, one obtains a positive feedback to an ecological catastrophe. The aim of this note is to give a mathematical method to estimate such a possibility and define key parameters that influence the process.

This melting process can be described by the Stefan approach [[1, 3]]. It is difficult to resolve the multidimensional Stefan problem numerically, moreover, a multi-dimensional Stefan layer is unstable [[1]]. To overcome these difficulties, we propose to use a phase transition theory based on the fundamental concept of the order parameter. When the parameter order $\phi$ is close to $-1$, the permafrost state is ice, if $\phi$ close to 1 this state is water. It is shown that this approach generalizes and improves the classical Stefan method [[1]]. An important advantage of the field approach is that the numerical realization becomes simpler and the layer width is not zero. Moreover, we use a natural assumption that the horizontal dimensions of the permafrost layer are much larger than the layer depth.

The main idea is as follows. Typically, the lakes have horizontal sizes $100 - 1000 m$, and a small depth. It is clear that temperature at the lake bottom is, nonetheless, essentially below than at lake boundaries where depths are smaller. One can expect thus that the lakes grow in horizontal directions. Therefore, we can apply 2-dimensional models of phase transitions to describe this lake slow extension. Notice that he phase transition is a microscopic process, the lakes have macroscopic objects. Thus we can assume that, locally, the lake boundary is a circle of a large curvature radius $R$. Then the melting front velocity can be investigated by asymptotics based on so-called the mean curvature motion [[6, 3]].

As a result, we obtain that the lake increasing rate is proportional to a constant, that can be empirically estimated, plus a small contribution proportional to $R^{-1}$. This gives a formula describing a dependence of methane production in Siberia on temperature. By this formula we show, by numerical simulations and analytical arguments, that the Armageddon effect depends crucially on a coefficient describing a positive feedback between temperature and methane concentration.

2 Phase Transition Approach

We denote the order parameter by $\phi$ and use the following equations for two unknown functions $\phi, u$, where $u$ is temperature [[1]]:

$$u_t = K \Delta u - b \phi_t,$$

(1)

$$\alpha \xi^2 \phi_t = \xi^2 \Delta \phi + a^{-1} g(\phi) + 2(u - \theta),$$

(2)

where $g = 0.5(\phi - \phi^3)$, $\theta$ is the temperature of the phase transition (temperature of melting), $a, \alpha, \xi, b, K$ are parameters. To explain the main idea beyond (1), (2), let us consider the 1D version of (2) removing the term with the temperature:

$$\alpha \xi^2 \phi_t = \xi^2 \phi_{zz} + a^{-1} g(\phi),$$

(3)
where we assume that the front moves along the $z$ axis. Let $u$ set $\alpha = 1$. Then (3) has a solution $\tanh((x - x_0)/\epsilon)$, where $\epsilon = a^{-1/2}\xi^{-1}$ is a parameter that defines the front width. To describe the influence of $2(u - \theta)$, we can apply the perturbation theory. The front appears at a point, where $u \approx \theta$, i.e., at the melting temperature.

In addition to (1), (2), we set standard boundary conditions on the depth and on a surface. The melting front position is defined by $\phi(x, y, l, t) = \theta$. Notice that the phase field approach is more effective numerically than the the Stefan problem method, and the corresponding initial boundary value problem is well posed.

### 3 Lake Area Growth and Methane Emission

Using above equations, one can obtain the relation

$$v(x, y, z, t) = \delta - \mu \kappa(x, y, z, t),$$

(4)

where $v$ is the normal melting front velocity at the point $(x, y, z)$, $\kappa$ is the mean front curvature at this point, $\mu$ is a positive coefficient.

Relation (4) can be checked experimentally. Let us consider the plane case. Then our fronts are curves. All front closed curves, which initially are not too distinguished of ellipses or circles, approaches for large times to circles. For circular fronts of the radius $R(t)$ eq. (4) takes the form

$$\frac{dR}{dt} = \delta - \mu R^{-1},$$

(5)

and can be resolved explicitly. The quantity $\delta$ can be expressed via microscopic parameters $a, \xi, b$. One can estimate $\delta$ by experimental data since $\delta$ determines the main contribution into lake area increasing.

To describe the methane emission rate, let us assume following [[4]] that locally the probability to emit a gaseous methane molecule is proportional to $p = \exp(-U_0/k_B u)$, where $u$ is the temperature, $k_B$ is the Boltzmann constant, $U_0$ is a parameter to adjust (a potential barrier for solid-gas transition). Then the total rate of the methane emission per unit time is given by

$$V_{meth} = \rho_0 \int \int \int_{D_M} \exp(-\frac{U_0}{k_B u}) dxdydz,$$

(6)

where $u = u(x, y, z, t)$ is the absolute temperature, $\rho_0$ is the methane density in the domain $D_M$, this domain is the methane emission zone depending on time $t$ formally defined by $D_M(t) = \{(x, y, z) : u(x, y, z, t) > \theta\}$.

Finally, the phase field theory gives a semi-heuristic model describing the methane emission by a lake system. We assume that the lakes are shallow and the lake growth reduces to the radiuses $R_i$ increasing. Indeed, the most of tundra lakes in Siberia are shallow (depth $1 - 10m$, typically $2 - 3m$) but large ($100m - 1000m$ in diameter). The melting and erosion lead to the diameter extension, the depth increases essentially slower. Therefore, we can apply a plane (2D ) model. Assume, in average, the lake form is close to circular. Indeed, these lakes developed as a result of longtime evolution, thus we can assume that original, possibly complicated, forms become circular. Then we can use eq. (4), where $\mu, \delta$ are positive coefficients that can be calculated from microscopic
parameters $a, \xi, b$ of the phase transition theory, but it is more reasonable to consider (7) as a phenomenological model, where they can be found by experimental data.

Assume there are a number of $R_i, i = 1, 2, ..., N$ lakes. The averaged lake size is $R = N^{-1} \sum_{i=1}^{N} R_i$, the total area is $S = \sum_{i=1}^{N} R_i^2$. Then the total methane emission rate per unit time is given by

$$V_{meth} = C dS \frac{dt}{dt} = 2CN(-\mu + \delta R),$$

where $\mu, C$ are constants, $R$ is the averaged lake radius. According to (6) the constant $C$ depends on the temperature $u$ as $C = c_0 \exp(-U_0(k_B u)^{-1})$, here $c_0$ is a positive coefficient.

Finally, we conclude with the following simplest semi-phenomenological relation

$$V_{meth} = \beta (dR - 1) \exp(-\bar{b}/\bar{u}(t)),$$

where $\beta$ is a constant that can be found from experimental data, $\bar{b} = U_0/k_B$ and $d$ are positive constants, the $\bar{u}$ is temperature averaged over the region.

This equation is in an accordance with phenomenological models. We can mention here two following models. The first is a result of observations under a large mire "Bachkarevskoe" within 1994-2002 [5]. The empirical formula is proposed (an analogous formula is obtained in [2]): $F_{methloc} = \exp(0.492 + 0.126(T_{12} - 273) - 0.057W)$, where $F$ is the flux measured in $Mgm^{-2}$ per hour, $T_{12}$ the temperature at $12cm$ depth, $W$ is the water level. To connect (8) with this relation, we can use the Taylor extension for $F_{methloc}$ at an averaged temperature $u = u_0$ assuming that the temperature deviations are small with respect to $u_0$.

Numerical results are as follows. We simulate a development of a lake system consisting of $N >> 1$ lakes. The initial $R_i(0)$ are distributed according the normal law $N(\bar{R}(0), \sigma)$. The $R_i(t)$ were computed then by (4), the parameters are taken from experiments. The emission rate is computed by (8), and we take into account a small positive feedback assuming that $u = u_0 + \gamma X$, where $X(t)$ is total methane mass in atmosphere, $\gamma > 0$.

The simulations show that there are two types of $X(t)$ dependencies. For small feedback methane-temperature coefficients $\gamma$ we observe a linear or weakly parabolic growth methane concentration. For larger $\gamma$ one can see a true 'Armageddon'. In this case firstly we see a linear growth, and after a very sharp increasing of $X$.

References


Paper P

Bifurcations in Large Free-Scale Networks

Authors: S. Vakulenko and M. Zimin.
Bifurcations in Large Free-Scale Networks

Sergey Vakulenko* and M. Zimin

Institute for Mechanical Engineering Problems, Russian Academy of Sciences, St. Peterburg, Russia.
*E-mail: vakulenfr@mail.ru
1 Introduction

In the last decades, much attention has been payed to large scale-free networks. Scale-free structures occur in metabolic and genetic networks, Internet, economical and communication systems [[1, 8, 5]]. Dynamical systems, defined by the networks, are important for many applications, in particular, for ecology [[2]]. We consider the following dynamical model, popular in neural and genetic network theory [[6, 4]]:

\[ x_i(t + 1) = \sum_{j=1}^{N} W_{ij} \sigma(x_j(t) - h_j) \]  

(1)

where \( x_i \) are unknown functions (gene or species concentrations, neuron activity), \( W_{ij} \) are entries of an interaction matrix, \( h_i \) are thresholds, \( N \) is the node number, \( \sigma \) is a fixed \( C^\infty \) smooth monotone increasing function such that \( \sigma(+\infty) = 1, \sigma(-\infty) = 0 \). We assume that the connections \( W_{ij} \) are realized on a graph having a scale-free structure. To precise this assumption, let us consider a directed graph \((V, E)\), where \( V = \{1, 2, ..., N\} \) is the set of nodes, \( E \) denotes the set of edges. An edge \((i, j)\) is belong to \( E \) if and only if \( W_{ij} \neq 0 \). We denote by \( C_{in}(n), C_{out}(n) \) the input and output degrees of the \( n \)-th node. Let us denote by \( P_{in}(k) \) the probability that a randomly chosen node has \( k \) input edges \((C_{in} = k)\). For scale-free graphs \((V, E)\) this probability has asymptotics \( P_{in}(k) = O(k^{-\gamma}), \gamma \in (2, 3), \) and similarly for \( P_{out}(k) \). We assume that \( W_{ij} \neq 0 \) only for pairs \((i, j)\) such that the corresponding edge in \( E: (i, j) \in E \). This means that weights \( W_{ij} \) satisfy some topological restrictions.

We show, by analytical methods, that all finite dimensional structurally stable dynamics with discrete time can be realized by networks (1), by an appropriate choice of parameters \( N, W, h \). This realization holds under scale-free restrictions on \( W \) (the corresponding graphs \((V, E)\) has a scale-free structure). A more precise mathematical formulation is as follows.

For any \( n \) and within arbitrary precision \( \epsilon > 0 \) any smooth \( n \)-dimensional map \( G: q \rightarrow G(q), q = (q_1, ..., q_n) \in D \), defined on a compact domain \( D \subset \mathbb{R}^n \), can be realized by a network (1) with appropriate parameters \( N, W, h \). In this network there exist nodes \( x_1 = q_1, ..., x_n = q_n \) of high connectivity (hubs) such that the time evolution of \( q \) is defined by \( q(t + 1) = G(q) + \epsilon \tilde{G}(q) \), where \( |\tilde{G}|, |\nabla \tilde{G}| \) are bounded by 1 in \( D \).

Notice that if the prescribed dynamics \( G \) is persistent under small perturbations (structurally stable), then attractors \( A_G \) of \( G \) and \( A_{G+\epsilon \tilde{G}} \) of the perturbed map are topologically equivalent, and the corresponding reduced dynamics on these attractors are also topologically equivalent. Assume \( A_G \) is a strange hyperbolic attractor. Then our network also produces a complicated chaotic dynamics (this follows from dynamical system theory [[9]])

The approximation procedure is constructive, and gives a method of the attractor control. This procedure admits a transparent interpretation: to obtain a given \( n \)-dimensional dynamics, \( DimD = n \), we take a sufficiently large network with \( n \) hubs and we assign, in an appropriate way, weights \( W_{ij} \) for interactions between hubs and some adjacent nodes. This creates a feedback loop. To explain the main idea, let us consider the case \( n = 1 \), i.e, an one-dimensional map \( G, q \in \mathbb{R} \), and \( D = [0, 1] \).

First, one approximates a given map \( G \) and derivatives \( \nabla G \) by a multilayered neural network on \( D \) within \( \epsilon \)-precision. As a result, one obtains some coefficients \( b_j, j = \)
1, ..., \(N_d\), the higher precision \(\epsilon\) is, the larger is the coefficient number \(N_d\). Second, we choose a highly connected node (a hub) in a large network, this node serves as a center. Then we choose \(N_d\) nodes (satellites) connected with this hub. Using obtained \(b_j\), we assign the weights for edges that connect the center and satellites. Then one can show that satellite state evolution is completely captured by the center \(q\)-dynamics.

Such structures can be interpreted as toy models of centralized systems, where we observe a small number of highly connected centers interacting with a large set of weakly connected nodes. They can appear in ecology, when a predator may specialize in many different species of prey (examples: human, leopard and dog).

An important question is robustness of such networks with respect to "mutations". We have established the following fact: the scale-free networks can realize complicated bifurcations of dynamics as a result of an elimination of a node of the minimal connectivity, \(C_{in}, C_{out} = 1\). These bifurcations are capable to trigger an attractor to another. The fact, that elimination of a hub can trigger dynamics, is obvious and well known [[2]].

Notice that for large \(N\) these bifurcations can be obtained only as a result of a target attack: if we eliminate a randomly chosen node, possibly that the network dynamics does not change essentially. Moreover, it is possible that even an elimination of many nodes does not influence essentially the network attractor. Only elimination of some special weakly connected nodes can lead to a sharp change of the network dynamics. Such nodes can be named "keystones", concept of keystone species is well known in ecology [[7]].

\section{Main Result}

To formulate main assertions in more precise mathematical terms, we need the notion of so-called \(\epsilon\)-realization of dynamics ([3]), applications to PDE and networks see [10]). Let us consider iterations defined by

\[ q_i(t + 1) = F_i(q), \quad q = (q_1, q_2, ..., q_n) \in D, \tag{2} \]

where \(D\) is a bounded domain of \(\mathbb{R}^n\) with a smooth boundary \(\partial D\), \(F(q)\) has a continuous gradient \(\nabla F\) in \(D\), and \(F(D) \subset D\). These conditions means that Eqs. (2) define a \(C^1\)-smooth dynamical system with discrete time. Let \(\epsilon > 0\) be a small number. We say that networks (1) \(\epsilon\)-realize the map \(F\) if:

- **Ri** for some \(N, W, h\) there exists a locally attracting \(n\)-dimensional invariant manifold \(\mathcal{M}\) for dynamics (1) defined by a \(C^1\) - smooth map \(x = x(q), \ q \in D\);

- **Rii** on \(\mathcal{M}\) map (1) reduces to a map, which is an \(\epsilon\)-small perturbation of (2): \(q(t + 1) = F(q) + \epsilon F(q), \ \text{where} \ \sup_{q \in D} |F|, \ \sup_{q \in D} |\nabla F| < 1\).

Recall that the set \(\Gamma\) is invariant under the map \(F\) if \(q \in \Gamma\) entails \(F(q) \in \Gamma\). Existence of the attracting invariant manifold in the phase space of system means that there are "hidden modes" controlling dynamics of whole system, at least in an open subdomain of the phase space. Mathematically, the \(\epsilon\)-realization implies the following. If dynamics (2) has a structurally stable local attractor \(A\) on \(\mathcal{M}\), then for sufficiently small \(\epsilon\) network (1) also has a topologically equivalent attractor \(\tilde{A}\), with a dynamics topologically equivalent to dynamics of (2) on \(A\). If the attractor \(A\) of (2) is not structurally stable, then we can only assert that trajectories of (2) are \(O(\epsilon T)\) - small perturbations of trajectories of (1).
within $T$ iterations, where $T << \epsilon^{-1}$. The number $N$ of involved nodes depends on $\epsilon$ and $F$.

Let us formulate conditions on graphs $(V, E)$ important for $\epsilon$-realizations. Consider an infinite sequence of graphs $(V_N, E_N)$, where $N = 1, 2, \ldots$. Let us formulate some auxiliary definitions. We say that a graph $(V_N, E_N)$ contains $(n, d)$ biclique, if there are two disjoint subsets $S_n = \{v_1, \ldots, v_n\} \subset V_N$ and $S_d = \{w_1, \ldots, w_d\} \subset V_N$ of nodes such that all $v_i$ are mutually connected with all $v_j$: both edges $(v_i, w_j)$ and $(w_j, v_i)$ lie in $E_N$. It is not essential exist there connections within the sets $S_n$ and $S_d$, or not.

Assume this graph sequence enjoys the following property:

**Property nH (to have n hubs)** For each integer $N_0$ there is a graph from the graph sequence having an $(n, d)$ biclique with $d \geq N_0$.

This can be interpreted as follows. Assume a sequence of growing graphs $(V_N, E_N)$ is emerged by a procedure (for example, by the preferential attachment [[1]]). Our property holds, for example, if there is a fixed set of $n$ "central" nodes connected with an infinitely increasing number $d(N)$ of other nodes, $d(N) \to \infty$ as $N \to \infty$. If $d = N - n$ this means that there are $n$ "central" nodes. All the other nodes are connected with them, such network can be named centralized network with $n$ centers.

The main assertion can be formulated as follows:

**Theorem.** Let $q(t + 1) = F(q(t))$ and $q(t + 1) = \tilde{F}(q(t))$ be two maps defined on $D \subset \mathbb{R}^n$ satisfying above assumptions. Let us consider a sequence of graphs $(V_N, E_N)$, $N = 1, 2, \ldots$ with $nH$-property. Then

**I** for each $\epsilon > 0$ there is a number $N(\epsilon, F)$ such that network (1), associated with $N$-th graph, $\epsilon$-realizes dynamics $q(t + 1) = F(q(t))$ on $D$, under an appropriate choice of weights $W_{ij}$ and thresholds $h_i$;

**II** for each $\epsilon > 0$ there is a number $N(\epsilon, F, \tilde{F})$ such that a network (1), associated with $N$-th graph with appropriate $W_{ij}$ and $h_i$, $\epsilon$-realizes dynamics $q(t + 1) = F(q(t))$. Moreover, if we remove a node $x_j$ from this network, the modified network $\epsilon$-realizes the second map $q(t + 1) = \tilde{F}(q(t))$. The input and output degrees of this node $x_j$ are minimal: $C_{in}(j) = C_{out}(j) = 1$.

These results show that ecological, genetical and neural systems evolved by the preferential attachment growth [[1]] are fragile with respect to eliminations of some key nodes. Due to social and ecological evolution, humanity becomes a dominant center acting on many species, and ecological networks become more centralized, this can increase a danger of ecological catastrophes. Events that seem firstly nonessential, for example, some species vanishing, can provoke such a catastrophe.

**References**


Paper Q

Goodness-of-Fit Tests Based on Distribution Characterizations, and their Efficiencies

Author: K. Volkova

This work was supported by the RFBR grant No. 10-01-00154-a and by the grant FZP No. 2010-1.1-111-128-033.
Goodness-of-Fit Tests Based on Distribution Characterizations, and their Efficiencies

Ksenia Volkova

St. Petersburg State University,
St. Petersburg, Russia.
E-mail: efrksenia@gmail.com
1 Tests of Normality Based on Shepp Property, and their Efficiencies.

In 1964 L. Shepp discovered that if \( X \) and \( Y \) are two independent centered normal rv’s with some variance \( \sigma^2 > 0 \), then the rv \( k(X, Y) := 2XY/\sqrt{X^2 + Y^2} \) has again the normal distribution \( \mathcal{N}(0, \sigma^2) \) (Shepp property, [9].)

[2] proved, that the Shepp property implies the characterization of the normal law in a broad class of distributions. Consider the class \( \mathcal{F} \) of df’s \( F \) satisfying the conditions: i) \( 0 < F(0) < 1 \) and ii) \( F(x) - F(-x) \) is regularly varying in zero with the exponent 1.

**Theorem Q.1**

Let \( X \) and \( Y \) be independent rv’s with common df \( F \) from the class \( \mathcal{F} \). Then the equality in distribution \( X \overset{d}{=} k(X, Y) \) is valid iff \( X \in \mathcal{N}(0, \sigma^2) \) for some variance \( \sigma^2 > 0 \).

Let \( X_1, \ldots, X_n \) be independent observations with zero mean and df \( F \). We are interested in testing the composite hypothesis \( H_0 : F \in \mathcal{N}(0, \sigma^2) \) for some unknown variance \( \sigma^2 > 0 \) against the alternatives \( H_1 \), under which the hypothesis \( H_0 \) is false.

Let \( F_n(t) = (n)^{-1} \sum_{i=1}^{n} 1\{X_i < t\} \), \( t \in \mathbb{R}^1 \) be the ordinary empirical df based on the sample \( X_1, \ldots, X_n \). To estimate the df of the random variable \( k(X, Y) \) we will use the \( U \)-empirical df \( H_n \) using the formula

\[
L_n(t) = \left( n \right)^{-1} \sum_{1 \leq i < j \leq n} 1\{k(X_i, X_j) < t\}, \quad t \in \mathbb{R}^1.
\]

It is known that the properties of \( V \)- and \( U \)-empirical df’s are similar to the properties of usual empirical df’s, see [3]. Hence for large \( n \) the df’s \( L_n \) and \( F_n \) should be close under \( H_0 \), and we can measure their closeness using some test statistics.

We suggest two scale-invariant statistics:

\[
I_n = \int_{\mathbb{R}^1} (L_n(t) - F_n(t))dF_n(t) \quad \text{and} \quad D_n = \sup_{t \in \mathbb{R}^1} |L_n(t) - F_n(t)|.
\]

Note that \( I_n \) is a \( U \)-statistic, while \( D_n \) is a supremum of a family of \( U \)-statistics.

**Advantages of the tests based on characterization:**

- Statistics \( I_n \) and \( D_n \) are unbiased due to characterization.

- The Kolmogorov-type test of normality based on the statistic \( D_n \) is consistent against any alternative. The integral test based on the statistic \( I_n \) is not necessarily consistent for any alternative, but in fact is consistent for many of them. We will consider common parametric alternatives such that statistic \( I_n \) will be consistent for all of them.

- It is well-known that non-degenerate \( U \)- and \( V \)-statistics are asymptotically normal (Hoeffding’s theorem, [4]). The statistic \( I_n \) is asymptotically equivalent to the non-degenerate \( U \)-statistic, therefore it is asymptotically normal. The statistic \( D_n \) converges in distribution to the supremum of absolute value of some centered Gaussian process with complicated covariance and can be in principle obtained from [10].
We find their limiting distribution under the null hypothesis and describe the asymptotic normality of considered statistics. The analysis of comparative properties of statistics can be fulfilled by evaluation of asymptotic relative efficiency. For this purpose we find the logarithmic large deviation asymptotics of both sequences of statistics under $H_0$, this allows us to calculate the local Bahadur efficiency of both statistics for some common parametric alternatives.

We use the notion of local exact Bahadur efficiency (BE) [1], [5], because the statistic $R_n$ has the non-normal limiting distribution, and therefore the Pitman approach to efficiency is not applicable. However, it is known that the local BE and the limiting Pitman efficiency usually coincide, see [5].

Let $\phi(x), \Phi(x)$ - respectively the density and df of standart normal law. We consider the standard alternatives used in testing of normality:

- Location alternative with the df $\Phi(x - \theta), x \in R^1$;
- Skew alternative in Azzalini’s sense with the density $2\phi(x)\Phi(\theta x), x \in R^1$;
- Contamination alternative with the df $(1 - \theta)\Phi(x) + \theta \Phi^2(x), x \in R^1$.

The result of calculation of local BE values we collect in the Table 1.

<table>
<thead>
<tr>
<th>alternative</th>
<th>efficiency of $I_n$</th>
<th>efficiency of $D_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>$3/\pi \approx 0.955$</td>
<td>$2/\pi \approx 0.637$</td>
</tr>
<tr>
<td>Azzalini’s skew</td>
<td>$3/\pi$</td>
<td>$2/\pi$</td>
</tr>
<tr>
<td>Contamination</td>
<td>0.417</td>
<td>0.313</td>
</tr>
</tbody>
</table>

We can remark that the efficiencies of our two tests for the composite hypothesis $H_0$ of normality under location alternative (respectively $3/\pi$ and $2/\pi$) coincide with the efficiencies under the same alternative of classical tests based on the Chapman-Moses statistic $\omega_n^1$ and Kolmogorov statistic when testing the simple hypothesis of normality (they are known since 1970-s).

2 Tests of Exponentiality Based on Rossberg’s Characterization, and their Efficiencies.

Consider a sample $X_1, \ldots, X_n$ of non-negative i.i.d. rv’s with continuous df $F$. We are testing the composite null hypothesis $H_0 : F \in \mathcal{E}(\lambda)$, where $\mathcal{E}(\lambda)$ denotes the class of exponential distributions with unknown scale parameter $\lambda > 0$.

We propose two new tests of exponentiality based on Rossberg’s characterization of exponential distribution [8], [7]. Its simplest formulation is as follows:

**Theorem Q.2**

Let $X_1, X_2$ and $X_3$ be i.i.d. rv’s with continuous df $F$. Two statistics $X_{2:3} - X_{1:3}$ and $\min(X_1, X_2)$ are identically distributed iff $F$ is the exponential df.
In accordance with this characterization we construct two $U-$empirical df’s $H_n$ and $G_n$ using the formulas

\[ H_n(t) = \left( \begin{array}{c} n \end{array} \right)^{-1} \sum_{1 \leq i < j < k \leq n} \mathbb{1}\{X_{2\{i, j, k\}} - X_{1\{i, j, k\}} < t\}, \quad t \geq 0, \]

\[ G_n(t) = \left( \begin{array}{c} n \end{array} \right)^{-1} \sum_{1 \leq i < j \leq n} \mathbb{1}\{\min(X_i, X_j) < t\}, \quad t \geq 0. \]

We suggest two statistics for testing of $H_0$:

\[ S_n = \int_0^\infty (H_n(t) - G_n(t)) \, dF_n(t) \quad \text{and} \quad R_n = \sup_{t \geq 0} |H_n(t) - G_n(t)|, \]

where $F_n$ is the usual empirical df. As previously the statistic $S_n$ is asymptotically equivalent to the non-degenerate $U$-statistic, therefore it is asymptotically normal. The statistic $R_n$ converges in distribution to the supremum of absolute value of some centered Gaussian process with complicated covariance, [10].

We consider the standard alternatives used in testing of exponentiality with the densities $g_1, g_2, g_3$:

- Weibull with $g_1(x, \theta) = (1 + \theta)x^\theta \exp(-x^{1+\theta})$, $x \geq 0$;
- Makeham with $g_2(x, \theta) = (1 + \theta(1 - e^{-x})) \exp(-x - \theta(e^{-x} - 1 + x))$, $x \geq 0$;
- linear failure rate with $g_3(x, \theta) = (1 + \theta x)e^{-x - \frac{1}{2}\theta x^2}$, $x \geq 0$.

The result of calculation of local BE values we collect in the Table ??.

**Table 2: Local Bahadur efficiencies of the statistic $S_n$ and $R_n$ under alternatives.**

<table>
<thead>
<tr>
<th>alternative</th>
<th>efficiency of $S_n$</th>
<th>efficiency of $R_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull</td>
<td>0.650</td>
<td>0.320</td>
</tr>
<tr>
<td>Makeham</td>
<td>0.452</td>
<td>0.207</td>
</tr>
<tr>
<td>Linear failure rate</td>
<td>0.119</td>
<td>0.047</td>
</tr>
</tbody>
</table>

In both examples we observe that the efficiencies for the Kolmogorov test are lower than for the integral test. However, it is the usual situation when testing goodness-of-fit, [5], [6], and this drawback partially compensated the fact, that Kolmogorov-type statistics consistent against any alternatives.

**References**


