MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

T a g u n g s b e r i c h t 01/1999

Inverse Problems in Statistics

03.01. - 09.01.1999

The conference was organized by F. Ruymgaart (Lubbock), W. Stute (Giessen) and Y. Vardi (New Brunswick). 43 participants attended the meeting, and 32 talks were given.

The aim of the conference was to bring together experts working in the field of inverse problems, with special emphasis on statistical methodology and applications. The unique facilities of the Oberwolfach institute provided an excellent opportunity for an exchange of ideas and significantly contributed to the success of the meeting.

Some topis which were discussed in greater detail were:

- Applications to Image Analysis
- Inverse Problems in Survival Analysis
- Curve Estimation under Constraints
- Problems in Tomography
- Statistical Aspects of Inverse Problems
- Deconvolution Techniques
- Algorithms for Inverse Problems

Bob R. S. ANDERSSEN

Linking Mathematics to Data with Exemplifications from Human Movement, Electromagnetic Induction in the Earth and Molecular Weight Distribution Determination

Operationally, applied mathematics is the dynamic activity of linking mathematics to data: for given data

(i) identify the question to be answered;

(ii) formulate a model which relates the data to the phenomenon of interest (i.e. the information which encapsulates the answer);

(iii) recover the phenomenon of interest from the available data; and finally,

(iv) construct an answer to the question from the resulting information.

The solution of “inverse problems” and the application of “statistics” are closely linked to all stages of this dynamic process:

(a) the model formulation may involve stochastic considerations;

(b) the input data to the model must, often, be deconvolved from available and related measurements;

(c) the recovery of an estimate of the phenomenon of interest involves the solution of an improperly posed problem and must be performed with finite and noisy data.

These and related matters, along with some of the challenges they pose, were illustrated and analysed within the frameworks of the following examples:

- the numerical differentiation of human movement data to recover accelerations on different scales in walking motion;
- the recovery of electrical conductivity as a function of depth within the Earth from measurements of the Earth’s transient magnetic field;
- the role of the molecular weight distribution as a molecular characterization in polymer dynamics including wheat-flour tough mixing and bread making;
- the deconvolution and analysis of field-flow data in the experimental recovery of molecular weight distributions.

Viktor BENEŠ

A Stereological Inverse Problem for a Function of Three Variables

The problem consists in stereological unfolding of spatial geometrical parameters of spheroidal particles from planar parameters observed in vertical uniform random section planes. A theoretical relation between corresponding probability densities as functions of three variables (size, shape factor, orientation) is derived. Results on real data from metallography are presented and statistical problems discussed.

Rudolf BERAN

Basis Economy in REACT Fits to Linear Models

REACT estimators for the mean of a Gaussian linear model use model-selection, shrinkage, and ideas from signal-processing to exploit the superefficiency loophole in classical parametric information bounds. REACT methods may also be viewed as smoothing techniques. The acronym sketches the steps in the methodology: Risk Estimation and Adaptation after Coordinate Transformation. If a linear combination of the first few vectors in the transformed regression basis closely approximates the unknown mean vector, then the asymptotic maximum risk of a monotone-shrinkage REACT estimator greatly undercuts that of the least squares estimator. In experiments on scatterplots found in the smoothing literature, REACT fits draw remarkable benefit from the economy of some natural regression bases.

P. BICKEL

Inference in Hidden Markov Models (HMM)

Let $X_1, \ldots, X_n$ be a realization of a K state Markov chain with transition probability function $\alpha_\theta(\cdot, \cdot)$ and $Y_1, \ldots, Y_n$ be conditionally independent given $X_1, \ldots, X_n$ with Y_i independent of $(X_k; Y_k)$ $k \neq i$ given $X_i = x$ and conditional density $g_\theta(y|x)$. Then $Y_1, \ldots, Y_n$ are distributed according to an (HMM). Under suitable conditions, if $l_\theta(y_1, \ldots, y_n)$ is the joint log likelihood we derive bounds on $l_\theta^{(j)}(y_1, \ldots, y_n)$, where $l^{(j)}$ is the jth derivative of the form $nM^j j!$ as well as limiting results for

$$\frac{1}{n} l_\theta^{(j)}(Y_1, \ldots, Y_n) \text{ and } \frac{1}{n}(l_\theta^{(1)}(Y_1, \ldots, Y_n) - E l_\theta^{(1)}(Y_1, \ldots, Y_n)).$$

These results are obtained both under the hypothesized model and also under the assumption that the Y_i's are generated by a geometrically mixing stationary process and are used to obtain consistency and asymptotic normality for the MLE. These results can be viewed as extensions with a simplified proof of results of Bickel, Ritov, Ryden (1998).

Nicolai BISSANTZ

Non-parametric Deprojection in Galaxy Research

Non-parametric deprojection methods are widespread in astrophysics. Some of the main applications include the search for Black Holes in the centres of galaxies and the analysis of surface-brightness and kinematical data from galaxies.
One important example is our Milky Way, which I will focus on in this talk. The DIRBE experiment on the COBE-satellite has produced maps of the Milky Way in several Near-Infra-Red (NIR)-bands. These maps have been corrected against several foregroundeffects, including zodiacal light, dust absorption and point sources (Spergel et al, 1996). We used this corrected data to produce three-dimensional light-distribution models of the Milky Way with non-parametric deprojection methods, including the Richardson-Lucy deprojection and penalized maximum likelihood estimation. The models that result from these fitting procedure, which consists of several steps, fit the observed data better than available parametric fits.

Laurent CAVALIER

Sharp Adaptation for Inverse Problems with Random Noise

Using the singular value decomposition (SVD), we transform a general model of inverse problem with random noise into an infinite-dimensional sequence space model. In this model, we construct an adaptive estimator that mimics the asymptotical behaviour of the best linear oracle chosen among a finite class. We prove that this estimator is sharp adaptive on different ellipsoids $\Theta(\alpha, Q)$ for α and Q unknown, i.e. it asymptotically attains the exact minimax L_2-risk.
We apply this method to construct sharp adaptive estimators in two examples of practical inverse problems: tomography and deconvolution.

Yair CENSOR

Dykstra's Algorithm for Finding the Projection of a Point onto an Intersection of Closed Convex Sets and its Extension to Bregman Projections

We describe Dykstra's algorithm (invented [1] in 1983, rediscovered [2] by Han in 1988, see also [3]) which deflects the current point prior to projecting it onto the next set and achieves in this way not just feasibility but convergence to the projection of the initial point onto the (assumed nonempty) intersection of a given finite family of closed convex sets in the Euclidean space.
The close relation of this algorithm with Hildreth's norm minimization method and other techniques is discussed. Recently [4,5] we extended this method to encompass Bregman projections. In doing so we show that Dykstra's algorithm with Bregman projections is actually the nonlinear extension of Bregman's primal-dual coordinate ascent row-action minimization method.
REFERENCES
[1] R.L. Dykstra, "An algorithm for restricted least squares regression", Journal of the American Statistical Association, 78 (1983) 837-842.
[2] S.P. Han, "A successive projection method", Mathematical Programming, 40 (1988) 1-14.
[3] T. Robertson, F.T. Wright, and R.L. Dykstra, ORDER RESTRICTED STATISTICAL INFERENCE, John Wiley & Sons, Chichester, UK, 1988.
[4] Y. Censor and S. Reich, "The Dykstra algorithm with Bregman Projections", Communications in Applied Analysis, 2 (1998) 407-419.
[5] L.M. Bregman, Y. Censor, and S. Reich, "Dykstra's algorithm is the nonlinear extension of Bregman's optimization method", Technical Report, March 23, 1998.

P.P.B. EGGERMONT

Maximum Penalized Likelihood Estimation

We consider/compare least squares and maximum penalized likelihood approaches to the standard nonparametric problem of estimating the density $f_0(x)$ of the random variable X from the i.i.d. sample $X_1, X_2, \ldots, X_n$. The least squares estimators we consider are solutions to

$$\text{(1)}\quad \begin{aligned} &\text{minimize} \int_{\mathbb{R}} |f(x)|^2 dx - 2\int_{\mathbb{R}} f(x) dF_n(x) + h^2 \parallel \mathcal{L}f \parallel_2^2 \\ &\text{subject to } f \in L^2(\mathbb{R}), f \text{ is a density} \end{aligned}$$

where $\mathcal{L}$ is some differential operator, h is the smoothing parameter. Here F_n is the empirical distribution function. The special choice $\mathcal{L}f(x) = f'(x)$ for $\mathcal{L}$ leads to the estimator

$$f_{nh}(x) = \mathcal{B}_h * dF_n = \frac{1}{n} \sum_{i=1}^{n} \mathcal{B}_h(x - X_i) \tag{2}$$

where $\mathcal{B}_h(x) = (2h)^{-1} \exp(-h^{-1}|x|)$ is the two-sided exponential. Thus, kernel density estimators in general may be thought of as least squares estimators. Since kernel estimators are very good by just about any criterion, it would seem that there is not much point in considering the maximum penalized likelihood estimator, the solution, denoted by f_{nh}, to

$$\text{(3)}\quad \begin{aligned} &\text{minimize } -\int_{\mathbb{R}} \log f(x) dF_n(x) + \int_{\mathbb{R}} f(x) dx + h^2 R(f) \\ &\text{subject to} f \in L^1(\mathbb{R}), f \geq 0. \end{aligned}$$

(note that the density constraint is <u>not</u> enforced.) Here $R(f)$ is the roughness penalization functional of GOOD (1971), $R(f) = \parallel (\sqrt{f})' \parallel_2^2$. KLONIAS (1982, 1984) derived bounds for the estimation error in terms of the Hellinger distance, which were not quite satisfactory. However, it can be shown that

$$\parallel f_{nh} - \mathcal{B}_{h/\sqrt{2}} * dF_n \parallel_1 = h^2 \parallel (\sqrt{f_{nh}})' \parallel_2^2,$$

and this implies the same order bounds on the L^1 error as for kernel estimators. Simulation experiments indicate that f_{nh} is practically indistinguishible from the kernel density estimator with the optimal Epanechnikov kernel. Thus the two approaches are not very different in the end.

Things are even more similar when considering the estimation of monotone or convex densities on $(0, \infty)$, by adding the monotonicity or convexity constraint to the minimization problems (1) and (3). For monotone densities, with $h = 0$, the problems (1) and (3) are equivalent [Grenander (1956)]. The equivalence still holds if F_n is replaced by $A_\lambda * dF_n$ for some kernel A with smoothing parameter λ (and $h = 0$). Moreover, something special happens. The map lcm defined on all densities on $(0, \infty)$ by

$$\text{lcm} : \varphi \longmapsto \text{ the derivative of the least concave majorant of } \int_0^x \varphi(t) dt$$

is a contraction in all L^p norms $(1 \leq p \leq \infty)$, as well as in Hellinger, Kullback-Leibler, and Pearson's χ^2 distances. This immediately implies error bounds on the estimator.

For convex density estimation the problems (1) and (3) with $h = 0$ are again equivalent, and likewise when F_n is replaced by $A_\lambda * dF_n$. This leads to error bounds on the estimator in L^2 error only.

For nonparametric deconvolution theoretical comparisons between the least squares and maximum penalized likelihood approaches are extremely difficult, but simulation studies strongly suggest the superiority of the latter.

Andrey FEUERVERGER

On Kugelverteilung Problems Revisited

We consider the classical "thin-slice" stereological unfolding problem of Wicksell (Biometrika, 1925) as well as its "thick-slice" version in which the section thickness is comparable to the actual radii of the spheres themselves. (Bach, 1967; Goldsmith 1967). It is pointed out that in a statistical sense, these two problems cannot be treated by analogy to each other, for example, the thin-slice version requires a heavier order of smoothing than classical kernel density estimators in order to attain optimal performance, while the thick-slice version does not! In that statistical sense, the thick-slice problem is hardly ill-posed at all. We also stress that the nice linear relations between the folded and unfolded distributions allow us to write an expectation with respect to one of these distributions as an expectation with respect to the other. This fact has been underutilized in statistical applications of Wicksell-type problems. Thus, for example, it is possible to obtain an unbiased estimator for the unfolded distribution and this may be used as the basis for statistical inference. Another possiblity is to implement ideas related to cross-validation for bandwidth selection. The properties of a number of such procedures are considered.

Alexander GOLDENSHLUGER

On Pointwise Adaptive Nonparametric Deconvolution

We consider estimating an unknown function f from indirect white noise observations with particular emphasis on the problem of nonparametric deconvolution. Nonparametric estimators that can adapt to unknown smoothness of f are developed. The adaptive estimators are specified under two sets of assumptions on the kernel of the convolution transform. In particular, kernels having the Fourier transform with polynomially and exponentially decaying tails are considered. It is shown that the proposed estimates possess, in a sense, the best possible abilities for pointwise adaptation.

Piet GROENEBOOM

Deconvolution and Late Stopping

Let $Z_1, \ldots, Z_n$ be a sample generated by the density h_0, given by

$$h_0(z) = \int k(z-y)dF_0(y),$$

where

$$k(x) = b^{-1}1_{[0,b]}(x), x \in \mathbb{R}.$$

The distribution function F_0 can be estimated by the nonparametric maximum likelihood estimator (NPMLE) $\hat{F}_n$ that maximizes the log likelihood

$$\sum_{i=1}^{n} \log\left\{\int k(Z_i - y)dF(y)\right\} = \sum_{i=1}^{n} \log\left\{\int k(Z_i - y)f(y)dy\right\}$$

over all *distribution functions* F. In this situation it makes no sense to estimate the density f_0 by straightforward maximum likelihood, analogously to the situation of a directly

observable sample $X_1, \ldots, X_n$, where straightforward maximum likelihood estimation of the density leads to point masses $1/n$ at the observation points.
On the other hand, the NPMLE $\hat{F}_n$ is a perfectly well-defined estimate of F_0 that can be computed by several methods, for example by EM with very late stopping. It can be proved that one can also use the NPMLE $\hat{F}_n$ for estimating f_0 by introducing

$$\hat{f}_{n,h}(t) = \int K_h(t-x)\, d\hat{F}_n(x) \qquad (1)$$

where $K_h(x) = h^{-1}K(x/h)$ for a continuously differentiable symmetric kernel K with support $[-1,1]$ that integrates to 1.
If f_0 is twice differentiable at $t \in (0, M)$, one can estimate $f_0(t)$ at rate $n^{-2/7}$, using (1) and taking $h \sim c \cdot n^{-1/7}$. This is the usual rate for estimating the *derivative* of a density, unless more restrictive smoothness conditions are imposed on the density. So, although we actually try to estimate a density, the inverse problem character of the estimation problem causes the slower rate. The asymptotic distribution of $n^{2/7}\{\hat{f}_{n,h_n}(t) - f_0(t)\}$, where $h_n \sim c \cdot n^{-1/7}$, can also be explicitly determined, see GROENEBOOM (1998). The result is used in ROY CHOUDHURY (1998) for "deblurring" images that are blurred by Poisson noise. References: GROENEBOOM, P. (1998). *Nonparametric estimators for inverse problems. Algorithms and Asymptotics.* Technical Report 344, Department of Statistics, University of Washington. ROY CHOUDHURY, K. (1998). *Additive Mixture Models for Multichannel Image Data.* Ph. D. Dissertation, University of Washington, Seattle.

Martin B. HANSEN

Nonparametric Bayes Inference for Concave Distribution Functions

A way of making Bayesian inference for concave distribution functions is introduced. This is done by uniquely transforming a mixture of Dirichlet processes on the space of distribution functions to the space of concave distribution functions. The approach also gives a way of making Bayesian analysis of multiplicatively censored data. We give a method for sampling from the posterior distribution by use of a Pólya urn scheme in combination with a Markov chain Monte Carlo algorithm. The methods are extended to estimation of concave distribution functions for incompletely observed data. Finally, consistency issues are touched upon.

N. W. HENGARTNER

A Universal Oracle Inequality for the L_1-Norm

We study the problem of estimating functions in a prescribed class $\mathcal{G}$, that are closest to either the true regression g_0, or its derivative g_0', in the L^1 distance. For both problems, we propose an estimator $\hat{g}$ for which the mean absolute error is bounded by

$$\frac{1}{n}\sum_{i=1}^{n} |\hat{g}(X_i) - g_0(X_i)| \leq 3 \inf_{g \in \mathcal{G}} \frac{1}{n}\sum_{i=1}^{n} |g(X_i) - g_0(X_i)| + R_n$$

and give bounds for the tail probability of the remainder that depend on the Kolmogorov entropy of the class $\mathcal{G}$. The methodology is general and readily generalizes to different L_p metric for $1 \leq p \leq 2$. These results are universal for they hold for any regression function g_0!

A. HERO

Asymptotic Minimax Methods for Incorporation of Uncertain Side Information into Penalized ML Image Reconstructions

We present a methodology for incorporating extracted MRI anatomical boundary information to improve the performance of penalized likelihood (PL) ECT image reconstruction and ECT tracer uptake estimation. Under the assumption of perfect errorless side information it is natural to use a spatially variant quadratic Gibbs penalty which enforces smoothness everywhere in the ECT image except across the MRI-extracted boundary of the ROI. When high quality estimates of the anatomical boundary are available and MRI and ECT images are perfectly registered, the performance of this Gibbs penalty method is very close to that attainable using perfect side information, i.e., an errorless anatomical boundary estimate. However when the variance of the MRI-extracted boundary estimate becomes significant this method performs poorly. We derive a modified Gibbs penalty function which accounts for errors in side information based on a min-max robustness approach. The resulting penalty is implemented with a set of averaged Gibbs weights where the averaging is performed with respect to a limiting form of the induced posterior distribution of the MRI boundary parameters. Examples will be presented for tracer uptake estimation using the SAGE version of the EM algorithm and a B-spline parameterisation of the anatomical boundary.

David HUNTER

Optimization Transfer Algorithms in Statistics

Optimization transfer is an iterative technique for solving difficult optimization problems. Using convexity arguments, it is often possible to construct, for a given objective function $L(\theta)$ and parameter value θ^k, a surrogate function $Q(\theta|\theta^k)$ with the following property: By driving uphill (or downhill, if the goal is minimization instead of maximization), one drives the value of $L(\theta)$ uphill (or downhill). Thus optimization of $L(\theta)$ is transferred to the surrogate function, which is constructed to be easier to optimize. As with the EM algorithm, which is a special case of optimization transfer, the resulting algorithms are numerically stable, simple to understand and code, and often very fast, where speed is measured not in number of iterations but in total computation required. The talk outlines general principles of optimization transfer, cites examples of their use, and presents a quasi-Newton method for accelerating their convergence.

Geurt JONGBLOED

Isotonic Inverse Estimation in Statistical Inverse Problems

When estimating a distribution function in a statistical inverse problem, plugging in some nonparametric estimator of the sampling distribution into the inverse relation often yields an estimator of the distribution function that does not satisfy the basic properties of distribution functions. Usually it is the monotonicity property that is violated. Based on such an unsatisfactory estimator, one can construct estimators that are monotone, using techniques from the theory of isotonic regression.
The (simple) general idea of such estimators will be discussed and illustrated using some examples.

H. P. LOPUHAÄ

Asymptotic Normality of L_1-Error of the Grenander Estimator

Let f be a decreasing density on $[0,1]$ and denote f_n the non-parametric maximum likelihood estimator for f. In Groeneboom (1985) an elegant proof, based on the so-called inverse process

$$U_n(a) = \sup\{t \in [0,1] : F_n(t) - at \text{ is maximal}\},$$

was given for Prakasa Rao's (1969) result on the limit distribution of $f_n(t_0)$ with $t_0 \in (0,1)$ fixed. Groeneboom also indicated how one might use the process U_n to obtain asymptotic normality of $\| f_n - f \|_1$. Although this result is often quoted and referred to as being well known, no rigorous proof has ever been given.
In this talk I will give a brief outline of a rigorous proof of asymptotic normality of $\| f_n - f \|_1$, which deviates from the guidelines given in Groeneboom (1985). This was joint work with P. Groeneboom and G. Hooghiemstra. We feel that this result is important since the problem of estimating a monotone density is closely related to several other inverse problem, e.g. estimating the distribution function of interval censored observations, or estimating a monotone hazard.

Bernard A. MAIR

Reconstruction of Positron Emission Tomography Images

This talk discusses new statistical reconstruction algorithms for positron emission tomography images. In particular, we will introduce several new algorithms for finding a Pearson type estimator instead of the usual maximum likelihood estimator. Some of these algorithms outperform the usual ML-EM algorithm. In addition we introduce a new statistical technique for dealing with data corrupted by attenuation. This method uses the emission and transmission prompts and blank scan data to reconstruct both emission and attenuation maps for chest phantoms. We will present the results of applying these algorithms to data obtained from actual PET scanners.

Axel MUNK

Model Checks in Indirect Regression Models

In this paper we propose a class of asymptotic tests for the validity of linear model assumptions of the regression function f in inverse estimation models $Y = \mathbf{K}f(V) + \epsilon$ under a random and fixed design assumption, respectively. In particular, this approach allows to test hypotheses which are L^2-neighborhoods of the classical null hypothesis $H : f \in \mathcal{L}$ where $\mathcal{L}$ denotes a specific parametric regression model. Various applications are considered in detail, which includes new test procedures for the direct case. These tests are obtained by preconditioning the inverse regression model with a suitable operator $\mathbf{T}$ which avoids direct estimation of f. Different choices of $\mathbf{T}$ are investigated in detail and their efficiencies are compared.

Georg NEUHAUS

On Testing Tumour Onset Times in a Lethal-Incidental Context

The asymptotic distribution of the combined test of Peto et al. (1980) under the null hypothesis and under local alternatives is derived. A corrected variance estimator is introduced and the implications of different cell numbers for the Mantel-Haenszel part of the combined test statistic are discovered. Moreover, it is shown that one may use data dependent cells.
Furthermore, a conditional version of the combined test is introduced being exactly valid under the strict null hypothesis and being asymptotically equivalent to its unconditional counterpart. By a simulation scenario from a real data situation it is demonstrated in what situations the asymptotics applies.

Robert D. NOWAK

A Multiscale Framework for Poisson Inverse Problems

Many important problems in science and engineering are well modeled by Poisson processes. Often, it is of interest to accurately estimate the intensities underlying observed Poisson data. This talk describes a maximum a posteriori (MAP) estimation method for linear inverse problems involving Poisson data based on a novel multiscale framework. The framework itself is founded on a carefully designed multiscale prior probability distribution placed on the "splits" in the multiscale partition of the underlying intensity, and it admits a remarkably simple MAP estimation procedure using an expectation-maximization (EM) algorithm. Unlike many other regularized or Bayesian approaches to this problem, the EM update equations for our algorithm have simple, closed-form expressions. Additionally, our class of priors has the interesting feature that the "non-informative" member yields the traditional maximum likelihood algorithm; other choices are made to reflect prior belief as to the smoothness of the unknown intensity. It can also be shown that the informative priors display 1/f spectral characteristics, which suggests that they may be very reasonable models for natural intensities functions like images.

Douglas NYCHKA

Numerical Weather Prediction

Prediction of meteorological conditions based on the previous state of the atmosphere and observations is an inverse problem. Here we outline some of the difficulties in transfering usual statistical methods to this context. Some issues are the sequential nature of weather prediction and the large size of the problem. We suggest an aproach based on mixtures of Gaussian densities as a form for the prior distribution of the atmospheric state. This structure helps to interpret current practice using ensembles of states and poses some interesting approximation problems for theoretical work.

Joachim OHSER

Stereological Unfolding of Particle Size Distributions

The stereological unfolding problem for systems of spatially distributed particles is discussed. It is assumed that all particles have the same shape but different sizes. In classical stereology the only information used comes from the sizes of the section profiles observed in a planar section of the particle system. Here we additionally consider either a shape parameter, or the complete shape information of the section profiles. Clearly, if the particles are spheres then the profile shape adds no information about particle size. If the particles are oblate spheroids, then the size and shape of the typical section ellipse are independently distributed. In general, however, the size and shape of the section profiles are not independent and thus the shape information can be used to refine the conventional stereological estimation of particle size distributions. An estimator of the particle size distribution is given which is best linear unbiased in some sense. As an example a system of cubic particles with random edge length is studied. Here additional shape information is provided by the number of vertices of the section polygons. In general, the shape of a section polygon is not completely characterized by a single shape parameter. Complete shape information is provided by the coordinates of the vertices of the section polygons. If the number of vertices of a section polygon is greater than 3 then the size of the corresponding cube can be computed directly from the polygon shape. In the triangle case only a lower bound of the cube size can be calculated, and this lower bound is always greater than the size of this triangle. Such geometrical considerations lead to several estimators of the cube size distribution which are compared using the mean square l^2-distance.

Sergei PEREVERZEV

Optimal Discretization and Degrees of Ill-Posedness for Inverse Estimation in Hilbert Scales in the Presence of Random Noise

The problem of inverse estimation of an unknown element x_0 from noisy observations $y_0 = A_{x_0} + \delta\xi$ in dependence of the natur of random noise ξ is considered. It is shown that a combination of a Tikhonov regularization estimator with some projection scheme is order optimal for a wide class of operators A acting along Hilbert scales.

Frits H. RUYMGAART

Some Results and Some Open Problems Regarding Inverse Statistical Inference

Two approaches to a general theory of inverse statistical inference will be briefly discussed. One is based on Halmos' version of the spectral theorem in the traditional model, where the unknown parameter is a smooth curve. The construction of estimators and questions like optimal convergence rates, weak convergence and asymptotic efficiency of smooth functionals can be dealt with. Some of these results as well as some open problems in this area will be considered. A very promising alternative approach suggested by Hall (personal communication) is to combine Halmos' spectral theorem in a suitable manner with a wavelet expansion. This approach might be particularly appropriate if the unknown parameter is not necessarily a smooth curve. A curious example, not covered by Donoho's wavelet-vaguelette method, is provided by convolution with the "boxcar". This example will be used to demonstrate both the potential and some of the prospective difficulties of this alternative.

A. W. VAN DER VAART

Censoring and Passive Registration

We discussed the estimation of the life time distribution of people in England in the period 1580-1800 from a data base that was put together through "familiy reconstitution". The data base, consisting of 100 000 entries, was prepared from handwritten registers from a large number of parishes in England, which recorded events such as births, marriages and deaths. Unfortunately, the times of death of almost 50 percent of the people in the data base are missing. This is thought to be due to moving of these persons to other parishes that are not in the data-base or whose records could not be connected to those of the parish of birth. Because the time of moving is not recorded in the data-base, the data cannot be analysed using standard techniques for censored data, such as the Kaplan Meier estimator. Following Gill we introduced a semiparametric model for which we discussed the asymptotic behaviour of the maximum likelihood estimators. In this model it is essential that the person is known to have been alive at a number of registered events, such as births of children, even though the person's death may be unobserved. The last moment that the person was known to be alive underestimates the time of censoring (i.e. moving), but the discrepancy can be modelled. When modelling the registration events as a Poisson process, this discrepancy is exponentially distributed and maximum likelihood estimation is shown to come down to deconvolving this exponential distribution, an inverse problem. To obtain our results we showed, among others, that the rate of convergence of the Grenander estimator for a monotone density relative to the uniform norm is of the order $(n/\log n)^{-1/3}$.

Erik VAN ZWET

Perfect Stochastic EM

In a missing data problem we observe the result of a (known) many-to-one mapping of an unobservable 'complete' dataset. The aim is to estimate some parameter of the distribution of the complete data. In this situation, the stochastic version of the EM algorithm is sometimes a viable option. It is an iterative algorithm that produces an ergodic Markov chain on the parameter space. The stochastic EM (StEM) estimator is then a sample from the equilibrium distribution of this chain. Recently, a method called 'coupling from the past' was invented to generate a Markov chain in equilibrium. We investigate when this method can be used for a StEM chain and give a simple example where this is indeed possible.

W.R. VAN ZWET

Statistical Estimation for the Contact Process

The contact process is a highly simplified model for the spread of an infection. At any time t, each point of the d-dimensional lattice is either infected or healthy. As time passes, the system changes according to the following dynamics: every infected site may infect each of its 2d immediate neighbors with Poisson rate m and becomes healthy with rate 1. All of the Poisson processes involved are independent. The process starts at time $t = 0$ with e.g. a single infected site at the origin.
Simple as this description may be, the contact process is a difficult object for probabilistic study. However, some basic facts have been discovered by probabilists over the past decades. Here we shall deal with the problem of statistical estimation of the parameter m on the basis of an observation of the process at a single time t. We propose an estimator and show that it is consistent and asymptotically normal as t tends to infinity.

Yehuda VARDI

Estimating Mixing Probabilities Subject to Lower and Upper Bound Constraints with Applications

Consider k *known* probability vectors, $h_i \in \mathbb{R}^m, i = 1, \ldots, k$, let f be an *unknown* probability vector $f \in \mathbb{R}^k$ with *known* lower and upper bounds a and b, respectively ($a, b \in \mathbb{R}^k$) and let $g \in \mathbb{R}^m$ be a probability vector obtained by mixing the h_i's with the probability-weights f_i.
We derive the maximum likelihood estimate (MLE) of the mixing probability vector f, subject to $a \leq f \leq b$, based on a random sample from g. The MLE is derived using the EM algorithm, and it generalizes the much studied EM algorithm in emission tomography linear-inverse problems with positivity constraints and many other applications to include lower and upper bound constraints on the vector f. Applications in statistics, signal recovery, finance, discrete tomography and more are discussed. Convergence of the iterated log-likelihood values, as well as the iterated probability vectors (f's), is proved.

Gerhard WINKLER

Recent Developments in Edge-Preserving Noise Reduction

Some recent and very recent methods and algorithms are presented, discussed and compared. The conclusions are illustrated by way of examples from medical imaging like fMRI and others.
We focus on approaches like the Bayesian, local M-smoothing, adaptive weights smoothing and chains of nonlinear Gaussian filters.

Cun-Hui ZHANG

Risk Bounds in Isotonic Regression

Nonasymptotic risk bounds are provided for maximum likelihood-type estimators of a nearly increasing unknown function, with general average loss at design points. These bounds are optimal up to scale constants in a certain sense, and they imply uniform $n^{-1/3}$-consistency for uniformly bounded unknown functions under mild assumptions on the stochastic structure of the data.

Jacques ZUBER

A Goodness-of-fit Test for Nonlinear Models Based on Nonparametric Techniques

The talk is devoted to goodness-of-fit tests for parametric possibly nonlinear heteroscedastic regression models. The test statistic is constructed using a marked empirical process based on residuals. We investigate the consistency of this test statistic and of the estimators needed to compute it. We illustrate our results with numerical experiments and comparisons to other tests.

Berichterstatter: Winfried Stute (Giessen)

Tagungsteilnehmer

Prof.Dr. Robert S. Anderssen
C.S.I.R.O.
Math. and Information Sciences
G.P.Box 664

Canberra ACT 2601

Dr. Eric Cator
Department of Mathematics and
Computer Science
Delft University of Technology
Mekelweg 4

NL-2628 CD Delft

Dr. Viktor Benes
Department of Probability and
Mathematical Statistics
Charles University
Sokolovska 83

18675 Praha 8
CZECH REPUBLIC

Dr. Laurent Cavalier
c/o Prof. A. Tsybakov
Lab. de Probabilities
Universite Paris 6
4, pl. Jussieu

F-75252 Paris Cedex 05

Prof.Dr. Rudolf J. Beran
Department of Statistics
University of California
367 Evans Hall

Berkeley , CA 94720-3860
USA

Prof.Dr. Yair Censor
Dept. of Mathematics and Computer
Sciences
University of Haifa
Mount Carmel

Haifa 31905
ISRAEL

Prof.Dr. Peter J. Bickel
Department of Statistics
University of California
367 Evans Hall

Berkeley , CA 94720-3860
USA

Prof.Dr. Paul Eggermont
Department of Mathematical Sciences
University of Delaware
501 Ewing Hall

Newark , DE 19716-2553
USA

Dr. Nicolai Bissantz
Astronomisches Institut
der Universität Basel
Venusstr. 7

CH-4102 Binningen

Gerrit Eichner
Mathematisches Institut
Universität Gießen
Arndtstr. 2

35392 Gießen

Dr. Dietmar Ferger
Institut für Mathematische
Stochastik
Technische Universität Dresden

01062 Dresden

Prof.Dr. Andrey Feuerverger
Dept. of Statistics
University of Toronto

Toronto, Ontario M5S 3G3
CANADA

Prof.Dr. Peter Gänßler
Mathematisches Institut
Universität München
Theresienstr. 39

80333 München

Prof.Dr. Alexander Goldenshluger
Dept. of Statistics
University of Haifa

Haifa
ISRAEL

Prof.Dr. Piet Groeneboom
Department of Mathematics and
Computer Science
Delft University of Technology
Mekelweg 4

NL-2628 CD Delft

Prof.Dr. Martin Hansen Bogsted
Dept. of Mathematics
Inst. for Electronic Systems
University of Aalborg
Fredrik Bajers Vej 7E

DK-9220 Aalborg

Prof.Dr. Nicolas Hengartner
Department of Statistics
Yale University
P.O.Box 208290

New Haven , CT 06520-8290
USA

Prof.Dr. Al Hero
Elc. Eng. & Comp. Science Dept.
The University of Michigan

Ann Arbor , MI 48109-2122
USA

Dr. David Hunter
Department of Statistics
The University of Michigan
1447 Mason Hall

Ann Arbor , MI 48109-1027
USA

Dr. Geurt Jongbloed
Faculteit Wiskunde en Informatica
Vrije Universiteit Amsterdam
De Boelelaan 1081

NL-1081 HV Amsterdam

Prof.Dr. Estate Khmaladze
Dept. of Mathematics
The University of New South Wales

Sydney NSW 2052
AUSTRALIA

Dr. Rik Lopuhaa
Faculty of Mathematics and
Informatics
Delft University of Technology
P. O. Box 356

NL-2628 BL Delft

Prof.Dr. Bernard A. Mair
Dept. of Mathematics
University of Florida
358 Little Hall

Gainesville , FL 32611-8105
USA

Dr. Axel Munk
Fakultät für Mathematik
Ruhr-Universität Bochum

44780 Bochum

Dr. Werner Nagel
Institut für Stochastik
Mathematische Fakultät
Universität Jena

07740 Jena

Prof.Dr. Georg Neuhaus
Institut für Mathematische
Stochastik
Universität Hamburg
Bundesstr. 55

20146 Hamburg

Prof.Dr. Robert Nowak
Dept. of Electrical and Computer
Engineering Building
Michigan State University
2120 Engineering Building

East Lansing , MI 48824-1226
USA

Dr. Michael Nussbaum
Weierstraß-Institut für
Angewandte Analysis und Stochastik
im Forschungsverbund Berlin e.V.
Mohrenstr. 39

10117 Berlin

Prof.Dr. Doug W. Nychka
Geophysical Statistic Project
National Center for Atomsph. Res.
Climate & Global Dynamics Division

Boulder , CO 80302
USA

Prof.Dr. Joachim Ohser
Institut für Techno- und
Wirtschaftsmathematik e.V. (ITWM)

67653 Kaiserslautern

Prof.Dr. Sergei Pereverzev
Institute of Mathematics of the
National Academy of Sciences of the
Ukraine
Tereshchenkivskastr. 3

252601 Kiev 4
UKRAINE

Prof.Dr. Rolf-Dieter Reiß
Fachbereich 6 Mathematik
Universität Siegen

57068 Siegen

Prof.Dr. Frits Ruymgaart
Dept. of Mathematics
Texas Tech. University
Box 41042

Lubbock , TX 79409
USA

Prof.Dr. Winfried Stute
Mathematisches Institut
Universität Gießen
Arndtstr. 2

35392 Gießen

Prof.Dr. Alexander B. Tsybakov
Laboratoire de Probabilites
Universite Paris 6
tour 56
4 place Jussieu

F-75252 Paris Cedex 05

Aad W. van der Vaart
Mathematisch Instituut
Vrije Universiteit
De Boelelaan 1081 a

NL-1081 HV Amsterdam

Prof.Dr. Yehuda Y. Vardi
Dept. of Statistics
Rutgers University
Hill Center, Busch Campus

New Brunswick , NJ 08903
USA

Prof.Dr. Wolfgang Weil
Mathematisches Institut II
Universität Karlsruhe

76128 Karlsruhe

Prof.Dr. Gerhard Winkler
GSF-Biomathematik und Biometrie
Postfach 1129

85758 Oberschleissheim

Prof.Dr. Cun-Hui Zhang
Department of Statistics
Rutgers University
110 Frelinghuysen rd.

Piscataway , NJ 08854

Dr. Jacques Zuber
Departement de Mathematiques
Ecole Polytechnique Federale
de Lausanne

CH-1015 Lausanne

Dr. Erik van Zwet
Department of Mathematics
University of Western Australia

Nedlands , WA 6907
AUSTRALIA

Prof.Dr. Willem R. van Zwet
Mathematisch Instituut
Rijksuniversiteit Leiden
Postbus 9512

NL-2300 RA Leiden