

PREFACE

NEW TRENDS IN STATISTICAL INFORMATION PROCESSING

This special issue of AISM is based on the symposium “New Trends in Optimization and Computational Algorithms (NTOC2001)” organized by the Institute of Statistical Mathematics (ISM) and held at Kyodai Kaikan, Kyoto, Japan, from December 9 to 13, 2001. The symposium focused on computational algorithms and methodologies which play crucial roles in discovering knowledge from massive and complex data sets in the 21st century. The symposium consisted of three sections: (i) Optimization Algorithms and Related Topics; (ii) Algorithms for Statistical Learning; (iii) From Optimization to High Dimensional Sampling and Integration. This special issue contains five papers contributed by those who gave talks or presented posters in NTOC 2001. They clearly represent new trends of statistics and information science in the 21st century and indicate some prominent directions for future research. Below we briefly introduce each paper.

The papers by Bousquet and by Weston, Schölkopf, Eskin, Leslie and Noble deal with theoretical and algorithmic aspects of statistical and machine learning.

Bousquet provides a new way to give a stochastic upperbound of the generalization error. Unlike the traditional PAC approach based on the VC dimension, he utilizes the Rademacher complexity and derives tighter bounds by means of the concentration inequality.

Weston *et al.* propose a new transformation method of kernel functions which are used in support vector machines and other kernel methods. The main contribution of this transformation is the reduction of extremely large diagonal elements in the kernel matrix, which are often harmful for classification and regression.

The papers by Doucet and Tadić and by Guo, Wang and Chen treat sequential Monte Carlo methods for complex and time-varying statistical models.

An important subject in the application of the sequential Monte Carlo method is how to incorporate the parameter estimation procedure. Doucet and Tadić develop novel algorithms for parameter estimation in state space models (or hyperparameter estimation in hierarchical Bayesian models). They discuss both on-line and off-line cases.

Efficient transmission of digital messages through complex and time-varying channels is a hot topic in statistical information processing. Guo *et al.* develop a method based on Bayesian model averaging and sequential Monte Carlo for adaptive detection in a fading channel. They model fading channels by wavelet expansions and calculate the averages over the models with a different number of coefficients. Their method seems suitable for time-varying situations.

The last paper by Tanaka deals with the analysis of massive data. Specifically, she has studied tick data of the Yen-Dollar exchange rate for 6 years consisting of ten million data points, and observes that the structure of the conditional probability of the increasing/decreasing of the rate is very stable throughout the examined term. The paper is a good example of the statistical analysis of tick data.

In addition to the five papers introduced above, eight more papers will be simultaneously published as the outcome of the symposium in a special issue of *Mathematical Programming Series B* edited by Hiroshi Konno and Takashi Tsuchiya. They deal with optimization without derivatives, independent component analysis, robust optimization, combinatorial optimization, portfolio optimization, infinite-dimensional opti-

mization, and interior-point methods.

Finally, we would like to express our cordial thanks to all the invited speakers, poster session contributors and participants of NTOC 2001 from various fields of the information/statistical/mathematical sciences who got together at Kyoto in December 2001 and brought this special issue to fruition.

Guest Editors

Kenji Fukumizu

Yukito Iba

Takashi Tsuchiya

Koji Tsuda