Statistisches Kolloquium

Zeit: Dienstag, 04.06.2013, 16.15 - 17.00

Ort: HKW 4 (sogen. „Toaster“), Raum 503

Vortragender: **Ioannis Ntzoufras** (Joint work with D.Fouskakis and D.Draper)
Department of Statistics
Athens University of Economics and Business

Titel: Power-Expected-Posterior Priors for Variable Selection in Gaussian Linear Models

Abstract

In the context of the expected-posterior prior (EPP) approach to Bayesian variable selection in linear models, we combine ideas from power-prior and unit-information-prior methodologies to simultaneously (a) produce a minimally-informative prior and (b) diminish the effect of training samples. The result is that in practice our power-expected-posterior (PEP) methodology is sufficiently insensitive to the size n* of the training sample, due to PEP's unit-information construction, that one may take n* equal to the full - data sample size n and dispense with training samples altogether. This promotes stability of the resulting Bayes factors, removes the arbitrariness arising from individual training-sample selections, and greatly increases computational speed, allowing many more models to be compared within a fixed CPU budget. In this we focus on Gaussian linear models and develop our PEP method under two different baseline prior choices: the independence Jeffreys (or reference) prior, yielding the J-PEP posterior, and the Zellner g-prior, leading to Z-PEP. The first is the usual choice in the literature related to our work, since it results in an objective model-selection technique, while the second simplifies and accelerates computations due to its conjugate structure (this also provides significant computational acceleration with the Jeffreys prior, because the J-PEP posterior is a special case of the Z-PEP posterior). We find that, under the reference baseline prior, the asymptotics of PEP Bayes factors are equivalent to those of Schwartz's BIC criterion, ensuring consistency of the PEP approach to model selection. We compare the performance of our method, in simulation studies and a real example involving prediction of air-pollutant concentrations from meteorological covariates, with that of a variety of previously-defined variants on Bayes factors for objective variable selection. Our PEP prior, due to its unit-information structure, leads to a variable-selection procedure that (1) is systematically more parsimonious than the basic EPP with minimal training sample, while sacrificing no desirable performance characteristics to achieve this parsimony; (2) is robust to the size of the training sample, thus enjoying the advantages described above arising from the avoidance of training samples altogether; and (3) identifies maximum-a-posteriori models that achieve good out-of-sample predictive performance. Moreover, PEP priors are diffuse even when n is not much larger than the number of covariates p, a setting in which EPPs can be far more informative than intended.

Für weitere Informationen wenden Sie sich bitte an:
Maria Kateri (Tel.: 97950, maria.kateri@rwth-aachen.de)