POLSKA AKADEMIA NAUK
Instytut Badań Systemowych

METHODS OF ESTIMATION OF RELATIONS OF: EQUIVALENCE, TOLERANCE AND PREFERENCE IN A FINITE SET

Leszek Klukowski

Warsaw 2011

SYSTEMS RESEARCH INSTITUTE
POLISH ACADEMY OF SCIENCES

Series: SYSTEMS RESEARCH
Volume 69

===

Series Editor:
Prof. dr hab. inż. Jakub Gutenbaum

Warsaw 2011

SYSTEMS RESEARCH INSTITUTE
POLISH ACADEMY OF SCIENCES

==

Leszek Klukowski

METHODS OF ESTIMATION OF RELATIONS OF: EQUIVALENCE TOLERANCE AND PREFERENCE IN A FINITE SET

Warsaw 2011

dr Leszek Klukowski
Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
email: Leszek.Klukowski@ibspan.waw.pl

Papers reviewers:

Prof. dr hab. inż. Ignacy Kaliszewski
Prof. dr hab. Tadeusz Trzaskalik

The work has been supported by the grant No N N111434937
of the Polish Ministry of Science and Higher Education

Printed in Polands
Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
www.ibspan.waw.pl

ISSN 0208-8029
ISBN 9788389475374

Main notations

$\mathbf{X} = \{x_1, ..., x_m\}$ $(3 \le m < \infty)$ – set of elements

$\mathbf{R}^{(e)}$ - equivalence relation

$\mathbf{R}^{(\tau)}$ – tolerance relation

$\mathbf{R}^{(p)}$ – preference relation

$\chi_1^{(\ell)*}, ..., \chi_n^{(\ell)*}$ $(\ell \in \{p, e, \tau\}; n \ge 2)$ - relation form (family of subsets)

$T_\upsilon^{(\ell)}(x_i, x_j)$ - the function expressing relation $\chi_1^{(\ell)*}, ..., \chi_n^{(\ell)*}$

$\widetilde{T}_\upsilon^{(\ell)}(x_i, x_j)$ - the function expressing relation $\widetilde{\chi}_1^{(\ell)}, ..., \widetilde{\chi}_n^{(\ell)}$

$g_{\upsilon k}^{(\ell)}(x_i, x_j)$ $(\ell \in \{e, \tau, p\}$; $\upsilon \in \{b, \mu\}$; $k = 1, ..., N)$ - pairwise comparisons

b – binary comparisons

μ - multivalent comparisons

$g_\upsilon^{(\ell, me)}(x_i, x_j)$ - median from the comparisons of a pair (x_i, x_j)

$F_X^{(\ell)}$ - the feasible set of optimization problem

$\hat{\chi}_1^{(\ell)}, ..., \hat{\chi}_{\hat{n}}^{(\ell)}$ - the estimator of the relation based on the sum of differences between the comparisons and the relation form

$\widehat{\chi}_1^{(\ell)}, ..., \widehat{\chi}_{\widehat{n}}^{(\ell)}$ - the estimator of the relation based on the sum of differences between medians from comparisons for each pair and the relation form

$\hat{T}_\upsilon^{(\ell)}(x_i, x_j)$ - the function expressing an estimate $\hat{\chi}_1^{(\ell)}, ..., \hat{\chi}_{\hat{n}}^{(\ell)}$

$\widehat{T}_\upsilon^{(\ell)}(x_i, x_j)$ - the function expressing an estimate $\widehat{\chi}_1^{(\ell)}, ..., \widehat{\chi}_{\widehat{n}}^{(\ell)}$

$\hat{\Delta}_\upsilon^{(\ell)} = \sum_{<i,j> \in R_m} \left| \hat{T}_\upsilon^{(\ell)}(x_i, x_j) - T_\upsilon^{(\ell)}(x_i, x_j) \right|$ - one-dimensional estimation error (estimator based on the sums of differences)

$\widehat{\Delta}_\upsilon^{(\ell)} = \sum_{<i,j> \in R_m} \left| \widehat{T}_\upsilon^{(\ell)}(x_i, x_j) - T_\upsilon^{(\ell)}(x_i, x_j) \right|$ - one-dimensional estimation error (estimator based on medians from comparisons)

$W_{\upsilon N}^{(\ell)*}$ - random variable expressing differences between comparisons

$g_{\upsilon k}^{(\ell)}(x_i, x_j)$ and values $T_{\upsilon}^{(\ell)}(x_i, x_j)$

$\widetilde{W}_{\upsilon N}^{(\ell)}$ - random variable expressing differences between comparisons

$g_{\upsilon k}^{(\ell)}(x_i, x_j)$ and values $\widetilde{T}_{\upsilon}^{(\ell)}(x_i, x_j)$

List of tables

List of graphs

The book presents the estimators of three relations: equivalence, tolerance, and preference in a finite set of data items, based on multiple pairwise comparisons, assumed to be disturbed by random errors. The estimators were developed by the author. They can refer to binary (qualitative), multivalent (quantitative) and combined comparisons. The estimates are obtained on the basis of solutions to the discrete programming problems. The estimators have been developed under weak assumptions on the distributions of comparison errors; in particular, these distributions can have non-zero expected values. The estimators have good statistical properties, including, especially importantly, consistency. Therefore, they produce good results in cases when other methods generate incorrect estimates. The precision of the estimators has been established with the use of simulation methods. The estimates can be validated in a versatile way. The whole estimation process, i.e. comparisons, estimation and validation can be computerized. The approach allows also for inference about the relation type – equivalence or tolerance, on the basis of binary data. Thus, it has features of data mining methods.

The estimators have been applied for ranking and grouping of data from some empirical sets. In particular, estimation of the tolerance relation (overlapping classification) was applied for determination of homogenous shapes of functions expressing profitability of treasury securities and was used for forecasting purposes.

ISSN 0208-8029
ISBN 9788389475374

SYSTEMS RESEARCH INSTITUTE
POLISH ACADEMY OF SCIENCES
Phone: (+48) 22 3810246 / 22 3810277 / 22 3810241 / 22 3810273
email: biblioteka@ibspan.waw.pl