# The Application of the Entrop Procedure

Uwe Blien[1], Friedrich Graef[2]

[1] Institute for Employment Research,
Regensburger Str. 104, D-90327 Nürnberg, Germany

[2] Institute of Applied Mathematics, University of Erlangen-Nuremberg,
Martensstr. 3, D-91058 Erlangen, Germany

**Abstract:** A procedure for the problem of recovering tabular data in case of incomplete or inconsistent information is presented. It generalizes the well kown RAS (or IPF) algorithm by allowing a wider class of constraints concerning the table entries such as equalities and inequalities over arbitrary cross sections. The theoretical background of the procedure is outlined and some examples of applications are reported.

# 1 Introduction

A frequently encountered problem in econometrics, empirical economics, and statistics is that of recovering matrices or multidimensional tables from incomplete information.

To problems of this type usually the RAS procedure (known from input-output analysis) is applied, which is identical to the Iterative Proportional Fitting Algorithm (IPF) used with log-linear models in statistics. With the RAS method a matrix is computed which is structurally similar to a given prior matrix and whose column and row sums attain prescribed values.

The ENTROP procedure presented here is a generalization of the RAS method. With ENTROP not only values for column and row sums can be prescribed. It may be applied whenever the information about the structure of the matrix (in the following also called *table*) to be estimated is given by a set of linear equations and/or inequalities with respect to the table entries. In addition, an a priori known reference table can be specified. The structure of the table resulting from this computational procedure will be as close as possible to that of the reference table.

ENTROP is a procedure minimizing the relative entropy. The result can be interpreted according to criteria from statistics and formal information theory (Shannon, Weaver (1949), Kullback (1968)). Under certain conditions the table estimated by this method may be interpreted as the most likely one.

Originally the ENTROP procedure had been developed as a tool for the Educational Accounting System of the Institute for Employment Research in Germany. Its purpose is the estimation of large matrices (with approximately 30000 elements) of transition rates between the labour market and the educational system.

Beyond this specific application it may serve many purposes. Typical tasks within the range of applicability of ENTROP are the *disaggregation of data* available only in a summarized form, *the weighting of random samples*, i.e. their adjustment to distributions known from official statistics, the *estimation of tables* from heterogeneous, incompatible and incomplete data, the *computation of transition matrices* of Markov processes, and the *construction of forecasts* (e.g. of the joint distribution of some variables from estimates of their univariate distributions)

# 2 Computing Tables

The computation of two-dimensional input-output tables may be considered as a representative example for the application of the method. Table 1 illustrates a matrix $X^t$ to be estimated.

Table 1: input-output matrix $X^t$

| | inputs to industries $1 - J$ | | | | $\Sigma$ |
|---|---|---|---|---|---|
| outputs of ind. $1 - I$ | $x_{11}$ | $x_{12}$ | $\ldots$ | $x_{1J}$ | $b_1^r$ |
| | $x_{21}$ | $x_{22}$ | $\ldots$ | $x_{2J}$ | $b_2^r$ |
| | $x_{31}$ | $x_{32}$ | $\ldots$ | $x_{3J}$ | $b_3^r$ |
| | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ | $\vdots$ |
| | $x_{I1}$ | $x_{I2}$ | $\ldots$ | $x_{IJ}$ | $b_I^r$ |
| $\Sigma$ | $b_1^c$ | $b_2^c$ | $\ldots$ | $b_J^c$ | $b^{rc}$ |

In input-output analysis the rows of the table are the outputs of industries. The columns are the inputs. The entry at row $i$ and column $j$ is that part of the total output of industry $i$ which is used as input to industry $j$ during a given time period $t$. In most cases the number of rows $I$ equals the number of columns $J$. Here, for generality it is assumed that $I$ is not necessarily equal to $J$.

The unknown values of the elements $x_{ij}$ of $X^t$ have to be estimated from heterogeneous information. If, e.g., the concrete numbers of the row and column totals are given, then the following equations have to be observed:

$$\textstyle\sum_{j=1}^{J} x_{ij} = b_i^r \quad \text{for all } i, \quad \text{and} \quad \sum_{i=1}^{I} x_{ij} = b_j^c \quad \text{for all } j \tag{1}$$

Frequently there is some information about the internal structure of $X^t$ which might be drawn from different surveys and other sources. It can be included into the estimation process if is possible to define appropriate linear equations and inequalities. If—as a simple example—an estimate of a single entry $x_{ij}$ stemming from a representative survey is given and it has to be expected that the value is affected by certain random sample and survey errors an inequality of the form

$$b^l \leq x_{ij} \leq b^u \tag{2}$$

may be adequate where $b^l$ and $b^u$ are the assumed lower and upper limits of the real value corresponding to an error estimate.

By an appropriate choice of these limits information with various degrees of reliability can be handled. The fact that a single $x_{ij}$ is known only from relatively vague judgements of experts can be treated by entering a relatively large difference between $b^l$ and $b^u$ in (2) or by introducing a lower or upper bound only. By contrast, if the given data can be regarded as "hard", for example because they are taken from an overall survey of the population, then by setting $b^l = b^u$ the inequality becomes an equation.

Specifying inequalities to include information affected by sample errors is important, because contradictions between different data sources may lead to intolerable consequences for the whole system which requires the consistency of all accounts.

If only aggregated information on $X^t$ is available, then sums have to be taken into account. The method developed for the estimation of tables works if the information on the table can be given the form of $K$ linear inequalities:

$$\sum_{i=1}^{I} \sum_{j=1}^{J} a_{kij} x_{ij} \leq b_k \tag{3}$$

for $k = 1 \ldots K$. A system (3) includes inequalities like (2) and—with the $a_{kij}$ being either equal to 0 or to 1—equations like (1). Thus, the row and colummn sums of the tables can also be specified as conditions which must be fulfilled by the estimates.

Under special conditions the system (3) could be solved exactly. In many cases there are not enough inequalities to obtain an exact solution. Thus an estimation procedure has to be applied.

In some cases an estimation can be based on additional information on the structure of the matrix. This is possible if a matrix $U^{t-1}$ for one year is already available, whereas only incomplete and heterogeneous information is given for the adjacent years. If it can be assumed that transition behaviour varies only relatively slightly from year to year, the unknown matrix should be estimated in accordance with the structure of the known matrix.

With a given reference matrix $U^{t-1}$ the estimation problem can be solved if it is possible to redefine it as an optimization problem, e.g. to minimize a distance measure $D(X^t, U^{t-1})$ with the inequalities (3) representing the constraints of the optimization process. Proceeding this way will also be appropriate if the reference matrix $U^{t-1}$ itself is the result of an estimation (then $U^{t-1} = X^{t-1}$), but is based on far more information than is available for the matrix $X^t$.

# 3 Minimizing the Relative Entropy

The above description of the estimation problem requires the definition of a distance measure $D(X,U)$ between tables $X$ and $U$ (cf. Blien, Graef (1991, 1992)). After considering the properties of some distance functions, the relative entropy

$$E_u(X) = \sum_{m=1}^{M} x_m \ln\left(\frac{x_m}{u_m}\right) \tag{4}$$

was chosen. Since the estimation procedure does not rely on a specific ordering of the table entries it is assumed for the following that they are linearly numbered by $m = 1, \ldots, M$ (with $M = I \cdot J$ in case of a two-dimensional table).

Measuring the degree of similarity between two tables by (4) has a long tradition in natural sciences and engineering[1]. Minimizing the relative entropy can also be justified from a statistical viewpoint. If the reference matrix has been normalized such that $\sum_m u_m = 1$, then $u_m$ can be interpreted as the probability of a certain object to be in state $m$. If a total number $N = \sum_m x_m$ of objects is distributed independendly over the possible states according to this probability law and $x_m$ is the number of objects occupying state $m$ then the probability of the table $X = (x_m)$ follows a multinomial distribution:

$$P(X) = \frac{N!}{x_1!x_2!\ldots x_M!} u_1^{x_1} u_2^{x_2} \ldots u_M^{x_M}$$

The logarithm of $P(X)$ is closely related to the relative entropy $E_u(X)$. Using the Stirling-formula $\ln(s!) \approx s\ln(s) - s$ it can be shown by some algebra that

$$\ln(P(X)) \approx N\ln(N) - E_u(X)$$

The natural logarithm being an increasing function the table $X$ maximizing $\ln(P(X))$ also maximizes $P(X)$. Thus, assuming a stochastically independent placement of the $N$ objects under the specified constraints (5) the table

[1] A particular impressive example is computer tomography in medicine, as the relationship to the problem under discussion is especially clear. In computer tomography x-ray pictures of an organ are taken from various angles. From these photographs a sectional picture of the corresponding organ is made. The resulting picture can be imagined as an extremely finely gridded table which contains the grey values for the individual dots. The x-ray photographs correspond to the row and column sums of the table. The picture is generated with an entropy optimizing algorithm. It shows the organ in question with all details, including any possible pathological changes. Other cases of application for entropy optimizing are pattern recognition in research on artificial intelligence. In particular new approaches usually summarized under the title *neuronal networks* make use of these methods (cf. eg. Hinton, Sejnovski (1987), Kosko (1992)).

minimizing the relative entropy is asymptotically equal to the table representing the distribution of objects with the highest probability[2].

For arbitrary $u_m \geq 0$ the optimization problem for table estimation may now be stated as follows:

*Minimize the function $E_u(X)$ subject to the constraints*

$$\textstyle\sum_{m=1}^{M} a_{km} x_m \leq b_k \quad \text{for } k = 1, \ldots, K \tag{5}$$

*and the nonnegativity constraints*

$$x_m \geq 0 \quad \text{for } m = 1, \ldots, M \tag{6}$$

The nonnegativity constraints are stated because all elements of $X$ should be equal to or greater than 0 (note that all $u_m \geq 0$ by definition). The general structure of the resulting table can be obtained via the Kuhn-Tucker-Karoush theorem (cf. for example Chiang (1974)). The conditions for the minimum are:

$$\begin{aligned}
\ln\left(\frac{x_m}{u_m}\right) + 1 + \sum_{k=1}^{K} \mu_k a_{km} &= 0 && (7)\\
\mu_k &\geq 0 && (8)\\
\mu_k\Big(\sum_m a_{km} x_m - b_k\Big) &= 0 && (9)
\end{aligned}$$

The $\mu_k$ are the dual variables. If we solve (7) for the $x_m$ then

$$x_m = u_m e^{-1} e^{-\sum_k \mu_k a_{km}} \tag{10}$$

From (10) some properties[3] of the estimated table can be seen:

[2] By some algebra it can be shown (cf. Blien, Graef (1991)) that the relative entropy is proportional to another distance measure, to Kullback's *information gain* (cf. Kullback (1968), see also Ireland, Kullback (1968)), which is identical to (4), but defined for probability distributions, i.e. matrices with component sum equal to 1. The information gain is common in applications of the formal information theory (originally founded by Shannon, cf. Shannon, Weaver (1949)). Another term for information gain is *minimum discriminant information.* Haberman (1984) discusses the case of continuous data. Golan, Judge et al. (1996) use information gain minimization for parameter estimation in underdetermined linear models.

[3] There is a growing literature about the properties of estimates obtained by this minimum information principle (see Wauschkuhn (1982), Kullback (1968), Snickars, Weibull (1977), Batten (1983), Batten, Boyce (1986)). One property should be noted: an estimation via this principle is approximately equivalent to an optimization of a weighted sum of squares, a modified Chi-square statistic:

$$\chi_Q^2 = \sum_m \frac{(p_m - q_m)^2}{q_m}$$

where the $p_m$ are the normalized $x_m$'s and the $q_m$ the normalized $u_m$'s. This relationship was proved by Kadas and Klafszky (1976, p. 442).

- nonnegativity: all $x_m \geq 0$

- conservation of zeroes: for each $u_m = 0$ the corresponding $x_m$ is also equal to zero.

## 4 The Entrop algorithm

To estimate a table our task is to minimize the relative entropy (4) under the constraints (5) and (6). There exists a variety of methods known from the theory of nonlinear optimization (cf. Gill, Murray, Wright (1981)) to achieve this, as the relativ entropy has the property of global convexity. To solve the problem some authors have already proposed special applications of the Newton-Raphson algorithm (Wauschkuhn (1982)), geometric programming (Kadas, Klafsky (1976)), stochastic optimization (Ablay (1987)) and simulated annealing (Paass (1988)).

The algorithm finally chosen was designed for the application in the Educational Accounting System (BGR) of the Institute for Employment Research. As a primary goal the computer program implementing this algorithm should handle large scale problems on computers with limited memory. In addition it should be a robust procedure with respect to inconsistent constraints.

The Entrop procedure developed under these premisses is an iterative algorithm of the row action type, i.e. only one out of the total set of constraints is used at each iteration step. It is based on so called entropy projections. These projections are nonlinear equations originally introduced by Censor (1982)[4] to maximize absolute entropy[5] It can be shown that they may also be applied to minimize the relative entropy with respect to arbitrary reference tables.

The algorithm is an iterative procedure. Its starting values are:

$$\begin{array}{ll} x_m = u_m & \text{for all } m \text{, and} \\ \mu_k = 0 & \text{for all } k. \end{array} \tag{11}$$

Every step in the iteration process includes the following operations:

1. Computation of the entropy projection on restriction $k$: compute a $\delta$ so that
$$\sum_m a_{km} x_m e^{\delta a_{km}} = b_k \tag{12}$$

2. Correction of the sign: If $\delta > \mu_k$, set $\delta = \mu_k$.

[4]See also Censor and Lent (1981) and Censor (1981). Censor's method is based on a general principle found by Bregman (1967).

[5]The (absolute) entropy of a table is $-\sum_m x_m \ln(x_m)$. It coincides with the negative relative entropy in the case of $u_m = 1$ for all $m$.

3. Updating the values for $x_m$ and $\mu_k$:

$$\begin{aligned} x_m &= x_m e^{\delta a_{km}} \quad \text{for all } m \\ \mu_k &= \mu_k - \delta \quad \text{for all } k \end{aligned} \tag{13}$$

By means of the criteria given by Censor and Lent (1981) it can be shown that the iteration process converges against the solution of (10), (8), and (9), if a solution exists. The procedure is rather simple.

If no reference matrix is present the ENTROP method can also be used. In this case calculations are made with all $u_m$ in the reference table set equal to 1. The method then attempts to occupy the resulting matrix as evenly as possible. Whether this is a reasonable proceeding depends on the specific problem. In this case the optimization of the relative and the absolute entropy are identical (see footnote 4).

# 5 Generalization of the RAS method

The ENTROP method contains as a special case the iterative proportional fitting algorithm (IPF) already familiar from loglinear models of statistics (Bishop, Fienberg et al. (1975)). The RAS method (cf. Stone (1962), Bacharach (1970)) used in input-output analysis (also known as the Deming-Stephan algorithm after Deming, Stephan (1940), cf. Bachem, Korte (1979)), is also included, as it is identical with the IPF. Gorman (1963) proved that the RAS method is also a procedure for entropy optimizing.

For the special case that the constraints consist of column and row sum prescriptions only, it can be seen easily that the steps carried out in the RAS method are identical with the computation of entropy projections in ENTROP.

If the constraints are sums

$$\sum_{m \in \mathcal{M}_k} x_m = b_k$$

over cross sections $\mathcal{M}_k$ of table entries the equations (12) can be solved explicitly for $\delta$. Substitution of $\delta$ in (13) then results in the recursion

$$x_m^{t+1} = \frac{x_m^t b_k}{\sum_{r \in \mathcal{M}_k} x_r^t} \tag{14}$$

where $t$ is the respective iteration step. Returning to double-index notation $ij$ for the row and column constraints (1) of input-output analysis the iteration formula (14) reads as

$$x_{ij}^{t+1} = \frac{x_{ij}^t b_j^c}{\sum_r x_{rj}^t} \tag{15}$$

for column constraints and

$$x_{ij}^{t+1} = \frac{x_{ij}^t b_i^r}{\sum_s x_{is}^t} \tag{16}$$

for row constraints which is precisely one of the iteration steps of the RAS or the IPF method.

# 6 Examples of Applications of the Entrop procedure

The Entrop procedure was developed for a task very similar to our example from input-output analysis. It was designed for the *Educational Accounting System* (*Bildungsgesamtrechnung*: BGR) of the Institute for Employment Research (in Germany). To fix ideas a simple application with artificial data will be presented first before giving an overview of the type and size of estimation problems within the framework of the BGR.

## 6.1 A simple example

Assume table 2 to be the input-output matrix of a fictitious economy at time period $t-1$. For the next period $t$ as a first approach assume only the sums of the column entries (the total input to the respective industries) and the row sums (the total output) to be known.

Table 2: Input-output table for time period $t-1$ and totals for period $t$

| | | Inputs to industries 1–5 | | | | | totals | totals at period $t$ |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | | |
| | 1 | 9 | 71 | 54 | 66 | 11 | 211 | 226 |
| outputs | 2 | 20 | 189 | 60 | 53 | 17 | 339 | 372 |
| of | 3 | 31 | 159 | 21 | 25 | 9 | 245 | 333 |
| indu- | 4 | 15 | 56 | 0 | 11 | 3 | 85 | 142 |
| stries | 5 | 1 | 3 | 5 | 5 | 1 | 15 | 50 |
| | 6 | 2 | 10 | 51 | 0 | 5 | 68 | 153 |
| totals | | 78 | 488 | 191 | 160 | 46 | 963 | |
| totals at $t$ | | 119 | 638 | 252 | 225 | 42 | | 1276 |

In estimating the entries of the table for time period $t$ it often is plausible to assume that the basic structure of the relationships between the different industries is nearly constant over time. Thus, the structure of the estimated table should be as similar as possible to that of the a priori matrix of table

2 while at the same time the row entries and column entries should sum up to the quantities known for period $t$.

It is well known that this can be done with the RAS algorithm. An application of the ENTROP method yields the same result, shown in table 3.

Table 3: Input-output table for time period $t$, estimated with the ENTROP method on the basis of table 2 and of the column and row totals for period $t$ (the result equals a RAS estimation)

| | | Inputs to industries 1–5 | | | | | totals |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | |
| | 1 | 10.9 | 78.2 | 48.2 | 80.9 | 7.8 | 226 |
| outputs | 2 | 24.8 | 213.4 | 54.9 | 66.6 | 12.3 | 372 |
| of | 3 | 46.5 | 217.3 | 23.3 | 38.0 | 7.9 | 333 |
| indu- | 4 | 27.0 | 91.8 | 0.0 | 20.1 | 3.2 | 142 |
| stries | 5 | 3.8 | 10.4 | 14.1 | 19.4 | 2.2 | 50 |
| | 6 | 5.9 | 27.0 | 111.5 | 0.0 | 8.6 | 153 |
| totals | | 119 | 638 | 252 | 225 | 42 | 1276 |

However, the RAS method only permits an estimation using prescribed row and column sums and an a priori matrix. The ENTROP algorithm allows the specification of general linear equations and inequalities serving as restrictions to the estimation. As an example let it be assumed that there exists additional information about the input-output relations for time $t$. It might be known from a survey that the output of industries 2 and 3 to industries 3 and 4 is at least 250. The output of industry 3 to industry 2 might be less than two times the output of industry 4 to industry 2. This can be stated by two additional inequalities:

$$x_{23} + x_{24} + x_{33} + x_{34} \geq 250 \tag{17}$$
$$x_{32} - 2x_{42} \leq 0 \tag{18}$$

Table 4 shows the result of an application of the ENTROP procedure, which includes table 2 as a reference matrix, the column and the row sums for time period $t$ and the two additional inequalities (17) and (18). The entries affected by (17) and (18) have been emphasized for easy comparison with table 3. Both inequalities are met. The sums of the row entries and column entries are the same as in tables 3 and the structure of table 4 is as similar as possible to that of table 2.

## 6.2 Application to the Educational Accounting System (BGR)

Demographic accounting approaches provide a means of relating the population in different states at the beginning and the end of a period. Such

Table 4: Input-output table for time period $t$, estimated with the ENTROP method on the basis of table 2, the column and row totals of period $t$ and additional information concerning the structure of the table

| | | Inputs to industries 1–5 | | | | | totals |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | |
| | 1 | 14.5 | 106.4 | 37.1 | 58.7 | 9.3 | 226 |
| outputs | 2 | 20.5 | 180.1 | **75.7** | **86.6** | 9.2 | 372 |
| of | 3 | 41.4 | **197.5** | **34.5** | **53.2** | 6.3 | 333 |
| indu- | 4 | 28.5 | **98.9** | 0.0 | 11.5 | 3.0 | 142 |
| stries | 5 | 5.4 | 15.2 | 11.6 | 15.0 | 2.9 | 50 |
| | 6 | 8.6 | 39.9 | 93.2 | 0.0 | 11.3 | 153 |
| totals | | 119 | 638 | 252 | 225 | 42 | 1276 |

approaches differ from cross-sectional analyses in that the gross flows of individuals into, within, and out of the states during that period are examined (cf. Stone (1981), Stone, Weale (1986), see also Land, Juster (1981b) and Land, McMillen (1981)).

A system of that kind, the *Educational Accounting System* (*Bildungsgesamtrechnung* - BGR) was developed by the Institute for Employment Research (IAB) (cf. Blien, Tessaring (1988) and (1992), Tessaring (1986) and (1987), Tessaring et al. (1991) and (1992)). In the BGR the German population is classified according to defined categories: pupils attending different kinds of general and vocational schools, apprentices, students, gainfully employed, unemployed and economically non-active individuals.

Thus, the BGR forms the basis for improved analyses as well as for forecasts on the relationship between education and the labour market. The long-term development of the labour supply in various segments of the labour market can be examined.

The relation between the stocks of adjacent years is given by the inflow-outflow matrix $X^t$. In $X^t$ the individuals are classified by their opening states in the rows and by their closing states in the columns. An element $x_{ij}$ of this matrix shows the number of people, who change from state $i$ to state $j$. The so called *stayers*, i.e. $x_{ij}$ with $i = j$ are included as well. The system is consistent, i.e.:

$$\sum_{j=1}^{n} x_{ij} = b_i^t \quad \text{for all } i, \quad \text{and} \quad \sum_{i=1}^{n} x_{ij} = b_j^{t+1} \quad \text{for all } j \tag{19}$$

The $b_i^t$ are the row totals (i.e. the states at $t$), the $b_j^{t+1}$ the column totals (i.e. the states at $t+1$), $n$ is the number of states.

Most of the BGR stock data were taken from German official statistics. Available transition data, however, stem from surveys and are often incompatible or affected by sample errors. In order to use such heterogeneous and incompatible information to estimate the matrix of the flows of individuals in the BGR the ENTROP procedure was developed.

The matrices estimated by the ENTROP method are very large. Since they contain about 30 000 elements in four dimensions (year $t$, year $t+1$, sex, age) it is shown that the ENTROP procedure is not restricted to two dimensional tables. The calculations are based on a-priori matrices. Additionally, they use about 2000 inequality and equality restrictions.

The procedure originally was programmed in FORTRAN on a BS2000 mainframe. A program written in C is available as well for Personal Computers running DOS, Windows, or OS/2 and UNIX workstations under the Sun Solaris operating system. The memory management of ENTROP algorithm is so efficient, that PCs can be used for most problems.

Other applications of the ENTROP procedure are in preparation, e.g. an analysis of regional differences of unemployment in Eastern Germany and an analysis of voting mobility.

## References

BACHARACH, M. (1970): Biproportional Matrices and Input-Output-Change, Cambridge

BACHEM, A. and KORTE, B. (1979): On the RAS-Algorithm. In: *Computing Vol. 23, 189 ff*

BATEY, P.W.J. and MADDEN, M. (1986) (eds): Integrated analysis of regional systems. London: Pion

BATTEN, D.F. (1983): Spatial Analysis of Interacting Economics. Boston etc.: Kluwer-Nijhoff

BATTEN, D. and BOYCE, D. (1986): Spatial interaction, transportation, and interregional commodity flow models. In: Nijkamp (1986)

BISHOP, Y., FIENBERG, S. and HOLLAND, P. (1975): Discrete Multivariate Analysis: Theory and practice. Cambridge, London

BLIEN, U. and GRAEF, F. (1991): Entropieoptimierungsverfahren in der empirischen Wirtschaftsforschung. In: *Jahrbücher für Nationalökonomie und Statistik 208/4, 399–413*

BLIEN, U. and GRAEF, F. (1992): ENTROP: A General Purpose Entropy Optimizing Method for the Estimation of Tables, the Weighting of Samples, the Disaggregation of Data, and the Development of Forecasts. In: Faulbaum, F. (1992) (ed): *Softstat '91. Advances in Statistical Software 3.* Stuttgart: Gustav Fischer

BLIEN, U. and TESSARING, M. (1988): Die Bildungsgesamtrechnung des IAB. In: Mertens, D. (ed): *Konzepte der Arbeitsmarkt- und Berufsforschung , 3rd ed..* BeitrAB 70, Nürnberg

BLIEN, U. and TESSARING, M. (1992): Transitions between education and the labour market in Germany. The application of the ENTROP procedure in the

Educational Accounting System. *Unpublished paper prepared for presentation at the EALE Conference at the University of Warwick.*

BREGMAN, L. (1967): The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Comp. Math. and Math. Phys. Vol. 7/3, 200ff*

CENSOR, Y. (1981): Row-actions methods for huge and sparse systems and their applications. In: *SIAM Review, Vol. 23*

CENSOR, Y. and Lent, A. (1981): An iterative row-action method for interval convex programming. *Journal of Optimization Theory and Applications, Vol. 34, 321ff*

CHIANG, A. (1974): Fundamental methods of mathematical economics. (2nd ed.), Tokyo etc.: Mcgraw-Hill Kogakusha

DEMING, W.E. and STEPHAN, F.F. (1940): On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics 11/1, March, 427ff*

DRENICK, R. and KOZIN, F. (1982) (eds): System Modeling and Optimization. *Lecture Notes in Control and Information Science 38*, Heidelberg etc.: Springer

GOLAN, A., JUDGE, G., and MILLER, D. (1996): Maximum Entropy Econometrics: Robust Estimation With Limited Data. Chichester etc.: Wiley

GORMAN, W. (1963): Estimating Trends in Leontief Matrices. A Note on Mr. Bacharach's Paper. *Discussion paper, Nuffield College*, Oxford

HABERMAN, S.J. (1984): Adjustment by minimum discriminant information. *The Annals of Statistics, 12/3, 971ff*

HARRIGAN, F. and MCGREGOR, P.G. (1988) (eds): Recent advances in regional economic modelling. London: Pion

HINTON, G.E. and SEJNOWSKI, T.J. (1986): Learning and Relearning in Boltzmann Machines. In: Rumelhart, D.E., McClelland, J.L. et al. (1986) (eds): *Parallel distributed processing.* Vol. 1: Foundations , Cambridge (Mass.), London: MIT Press

IRELAND, C. and KULLBACK, S. (1968): Contingency tables with given marginals. *Biometrika 55, 179ff*

JUSTER, F.T. and LAND, K.C. (1981) (eds): Social Accounting Systems. Essays on the State of the Art. New York etc.: Academic Press

KADAS, S.A. and KLAFSKY, E. (1976): Estimation of the Parameters in the Gravity Model for Trip Distribution: a New Model and Solution Algorithm, *Regional Science and Urban Economics Vol. 6, 439ff*

KOSKO, B. (1992): Neural networks and fuzzy systems. Englewood Cliffs: Prentice Hall

KULLBACK, S. (1968): Information Theory and Statistics , 2nd ed., New York: Dover

LAND, K.C. and MCMILLEN, M.M. (1981): Demographic Accounts and the Study of Social Change, In: Juster, Land (1981)

LAND, K.J. and JUSTER, F.T. (1981): Social Accounting Systems: An Overview. In: Juster, Land (1981)

LEVINE, R.D. and TRIBUS, M. (1979) (eds): The maximum entropy formalism. Cambridge (Mass.), London: MIT-Press

NIJKAMP, P. (1986) (ed): Handbook of regional and urban economics, Vol. I: Regional economics. Amsterdam etc.: North-Holland

PAASS, G. (1988): Stochastic Generation of a Synthetic Sample from Marginal Information. *Gesellschaft für Mathematik und Datenverarbeitung, Arbeitspapiere der GMD Nr. 308*. St. Augustin

SHANNON, C.E. and WEAVER, W. (1949): The Mathematical Theory of Communication. Urbana: University of Illinois Press

SLATER, P.B. (1992): Equilibrium and nonequilibrium statistical thermodynamical approaches to modeling spatial interaction dynamics. *Environment and Planning A, vol. 24: 441–446*

SNICKARS, F. and WEIBULL, J. (1976): A minimum information principle. *Regional Science and Urban Economics Vol. 7, 137ff*

STONE, R.N. (1962): Multiple classifications in social accounting. *Bulletin de l'Institut International de Statistique, Vol. 39, No. 3, 215ff*

STONE, R.N. (1981): The Relationship of Demographic Accounts to National Income and Product Accounts. In: Juster, Land (1981)

STONE, R.N. and WEALE, M. (1986): Two populations and their economies. In: Batey, Madden (1986)

TESSARING, M. (1986): Educational accounting in the Federal Republic of Germany. In: Parkes, D.L., Sellin, B. and Tessaring, M.: *Education/ training and labour market policy.* Gravenhage: SVO

TESSARING, M. (1987): Demographic aspects of educational expansion and labour-force development in the Federal Republic of Germany. *European Journal of Population Vol 3: 327–358*

TESSARING, M., BLIEN, U., FISCHER, G., HOFMANN, I. and REINBERG, A. (1990): Bildung und Beschäftigung im Wandel. Die Bildungsgesamtrechnung des IAB. *BeitrAB 126.* Nürnberg

TESSARING, M., REINBERG, A., FISCHER, G. and SCHWEITZER, C. (1992): Übergänge zwischen Bildung, Ausbildung und Beschäftigung und die Entwicklung der Qualifikationsstruktur. *Forthcoming paper.* Nürnberg

THEIL, H. (1967): Economics and Information Theory. Amsterdam

WAUSCHKUHN, U. (1982): Anpassung von Stichproben und n-dimensionalen Tabellen an Randbedingungen. München, Wien: R. Oldenbourg

WILSON, A. (1970): Entropy in urban and regional modelling. London: Pion Ltd.