Budapest Bridges Benchmarking

András Farkas

Óbuda University, Faculty of Economics
1084 Budapest, Tavaszmező u. 17, Hungary
e-mail: farkas.andras@kgk.uni-obuda.hu

Abstract: This paper is concerned with the comparison of different scaling methods which are applied to a complex bridge evaluation problem. It is shown that both tangible and intangible data and satisfaction of multiple criteria are essential to the success of such projects. Some new inconsistency measures for the matrices emerging in the decision making process are also used. A detailed numerical analysis of the results is presented.

Keywords: analytic hierarchy process, least-squares method, bridge design

1 Introduction and problem description

Bench-mark means altitude (fix) point or level in the field of geology. In management sciences this term is defined as a standard or point of reference by which something can be measured or judged. The term c*ompetitive benchmarking* involves analyzing the performance and practices of best-in-class enterprises. Their performance becomes a benchmark to which a firm can compare its own performance and their practices are used to improve that firm's practices. The potential benefits of this process are substantial. Not only is a firm able to recognize and adopt world-class standards, but it learns how these standards can be met. Although these goals may be challenging, in order to be successful, the management should aim to achieve them.

Benchmarking also plays an important role in design, technology, construction and architecture. In this paper we will investigate the use of different scaling methods for these purposes. Both normative and descriptive type methods will be used for ranking the alternatives. Comparisons will be made among these techniques concerning the errors inherent in these subjective information based procedures. Three methods will be applied: the multi-criteria technique for systems' evaluation (MTSE; see [6],[7]), the eigenvector method which is incorporated in the famous analytic hierarchy process (AHP; see [12],[13] and an extremal approach called least-squares method (LS, see

[1],[10]). An extension of the LS method for improving its quality of assessment, in the form of a recursive LS algorithm will also be discussed briefly (LSR; see [9]).

Recently, a considerable attention has shifted to rank and prioritize bridges by bridge management professionals. Some remarkable theoretical developments and real-world applications have been appeared in the literature in this topic, most notably [3].

The following benchmarking study came from the author's teaching experience in the fields of engineering and economics. Groups of students examined the ranking and evaluation problem of bridges serving urban transportation by stretching over the river Danube at the capital city of Budapest. Four bridges were selected for the analysis. They are displayed and coded below:

C **= Chain bridge**

E **= Elisabeth bridge**

M **= Margaret bridge**

A **= Árpád bridge**

A fictitious bridge having the desired (ideal) characteristics has also been defined and denoted by ***R***. This object will be referred to a benchmark, that is a standard or a reference point by which these bridges will be measured or judged. As a result of conducted surveys, the use of the Delphi method and an extensive literature review (for finding the engineering characteristics), the following set of the evaluation criteria C_i has been established:

C_1 = ***Urban traffic accessibility***: Easy access from both sides of the city.

C_2 = ***Multifunctional capabilities***: Usability for different transportation means.

C_3 = ***Lack of Environmental Impact***: No ecological harm/air pollution/noise, etc.

C_4 = ***Structure***: Construction, bridge type and traffic safety.

C_5 = ***Aesthetics appeal***: Architectural attractiveness.

C_6 = ***Lighting performance***: The quality of lighting and illumination.

C_7 = ***Cold allowance***: Minimum temperature specified for the structure.

C_8 = ***Engineering characteristics***: An averaged dimensionless technical measure.

C_9 = ***Traffic flow capacity***: Maximum number of crossing vehicles per hour.

C_{10}= ***Maintenance***: Cost of inspection, potential major repairs and routine tasks.

From these enumerated items, it is apparent, that both tangible attributes (measurable characteristics) and intangible attributes (where subjective judgements are needed) comprised the set of criteria.

2 The use of the MTSE method for benchmarking

In this Section a multi-attribute utility model (MAUT) called Multi-criteria Technique for Systems' Evaluation (MTSE) for determining the preferences, and thus the priority ranking of the four bridges is applied [6].

A characteristic feature of most of these models is the postulation that the preference of an individual towards a choice object is related to its "distance" from his/her ideal object which is usually a hypothetical object (see e.g. [11]). The closer the object is to the ideal one, the greater the preference towards it. The distance is a compound measure which takes into account the location of each object on several attributes (criteria) which characterize the object. Given *n* alternatives for an object of similar type, each characterized by *m* attributes, the general form of the model can be described by the function [11]:

$$D_j = \sum_{i=1}^{m} w_i d_{ji} + \varepsilon_j, \quad j = 1,2..,n, \tag{1}$$

where D_j is the *overall distance* of alternative j from the ideal one, w_i is the weight of attribute i, d_{ji} is the *distance* of the jth alternative from the ideal point on attribute i and $\boldsymbol{\varepsilon}_j$ is an error term. It is favored, that d_{ji} satisfies the metric properties, for example, let it correspond to the squared Euclidean distance.

MTSE was designed to incorporate both tangible and intangible attributes. Its data matrix is partitioned into four blocks. Every criterion is then assigned to the block that represents the associated scale of measurement. The numbers may appear in forms of binary variables, ranks and quantitative data usually with different units of measurement. A weighting number can be assigned to each criterion to measure its relative importance. The preference order of the objects is determined by the particular ratings received from the respondents. Each alternative is compared to the reference object. The "best" alternative becomes the one which is closest to this "ideal" object.

The decision makers' preferences are expressed by the differences between the objects on a [0–100] point *interval* scale. Their priority ranking results directly from the order of magnitude of their *relative standings*. For the aggregation of the individual rankings into a composite ranking the minimum variance method is proposed (see [2]).

The numerical values of the evaluation process for the selected Budapest bridges and those of the reference bridge are presented in Table 1.

Table 1. Evaluation data for the bridge benchmarking problem

Attribute	Scale	***C***	***E***	***M***	***A***	***R***	Weight
C_1	Nominal [0 or 1]	0	0	1	1	1	0.10
C_2	Nominal [0 or 1]	0	1	0	1	1	0.10
C_3	Nominal [0 or 1]	0	0	0	0	1	0.10
C_4	Ordinal [1-5]	2.5	3.5	4	3.5	5	0.10
C_5	Ordinal [1-5]	5	4	3	2.5	5	0.10
C_6	Ordinal [1-5]	2	3	3.5	5	5	0.10
C_7	Interval [0-–45] [˚C]	–22	–30	–35	–40	–45	0.10
C_8	Ratio [real] [dim.less]	4.2	5.1	3.2	7.6	8.7	0.10
C_9	Ratio [real] [unit/hour]	1000	1450	1300	1600	1800	0.10
C_{10}	Ratio [real] [BiHUF/y]	0.8	0.6	0.5	0.4	0.2	0.10

Observe in Table 1 that criteria C_1, C_2 ,C_3 are measured on a nominal scale, C_4, C_5, C_6 on an ordinal scale, C_7 on an interval scale and C_8, C_9, C_{10} on a ratio scale. On the ordinal scale the frequency of preference is given for these attributes, rather than their ranks, since we have applied a ten-point scale [1, 1.5, 2, 2.5, ..., 5] for the evaluation. C_7 represents temperature in degrees Celsius with the lowest temperature being the most favourable. The second to last column of the data matrix contains the scores (ratings) of the reference object ***R***. Notice also that for each criterion the weight numbers are identical. Hence, they equal 0.10. This choice is explained by simplicity reasons. In such a way, an average satisfaction of the stakeholders can be achieved, as they usually have wholly different perspectives and interests.

Numerical computations were made by MAROM [5]. The formal description of MTSE with its mathematical background can be found in [6] and [7]. The priority scores p_j's of the relative standings s_j, $j=1,\ldots,n$, of the bridges are given by the composite vector:

$$\mathbf{p} = \begin{bmatrix} 100.00 \\ 79.43 \\ 55.76 \\ 53.63 \\ 25.98 \end{bmatrix} \begin{matrix} \rightarrow R \\ \rightarrow A \\ \rightarrow E \\ \rightarrow M \\ \rightarrow C \end{matrix} .$$

By definition, the reference bridge (benchmark) received 100%. If we understand the up-to-date requirements of a modern metropolis like Budapest is, then, the priority ranking of the bridges under study must not come as a surprise to us.

By forming pairwise ratios from the components p_j, of the vector **p**, so that s_i/s_j for i j, $i,j=1,2,...,n$, and $s_{ii}=1$ for $i=1,...,n$, a positive matrix $\mathbf{S}=[s_{ij}]$ and an element-wise positive vector $\mathbf{s}=[s_i]$, $i=1,\ldots,n$, can be constructed entries of which are the appropriate ratios of the priority scores. The entries of **s** are usually normalized so that the sum of the elements is unity.

$$\mathbf{S} = \begin{bmatrix} 1 & 0.4659 & 0.4844 & 0.3271 & 0.2598 \\ 2.1464 & 1 & 1.0397 & 0.7020 & 0.5576 \\ 2.0644 & 0.9618 & 1 & 0.6752 & 0.5363 \\ 3.0572 & 1.4245 & 1.4810 & 1 & 0.7943 \\ 3.8491 & 1.7934 & 1.8646 & 1.2590 & 1 \end{bmatrix},$$

$$\mathbf{s} = \begin{bmatrix} 0.3177 \\ 0.2523 \\ 0.1771 \\ 0.1704 \\ 0.0825 \end{bmatrix} \begin{matrix} \rightarrow R \\ \rightarrow A \\ \rightarrow E \\ \rightarrow M \\ \rightarrow C \end{matrix} .$$

Observe here that by forming ratios from the elements of **p**, this interpretation leads to the derivation of a *transitive* matrix **S** termed the matrix of the relative standings, since the cardinal consistency condition, $s_{ij} \cdot s_{jk} = s_{ik}$ holds for all $i,j,k=1,2,\ldots,n$. The vector **s** is thus the principal right eigenvector of matrix **S**. We note that the components of **s** have been reordered such that the priority ranking of the bridges to appear in a descending order. From here on, all through the article, this practice will be followed.

3 The use of the AHP method for benchmarking

In this Section, a descriptive method is used for benchmarking. Being a multiple criteria scaling method, the AHP represents a basic approach to decision making. It was founded by Saaty T.L. [12]. An overview of the AHP methodology and a concise discussion of the mathematical basics can be found in [8]. In this approach, each respondent should take $n(n-1)/2$ pairwise comparison judgments for the pairs of alternatives using a scale: [1/9,…,1/2,1,2,…,9]. Then, these ratio estimates become the elements of a positive $n\times n$ matrix $\mathbf{A}=[a_{ij}]$ called a *pairwise comparison matrix* (PCM). Here an entry a_{ij} from R^n represents a *ratio*, i.e., a_{ij} indicates the strength with which alternative A_i dominates alternative A_j with respect to a given criterion C_k. Matrix **A** is a *symmetrically reciprocal* (SR) matrix since its entries satisfy $a_{ij} \cdot a_{ji}=1$ for $i\neq j$, $i,j=1,2, ...,n$, and $a_{ii}=1$, $i=1,2, ...,n$.

The basic objective is to derive implicit *weights* (the priority scores), $w_1,w_2, ...,w_m$, with respect to each criterion C_k. A vector of the weights, $\mathbf{w}=[w_i]$, $w_i>0$, $i=1,...,n$, may be determined by using the eigenvalue-eigenvector method: $\mathbf{Aw}=\lambda\mathbf{w}$. Saaty proved [12] that the priority score of an alternative, what he called the *relative dominance* of the ith alternative A_i, is the ith component u_i of the principal right (Perron) eigenvector **u** of matrix **A**, even if the PCM is *not* transitive, i.e. where sets of distinct alternatives can be found for which $a_{ij} \cdot a_{jk}\neq a_{ik}$, for $i,j,k=1,2, ...,n$. For a transitive (consistent) **A** the maximal eigenvalue is: $\lambda_{\max}=n$, whereas for a nontransitive (inconsistent) **A**: $\lambda\max>n$. (see the proofs in [13]). The eigenvector method provides the true relative dominances of the alternatives only if **A** is a *transitive* matrix. In reality, however, an individual cannot give his/her estimates in a perfectly consistent way. Recognizing this fact, Saaty [12] proposed an *index* for measuring the *inconsistency* of a PCM in the form: $\mu(\mathbf{A})=(\lambda_{\max}-n)/(n-1)$. To compute the components of the *overall priority scores*, $\pi_1,\pi_2,...,\pi_n$, for a given set of alternatives the AHP utilizes an *additive* type aggregation function: $\pi_i = \sum_{k=1}^{m} c_k \cdot w_{ik}$, $i=1,2, ...,n$.

In this study, the aggregated PCM with the right hand side Perron eigenvector elicited from the respondents is obtained as:

$$\mathbf{A} = \begin{bmatrix} 1 & 1/2 & 1/2 & 1/3 & 1/4 \\ 2 & 1 & 1 & 1 & 1/2 \\ 2 & 1 & 1 & 1/2 & 1/2 \\ 3 & 1 & 2 & 1 & 1 \\ 4 & 2 & 2 & 1 & 1 \end{bmatrix},$$

$$\mathbf{u} = \begin{bmatrix} 0.3146 \\ 0.2619 \\ 0.1832 \\ 0.1573 \\ 0.0830 \end{bmatrix} \begin{matrix} \to R \\ \to A \\ \to E \\ \to M \\ \to C \end{matrix} .$$

Observe here, that although there are only slight inconsistencies in the responses, **A** is *not* a transitive matrix, as $\lambda_{\max}(\mathbf{A})$=5.0523. The inconsistency index is: $\mu(\mathbf{A})$=0.0131.

4 The use of the LS method for benchmarking

The objective here is to produce the 'best' transitive (rank one) matrix approximation **B** to matrix **A** where the 'best' is assessed in a least-squares (LS) sense. For that, the following expression should be minimized:

$$S^2(\mathbf{w}) := \|\mathbf{A}-\mathbf{B}\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij}-\frac{w_j}{w_i}\right)^2, \tag{2}$$

where the subscript F denotes the Frobenius norm; i.e. the *error*, which is the square root of the sum of squares of the elements. A *stationary vector* **w** (a local minimum) of the error functional $S^2(\mathbf{w})$ is a (positive) solution to a set of inhomogeneous nonlinear equations (see [10]). Using the Newton-Kantorovich numerical optimization method for carrying out the necessary computations we may obtain the rank one matrix **B** and the associated priority vector $\mathbf{w}^{-1}$. In approximating matrix **A** we get

$$\mathbf{B} = \begin{bmatrix} 1 & 0.4693 & 0.5309 & 0.3234 & 0.2556 \\ 2.1307 & 1 & 1.1312 & 0.6891 & 0.5446 \\ 1.8837 & 0.8841 & 1 & 0.6092 & 0.4815 \\ 3.0920 & 1.4512 & 1.6415 & 1 & 0.7903 \\ 3.9122 & 1.8361 & 2.0769 & 1.2653 & 1 \end{bmatrix},$$

$$\mathbf{w}^{-1} = \begin{bmatrix} 0.3255 \\ 0.2573 \\ 0.1773 \\ 0.1567 \\ 0.0832 \end{bmatrix} \begin{matrix} \to R \\ \to A \\ \to E \\ \to M \\ \to C \end{matrix} \; .$$

The interested reader may turn to [1] or to [10] for a detailed discussion of the least-squares optimization method for SR matrices.

In order to improve the adjustment of the right hand side principal eigenvector $\mathbf{w}^{-1}$ of matrix **B** to the Perron eigenvector **u** (and also to **s**) of the matrices **A** and **S** a recursive least-squares (LRS) algorithm has been developed [9]. The LRS generates a series of transitive matrices and converges to the residual matrices, denoted by $\mathbf{B}^*_k$ and $\mathbf{H}_k$, k=0,1,2,…,q, (derived from **A)**, which have useful properties for this adjustment.

In the present paper, two new measures for the perturbation of SR matrices are used to characterize the magnitude and the variability of the average inconsistency of matrix **A**. These measures are the geometric mean and the geometric standard deviation of the appropriate values of the residual matrix $\mathbf{H}_{k=q}$. This matrix yields in the last step q of the iteration as function of a prescribed arbitrarily small level of accuracy $\varepsilon>0$.

Running the LSR on “Mathematica” starting from the matrix **A**, a convergent process has produced the following transitive matrix and the associated priority vector at the last step of the iteration, $k=q=5$, in a limiting sense:

$$\mathbf{B}_5^* = \begin{bmatrix} 1 & 0.4605 & 0.5297 & 0.3225 & 0.2648 \\ 2.1718 & 1 & 1.1504 & 0.7004 & 0.5752 \\ 1.8879 & 0.8693 & 1 & 0.6089 & 0.5000 \\ 3.1006 & 1.4277 & 1.6424 & 1 & 0.8212 \\ 3.7757 & 1.7386 & 2.0000 & 1.2178 & 1 \end{bmatrix},$$

$$\mathbf{w}^{-1*} = \begin{bmatrix} 0.3163 \\ 0.2598 \\ 0.1819 \\ 0.1582 \\ 0.0838 \end{bmatrix} \begin{matrix} \to R \\ \to A \\ \to E \\ \to M \\ \to C \end{matrix} .$$

The residual matrix $\mathbf{H}_5$, wherein the deviations of the elements from the value of 1 are the perturbations of the entries of matrix **A**, yields

$$\mathbf{H}_5 = \begin{bmatrix} 1 & 1.0859 & 0.9439 & 1.0335 & 0.9439 \\ 0.9209 & 1 & 0.8693 & 1.4277 & 0.8693 \\ 1.0594 & 1.1504 & 1 & 0.8212 & 1 \\ 0.9676 & 0.7004 & 1.2178 & 1 & 1.2178 \\ 1.0594 & 1.1504 & 1 & 0.8212 & 1 \end{bmatrix}.$$

One obvious verification concerning the preservation of the spectral properties of matrix **A** is justified by the fact that $\lambda_{max}(\mathbf{H}_5)=\lambda_{max}(\mathbf{A})=5.0523$.

5 Concluding remarks

In this project a multi-criteria evaluation problem for benchmarking of bridges has been addressed. A comprehensive numerical analysis has demonstrated that there is *no contradiction* between the applied multi-attribute scaling methods of normative and descriptive types. Saaty himself stated that "measurements from an interval scale may be converted to ratios and used as priorities if there is adequate justification for using them in that manner" ([14], p.260).

In this study, the **Árpád bridge** was ranked first as the best bridge design in Budapest among the four selected bridges under investigation. The examinations were based on a complex multi-criteria evaluation of the bridges. The backward of the other bridges (among them were the **Chain bridge** and the **Elisabeth bridge** contrary to their world famous recognition) may be attributed to their lack of suitability to the requirements and new challenges of the modern era. Remarkably, all scaling methods applied in this study have produced the same priority ranking of the alternatives. A more subtle analysis for a possible improvement on the single attributes of every bridge in relation to the standards of the benchmark bridge ***R***, however, is a subject of future research.

If one is willing to accept the reasoning for an itemized approach and the necessity of an adequate assignment of the attributes to the appropriate scales of measurement, then we might consider the final results derived by the MTSE method as a reference point (benchmark).

Determining the ‘goodness’ of matrix approximations to matrix **S** in a least-squares sense**,** we found for the different approaches that

$$S(\mathbf{u}) := \left\|\mathbf{S}-\mathbf{A}\right\|_F = 0.8946,$$

$$S(\mathbf{w}) := \left\|\mathbf{S}-\mathbf{B}\right\|_F = 0.3685,$$

$$S(\mathbf{w}^*) := \left\|\mathbf{S}-\mathbf{B}_5^*\right\|_F = 0.3433.$$

From the above results, it is straightforward that the LSR method produced the least error in approximating the matrix **S**. Thus, indeed, matrix $\mathbf{B}^*_5$ and its associated priority vector $\mathbf{w}^{-1*}$ provide us the best approximations. This fact can also be observed through a direct numerical comparison of the corresponding eigenvector elements.

The magnitude and the variability of the average inconsistency of matrix **A** can be characterized by the geometric mean and the geometric standard deviation of the entries which appear in the upper triangle of the residual matrix $\mathbf{H}_5$. This latter matrix yielded in the last step of the iteration with a prescribed level of accuracy: $\varepsilon=10^{-6}$. With these matrix elements the measures of the level of inconsistency of the original matrix **A** can be computed. They are: $g_5(\mathbf{A})=1.0076$ and $s_{g5}(\mathbf{A})=1.1741$, respectively. After the initial step of the iteration, for the matrix $\mathbf{H}_1$ these measures were equal to: $g_1(\mathbf{A})=1.0254$ and $s_{g1}(\mathbf{A})=1.1749$, respectively. The smaller values of $g_5(\mathbf{A})$ and $s_{g5}(\mathbf{A})$ indicate that the “true” errors which have actually been committed by the respondents can be obtained by using the LSR method, if we wish to find the best adjustment of the right hand side principal eigenvector $\mathbf{w}^{-1}$ of matrix **B** to the Perron eigenvector **u** (and also to **s**) of the matrices **A** and **S**.

Finally, it is essential to express our strong conviction that the approach presented in Section 4 would facilitate the practitioners to improve the original PCM in real world applications obtained through the judgmental process of the AHP method. In this way, a respondent is capable of finding the appropriate priority vector of the alternatives that best suits his/her needs.

References

[1] Chu,A.T.W., Kalaba,R.E. and Springarn,K. “A comparison of two methods for determining the weights of belonging to fuzzy sets”, *Journal of Optimization Theory and Applications*, **27**, (1979), pp.531-538.

[2] Cook,W.D. and Seiford,L.M. “On the Borda-Kendall consensus method for priority ranking problems”, *Management Science*, **28**, (1982), pp.621-637.

[3] Dabous,S.A. “A multi-attribute ranking method for bridge management”, Engineering, Construction and Architecture Management, 17, (2010), pp.282-291.

[4] Dyer,J.S. and Sarin,R. ”Measurable multi-attribute value functions”, Operations Research, 27, (1979), pp.810-822.

[5] Farkas,A. MAROM. (MultiAttRibute Object Measurement). A decision support system and software package. International Management Center (IMC), Budapest, 1993.

[6] Farkas,A. “Priority ranking methods: A survey and an extension”, in: Business Research and Management Challenges in Central and Eastern Europe. (ed. Peter,S.), International Management Center, Budapest, (1994), pp.74-94.

[7] Farkas,A. “Die kardinale Messung von Verbraucherpräferenzen”, Zeitschrift für Marktforschung und Marketing. (Planung & Analyse). 21, (2006), pp.75-78.

[8] Farkas,A. “Multicriteria comparison of bridge designs”, Acta Polytechnica Hungarica. 8, (2011), pp.173-191.

[9] Farkas,A. and Rózsa,P. ”A recursive least-squares optimization algorithm for pairwise comparison matrices”, Unpublished manuscript. (2010), p.25

[10] Farkas,A., Lancaster,P. and Rózsa,P. ”Consistency adjustments to pairwise comparison matrices”, Numerical Linear Algebra with Applications, 10, (2003), pp. 689-700.

[11] Horsky,D. and Rao,M.R. “Estimation of attribute weights from preference comparisons”, Management Science, 30, (1984), pp.801-822.

[12] Saaty,T.L. “A scaling method for priorities in hierarchical structures”, Journal of Mathematical Psychology”, 15, (1977), pp.234-281.

[13] Saaty,T.L. The Analytic Hierarchy Process. McGraw Hill, New York, 1980.

[14] Saaty,T.L. "An exposition of the AHP in reply to the paper 'Remarks on the Analytic Hierarchy Process' ", Management Science, 36, (1990), pp.259-268.