

Statistical methods applied for materials selection

C. Târcolea, A. Paris, A. Demetrescu-Târcolea

Abstract. In the design process the choice of materials for a new product is made from a very large spectrum. Usually in the first stage, for each material, it is taken into consideration, as many attributes as possible is. In the second stage it is taken into consideration only a few "virtual" attributes (usually two or three), given by principal component analysis. The main idea is exemplified by the particular case of fiberglass-reinforced thermoplastics.

M.S.C. 2000: 62H25, 62J15.

Key words: Principal component analysis, artificial factors.

1 Introduction

"The goal of design is create products that perform their function effectively, safely, at acceptable cost" (Ashby, 1999) [1]. The choice of a material for a product is not based on all the attributes, is based on a combination of properties. In this paper it is presented a method of reduction of attributes to a combination of a few artificial factors. The selection of the materials group uses the following premises: -the taking into consideration of an ever larger spectrum of materials, otherwise expressed, defining as complete as possible the selection range -the division of materials in "equivalence classes" otherwise expressed as the classification of materials thus obtaining a significant reduction of the selection space, the initial chaotically multitude is reduced to smaller number of classes, the number of each class having common properties, processes accompanied or not by the elimination or some materials taken into consideration - the establishing of the same simple and efficient methods of selection, with the possibility of implementing on a computer [1], [2], [8].

A possibility of getting down the properties range is the taking into consideration of the correlation between the properties, studying thus only a few independent properties to the possible extent. The introduction of the up-to-date calculation enables the data stocking concerning the properties, from the standard version, under the normal temperature and the time $t = 0$, up to more complete versions, taking into account their variability as against the temperature, time, etc.[8].

The classification of materials is often empirically made, for example in six groups [2,8] metallic materials, polymers, elastomers, glasses, ceramics and compound

materials. Another variant [8] uses as criterion the cristallinity of materials (crystalline or non-crystalline and compound) each class being grouped in underclasses as per the structure type, way of obtaining, etc.

2 Application of principal component analysis

The PCA is a standard technique to reduce multivariate data sets in a subspace of small dimension, in our case a trivariate one [3,10]. The number of observable attributes gives the dimension of the initial representation space of the objects [6,7]. Instead of realer attributes, the PCA uses new factors, but artificial ones. In the present paper, the dimensional reduction of attributes for a family of materials follows the Jöreskog's researches [4,5].

Let only 9 materials (objects), each of them having 6 attributes (table 1). The resulted three-dimensional subspace gives a minimal deformation of distances through projection. As illustration let's take an application from "Reinforced Plastics Handbook" p. 889 [9], concerning selecting plastics and process: for nine glass fiber (30wt%) reinforced thermoplastics are presented six important mechanical and cost properties for a detailed analysis.

Prop. / Materials	P1	P2	P3	P4	P5	P6
ABS	1.29	100 (14.5)	7.58 (1100)	102 (215)	69 (1.3)	1.4
POM(c.c.)	1.63	134 (19.5)	9.65 (1400)	166 (330)	96 (1.8)	2.2
ETFE	1.89	97 (14)	7.24 (1050)	238 (460)	320 (6)	19
PA 6	1.37	159 (23)	8.27 (1200)	216 (120)	123 (2.3)	1.9
6/6	1.38	179 (26)	8.96 (1300)	254 (490)	107 (2)	2.7
6/6(f.r.)	1.59	134 (19.5)	7.58 (1100)	243 (470)	69 (1.3)	(3.3)
6/10, 6/12	1.30	148 (21.5)	7.58 (1100)	214 (117)	128 (2.1)	1
PPA	1.43	221 (32)	11.4 (1650)	285 (545)	118 (2.2)	6.2
LCP	1.57	162 (23.5)	12.8 (3120)	318 (614)	85 (1.7)	32

Table 1. Analyzed properties of nine thermoplastics

The matrix X is a 6×9 dimensional matrix, attributes/materials:

$$X = \begin{pmatrix} 1,29 & 100 & 7,58 & 102 & 69 & 1,4 \\ 1,63 & 134 & 9,65 & 166 & 96 & 2,2 \\ 1,89 & 97 & 7,24 & 238 & 320 & 19 \\ 1,37 & 159 & 8,27 & 216 & 123 & 1,9 \\ 1,38 & 179 & 8,96 & 254 & 107 & 2,7 \\ 1,59 & 134 & 758 & 243 & 69 & 3,3 \\ 1,3 & 148 & 7,58 & 214 & 128 & 1 \\ 1,43 & 221 & 11,4 & 285 & 118 & 6,2 \\ 1,57 & 162 & 12,8 & 318 & 85 & 32 \end{pmatrix}$$

The standardized matrix of X , denoted Z , is:

$$Z = \begin{pmatrix} -1,0522 & -1,25 & -0,3398 & -1,953 & -0,715 & -0,5938 \\ 0,69766 & -0,369 & -0,3315 & -0,947 & -0,363 & -0,5189 \\ 2,03579 & -1,328 & -0,3411 & 0,185 & 2,5555 & 1,05346 \\ -0,6405 & 0,2794 & -0,337 & -0,161 & -0,012 & -0,547 \\ -0,589 & 0,7979 & -0,3342 & 0,437 & -0,22 & -0,4721 \\ 0,49179 & -0,369 & 2,66659 & 0,264 & -0,715 & -0,416 \\ -1,0007 & -0,006 & -0,3398 & -0,192 & 0,0536 & -0,6312 \\ -0,3317 & 1,8867 & -0,3245 & 0,924 & -0,077 & -0,1446 \\ 0,38886 & 0,3572 & -0,3188 & 1,443 & -0,507 & 2,2702 \end{pmatrix}$$

Based on matrix equation

$$R = \frac{1}{8} Z_{tr} * Z, \tag{1}$$

it results that:

$$R = \begin{pmatrix} 1 & -0,36956 & 0,184942 & 0,274144 & 0,657701 & 0,549213 \\ -0,3696 & 1 & -0,13352 & 0,626218 & -0,32582 & -0,03745 \\ 0,18494 & -0,13352 & 1 & 0,103873 & -0,27123 & -0,1519 \\ 0,27414 & 0,62622 & 0,103873 & 1 & 0,139821 & 0,610039 \\ 0,6577 & -0,32582 & -0,27123 & 0,139821 & 1 & 0,317506 \\ 0,54921 & -0,03745 & -0,1519 & 0,610039 & 0,317506 & 1 \end{pmatrix}$$

It is obvious that R is the correlation matrix. The proximity between attributes is expressed in terms of correlations. The inverse matrix of R is:

$$R^{-1} = \begin{pmatrix} 4,24886 & -2,28953 & -3,00663 & 3,409168 & -3,62451 & -3,80489 \\ -2,2895 & 16,1683 & 8,535721 & -18,9344 & 7,857795 & 12,21537 \\ -3,0066 & 8,53572 & 6,48209 & -10,6674 & 5,564535 & 7,696317 \\ 3,40917 & -18,9344 & -10,6674 & 24,00877 & -9,64994 & -15,7842 \\ -3,6245 & 7,8578 & 5,564535 & -9,64994 & 6,605492 & 6,919718 \\ -3,8049 & 12,2154 & 7,696317 & -15,7842 & 6,919718 & 12,14818 \end{pmatrix}$$

It obtains: $\text{diag}(R^{-1}) = \text{diag}(4.2488; 16.168; 6.482; 24.008; 6.605; 12.148)$

$$\text{diag}(R^{-1})^{\frac{1}{2}} = \text{diag}(2.061; 4.021; 2.546; 4.899; 2.570; 3.485)$$

$$\text{diag}(R^{-1})^{-1} = \text{diag}(0.235; 0.0618; 0.154; 0.0416; 0.151; 0.082).$$

Jöreskog’s method considers that theoretical covariance matrix of the standardized attributes, V , is factorized as:

$$V = L_{tr} * L + \text{diag}(s_1^2, \dots, s_6^2). \tag{2}$$

For the uniqueness of the estimators of the matrices from the decomposition of V it is supposed that the variances are direct proportional with the inverse values of the diagonal elements of the matrix V^{-1} ; that is:

$$\text{diag}(s_1^2, \dots, s_6^2) = m(\text{diag}(V^{-1}))^{-1}, \tag{3}$$

where m is the proportional parameter. Further on, it is denoted by R^* the following matrix:

$$R^* = \text{diag}(R^{-1})^{\frac{1}{2}} * R * \text{diag}(R^{-1})^{\frac{1}{2}} \tag{4}$$

It results:

$$R^* = \begin{pmatrix} 4,248862 & -3,06309 & 0,970577 & 2,768855 & 3,484313 & 3,945777 \\ -3,06309 & 16,16834 & -1,3669 & 12,33796 & -3,36711 & -0,52491 \\ 0,970577 & -1,3669 & 6,48209 & 1,295822 & -1,77477 & -1,34791 \\ 2,768855 & 12,33796 & 1,295822 & 24,00877 & 1,760794 & 10,41833 \\ 3,484313 & -3,36711 & -1,77477 & 1,760794 & 6,605492 & 2,844202 \\ 3,945777 & -0,52491 & -1,34791 & 10,41833 & 2,844202 & 12,14818 \end{pmatrix}.$$

The eigenvalues of this matrix are given in the table 2 and the first three eigenvectors in the table 3:

36,14978 19,16469 8,017578 5,002533 1,114291 0,212852

Table 2. Eigenvalues of the matrix R^*

Because

$$\frac{\lambda_1 + \lambda_2 + \lambda_3}{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6} \approx 0.91 > 0.90$$

it can be considered only the 3-dimensional subspace.

T	U	W
-0,0726	0,379876	-0,13195
-0,46987	-0,62708	0,215822
0,002525	-0,01	-0,89147
-0,80398	0,075068	-0,18658
-0,03727	0,407536	0,301071
-0,3552	0,53912	0,125862

Table 3. The values of the first three eigenvectors of the matrix R^*
An estimator for the parameter m is the average of the remained eigenvalues:

$$m = \frac{1}{3}(\lambda_4 + \lambda_5 + \lambda_6) = 2.11$$

The parameters of interest are the variances, given by:

$$s_i^2 = \frac{m}{r^{ii}}, \quad i \in \overline{1,6} \tag{5}$$

where r^{ii} is the element ii of the matrix R^{-1} .

In the present case, it is obtained the values of s_i^2 (table 4):

$$0,49660356 \quad 0,130502 \quad 0,325512 \quad 0,087885 \quad 0,319431 \quad 0,173689$$

Table 4. Values of s_i^2

Further on, it should be calculate:

$$k_i^2 = 1 - s_i^2, \quad i = \overline{1,6} \quad (6)$$

The results are given in the table 5:

$$0,50339644 \quad 0,869498 \quad 0,674488 \quad 0,912115 \quad 0,680569 \quad 0,826311$$

Table 5. Values of k_i^2

The three eigenvectors, denoted as T, U, W associated with the first three chosen eigenvalues are determined so that:

$$\|T\|^2 = \lambda_1 - \hat{m}, \quad \|U\|^2 = \lambda_2 - \hat{m}, \quad \|W\|^2 = \lambda_3 - \hat{m}. \quad (7)$$

The matrix of these eigenvectors is:

$$A = \begin{pmatrix} -0,42355 & 1,568786 & -0,32071 \\ -2,74139 & -2,58965 & 0,524566 \\ 0,014732 & -0,0413 & -2,16675 \\ -4,69073 & 0,310011 & -0,4535 \\ -0,21747 & 1,683015 & 0,731768 \\ -2,07237 & 2,226422 & 0,305914 \end{pmatrix}$$

In the next step it is obtained the loading matrix:

$$L = (\text{diag}(R^{-1}))^{-\frac{1}{2}} * A, \quad (8)$$

$$L = \begin{pmatrix} -0,20548 & 0,761075 & -0,15559 \\ -0,68177 & -0,64403 & 0,130457 \\ 0,005786 & -0,01622 & -0,85104 \\ -0,95732 & 0,063269 & -0,09255 \\ -0,08461 & 0,65484 & 0,284722 \\ -0,59458 & 0,638781 & 0,08777 \end{pmatrix}$$

It results:

$$k_i^2 \cong l_{i1}^2 + l_{i2}^2 + L_{i3}^2, \quad \text{for } i \in \overline{1,6}, \quad (9)$$

which confirm the correctness of the model.

3 Conclusions

Revising the up-to-date problems of the materials selection, it was emphasized on the statistical approach, starting from the bi-or three dimensional graphical drawing of some properties. The design engineer should have in this manner at his disposal an effective system, optimum for the selection of the material. The ellipse or ellipsoid

may be built, as in the shown method, all the possibility of assuming some different probabilities for the random vector to take the values in the area bounded by the ellipse or ellipsoid.

The application of this model will simplifies the materials design and there are many other possible extensions in the design process. Furthermore, a mixture between Jöreskog' method and materials design principles is useful for the developments.

References

- [1] M. F. Ashby, *Materials selection in mechanical design*. Butterworth-Heinemann Edition, Oxford GB, 1999.
- [2] M. F. Ashby, K. Johnson, *Materials and Design, the Art and Science of Materials Selection in Product Design*, Butterworth Heinemann, Oxford, GB, 2002.
- [3] M. Dumitrescu, A. Batatorescu, *Applied Statistics using the R-System*, Editura Universitatii din Bucuresti, 2006.
- [4] J. Hartung, B.Elzelt, *Multivariate Statistik*, Oldenbourg Verlag, München, Germany 1986.
- [5] K. G. Jöreskog, *Statistical Estimation in Factor Analysis, a New Technique and Its Foundations*, Ed. Almqvist&Wiksell, Uppsala, 1963.
- [6] L. Lemnete-Ninulescu, *A moment problem with values positive definite matrices*, BSG PROCEEDINGS 14, (MENP-4), 2006, Bucharest, ROMANIA, Geometry Balkan Press (2007), 95-98.
- [7] M. Migdalovici, D. Baran, *Theoretical research the Floquet stability theorem with applications*, BSG (MENP-4), 2006, Bucharest, ROMANIA, Geometry Balkan Press (2007), 106-110.
- [8] A. Paris, S. Ionescu, P. Szel, C. Târcolea, *Statistical approach in the selection of materials for engineering products*. Scientific Seminar "Modern mechanical design" Bucharest, U.P.B. (1991), 97-106.
- [9] D. V. Rosato, *Reinforced Plastics Handbook*, Elsevier Edition, B.V. 2005.
- [10] C. Târcolea, A. Paris, *The Jöreskog technique applied for materials design*, Proceedings of the 17-th, ICMaS, Ed Acad. Romane, (2008), 309-312.

Author's address:

Constantin Târcolea
University Politehnica of Bucharest, Faculty of Applied Sciences,
Department of Mathematics-Informatics I,
Splaiul Independentei 313, RO-060042, Bucharest, Romania.

Adrian Paris,
University Politehnica of Bucharest, IMST-Faculty Department of Technology,
Splaiul Independentei 313, RO-060042, Bucharest, Romania.
E-mail: sparis@rdslink.ro

Adina Demetrescu-Târcolea
Gruneburgplatz 49, 60629 Frankfurt am Main, Germany.
E-mail: adina_tarcolea@yahoo.com