Journal of University of Duhok, Vol. 24, No.1 (Pure and Eng. Sciences), Pp 39-45, 2021
https://doi.org/10.26682/sjuod.2021.24.1.6

MATHEMATICAL TREATMENT OF DUAL SCALING

RAFAL SAUD IDHAM and KHAWLA MUSTAFA SADIK AL-BAKIR
Dept. of Mathematics, College of Education Pure Science, University of Mosul-Iraq

(Received: November 30, 2020; Accepted for Publication: March 7, 2021)

ABSTRACT

On the basis of the Preparations so far, we are now ready to look at a formal derivation of dual scaling. As you may have guessed from the many names for the quantification method and from the earlier discussion of the many ideas behind dual scaling, there are a number of ways to formulate it. This paper contains a detailed description of one of them. Some readers may find this paper too technical η Many formulas are presented in this paper occasionally without sufficient background information .However, once of the basic idea is understood, technical matters can be mastered through self-directed endeavor, and this paper should serve such guide by offering bird's eye view of mathematical treatment of dual scaling.

OPTIMAL SOLUTION

The Principle component analysis, in which the object is to determine weights for categories so as to maximize the variance of the composite scores, the analysis - of - variance approach , in which category weights are determined so as to maximize the between – row (column) sum of squares, relative to the total sum of squares , that is the squared correlation ratio, the bivariate correlation approach , in which the product – moment correlation between responses weighted by row weights and those by column weights maximized ,and the simultaneous linear regression approach ,in which row and column weights are determined so as to make the regression of rows on columns and the regression of columns on rows simultaneously linear .This decomposition is a basis for the one –way analysis ,of variance and when data are continues ,the decomposition is uniquely determined in contrast when dated are categorical , relative magnitudes of those sums of squares vary , depending on weights given to the categories .In other words, one can manipulate weights to change the relative magnitudes of Each sum of squares. Dual scaling chooses the weights in a particular way, the weights are determined in such a way that the ratio SS_b/SS_t that is the squared correlation ratio η^2. be a maximum,In our task is to determine weights X_j so as to maximize η^2 ,(1) In terms of matrix notation see equation (1) ,(2) ,(3) and (4)

	Good	Averg	Poor	
White	f_{11}	f_{12}	f_{13}	$y_1 = \sum x_j f_{1j}$
Green	f_{21}	f_{22}	f_{23}	$y_2 = \sum x_j f_{2j}$
Brown	f_{31}	f_{32}	f_{33}	$y_3 = \sum x_j f_{3j}$
Total	$\sum f_{j1}$	$\sum f_{j2}$	$\sum f_{j3}$	$\sum\sum x_j f_{ij} = y_t$

$$ss_b = \Sigma(y_j^2 / f_j) - y_t^2 / f_t$$

$$ss_w = \Sigma\Sigma y_{ij}^2 - \Sigma(y_j^2 / f_j)$$

$ss_t = ss_b + ss_w$

$\frac{ss_b}{ss_t} = \eta^2$

If we have a matrix we have to equation (3,8)

F= (f_{ij}):the Data matrix

f_r :the vector of row totals of F

f_c:the vector column totals of F

D_r:the diagonal matrix with row totals in the main diagonal

D_c: A diagonal diagonal matrix with column totals in the main diagonal[1]

γ: avector of weights for the rows of F

χ: avector of weights for the columns of F

f_t: the sum of element of F, that is $\sum\sum f_{ij}$

rafal.suood.99@gmail.com

https://doi.org/10.26682/sjuod.2021.24.1.6

Table 1

Teacher	Good	Averge	Poor	Total
White	1	3	6	10
Green	3	5	2	10
Brown	6	3	0	9
	10	11	8	29

(2)

An illustrative example:

Let us follow the entire process of dual scaling, using the data in Table 1

That is the data used to illustrate the method of reciprocal averages. Let us assign unknown $\mathcal{X}_1$ $\mathcal{X}_2$ and $\mathcal{X}_3$ to three categories good, average and poor, respectively. Since the origin and the unit of quantified data are arbitrary, it is customary to see the sum of weighted responses and the sum of squares of weighted responses equal to zero and fi

(the total number of responses), respectively:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}\,\mathcal{X}_j = 0 \qquad \dots \quad (1)$$

$$\sum_{i=1}^{n}\sum_{j=1}^{m} f_{ij}\,\mathcal{X}^2{}_j = f_i \qquad \dots \quad (2)$$

The choice of the origin by (1) makes the average of the weighted response zero , eliminating. The correction term

$$SS_t = 10\mathcal{X}_1^2 + 11\mathcal{X}_2^2 + 8\mathcal{X}_3^2 \quad (3)$$

$$SS_b = \frac{(x_1 + 3x_2 + 6x_3)^2}{10} + \frac{(3x_1 + 5x_2 + 2x_3)^2}{10} + \frac{(6x_1 + 3x_2)^2}{9} \qquad \dots (4)$$

2

(3),(4)

In the current example we have

$$F = \begin{pmatrix} 1 & 3 & 6 \\ 3 & 5 & 2 \\ 6 & 3 & 0 \end{pmatrix}$$

$$f_r = \begin{pmatrix} 10 \\ 10 \\ 9 \end{pmatrix} \quad , \quad f_c = \begin{pmatrix} 10 \\ 11 \\ 8 \end{pmatrix}$$

$$D_r = \begin{pmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 9 \end{pmatrix} \quad , \quad D_c = \begin{pmatrix} 10 & 0 & 0 \\ 0 & 11 & 0 \\ 0 & 0 & 8 \end{pmatrix} \quad , \quad f_t = 29$$

$$,y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \quad , \quad \mathcal{X} = \begin{pmatrix} \mathcal{X}_1 \\ \mathcal{X}_2 \\ \mathcal{X}_3 \end{pmatrix}$$

rafal.suood.99@gmail.com

https://doi.org/10.26682/sjuod.2021.24.1.6

$$1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

[3] (8)

$$f_c' x = 0 \quad \ldots \quad (5)$$

$$x' D_C x = f_t \quad \ldots \quad (6)$$

$$SS_t = x' D_C x \quad \ldots \quad (7)$$

$$SS_b = x' F' D_r^{-1} F x \quad \ldots \quad (8)$$

$$^1 F' D_r^{-1} F x = (x_1, x_2, x_3) \begin{pmatrix} 1 & 3 & 6 \\ 3 & 5 & 3 \\ 6 & 2 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{10} & 0 & 0 \\ 0 & \frac{1}{10} & 0 \\ 0 & 0 & \frac{1}{9} \end{pmatrix} \begin{pmatrix} 1 & 3 & 6 \\ 3 & 5 & 2 \\ 6 & 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

rafal.suood.99@gmail.com

(5)

Carry out the remainder of this numerical verification step by step . You will find the expression to be identical with Equation (4).

The object of dual scaling in this approach is to determine $\mathcal{X}$ so as to maximize the squared correlation ratio η^2 , which is given by SS_b/SS_t.

This problem can be handled in two ways one is to maximize SS_b/SS_t in terms of , and the other is to maximize SS_b subject to the condtion that $SS_t = f_t$ as indicated by equation (7) . Both lead to the same outcome. A widely used Procedure is the later one.

which is handled by the Well-Known Lagrange method of unknown multipliers .

Now , in the Lagrange method , the objective function for optimization is called Lagrange function which is in the present example given by :

$$Q(\mathcal{X},\lambda) = SS_b - \lambda(SS_t - f_t)$$
$$= x'F'D_r^{-1}Fx - \lambda(x'D_Cx - f_t)$$

(6)The function Q has two unknowns , $\boldsymbol{\mathcal{X}}$ and λ and we wish to determine $\mathcal{X}$ and λ so as to maximize Q , Following the standard procedure

for maximizing quadratic function, we set the partial derivatives of Q with respect to $\boldsymbol{\mathcal{X}}$ and y equal to zero .

$$\frac{\partial Q}{\partial x} = 2F'D_r^{-1}Fx - 2\lambda D_cx = 0 \quad \dots \quad (9)$$

$$\frac{\partial Q}{\partial \lambda} = x'D_Cx - f_t = 0$$

The first formula can be written as

$$F'D_r^{-1}Fx = \lambda D_Cx \quad \dots \quad (10)$$

If we premultiply both sides of this expression by x'we obtain

$$x'F'D_r^{-1}Fx = \lambda x'D_Cx$$

$$\lambda = \frac{x'F'D_r^{-1}Fx}{x'D_Cx} = \eta^2$$

(7)

Thus, the unknown multiplier λ is nothing but the squared correlation ratio that we want to maximize .

There for equation(9) can be written

$$(F'D_r^{-1}F - \eta^2 D_C)x = 0$$

In our example

$$D_c^{-1}F'D_r^{-1}F = \begin{pmatrix} 0.5 & 0.38 & 0.12 \\ 0.35 & 0.40 & 0.25 \\ 0.15 & 0.33 & 0.50 \end{pmatrix}$$

To arriveat the standard form we will use the following transformation

$$B = D_r^{-\frac{1}{2}}FD_c^{-\frac{1}{2}} \quad \dots \quad (11)$$

$$B'B = D_c^{-\frac{1}{2}}F'D_r^{-1}FD_c^{-\frac{1}{2}}$$

And setting $D^{\frac{1}{2}}x = w$

We obtain the following eigenequation in the standard form .

$$(B'B - \eta^2 I)w = 0 \quad \dots \quad (12)$$

Equation (12) can be written as

[5] $(0.5000 - \eta^2)w_1 + 0.3623w_2 + 0.1342w_3 = 0$

$0.3623w_1 + (0.4000 - \eta^2)w_2 + 0.2985w_3 = 0$

$0.1342w_1 + 0.2985w_2 + (0.5000 - \eta^2)w_3 = 0$

Eigenequation (12) has two trival equation

1: w =0 … (13)

It is obvious that(17) is solution .But ,as we have already this solution is avoided by solving the characteristic equation that is

$$|B'B - \eta^2 I| = 0$$

Trival solution[2]

w= $D_c^{\frac{1}{2}}1$ that is $\mathcal{X}$=1 … (14)

In this case the corresponding eigenvalue η^2 is 1 . This can easily be shown as fallows .first rew rite equation (12) as

$$B'Bw = \eta^2 w \quad \dots \quad (15)$$

Then if we substute

w= $D_c^{\frac{1}{2}}1$

For the left – hand side of equation (15) we obtain

$$B'Bw = D_c^{-\frac{1}{2}}F'D_r^{-1}FD_c^{-\frac{1}{2}}D_c^{\frac{1}{2}}1$$

$$D_c^{-\frac{1}{2}}F'D_r^{-1}F1$$

$F1 = f_r$ then $D_r^{-1}f_r = 1$

$$D_c^{-\frac{1}{2}}F'1 = D_c^{-\frac{1}{2}}f_c$$

$$= D_c^{\frac{1}{2}}1 \quad \dots \quad (16)$$

rafal.suood.99@gmail.com

By subsitut w= $D_c^{\frac{1}{2}}$ 1 for the right hand to equation (14)which gives η^2 $D_c^{\frac{1}{2}}$ 1
And compare this equation with equation(15) then η^2= 1 when w= D_c1 , Wher $\boldsymbol{\mathcal{X}}$=1
This allow us to put columns (categories) in (10) do even SS_w and SS_b = zero then that $\mathcal{X}$=1 amount to putting all the columns(categories) in to one, making both SS_w and SS_b Zeros , hence no information
to analyze for the move this"Solution" vector dose not satisfy the condition that$F'_c\ \mathcal{X} = 0$.
Thus we must eliminate the second trivial solution from our computation .To carry out this task ,the following theorem relevant :
LaGrange's Theorem (Rao 1965) : Let G be any square matrix of order n and rank m> 0 , and p and q be column vector such that P' G q is not zero.
Then, the residual matrix G_1 is exactly of rank m-1 , where

$$G_1 = G - Gq\ \ (P'\ \mathrm{G\,q})^{-1}\ P'\ \mathrm{G} \quad \dots (17)$$

We can use this formula to eliminate the second trivial solution from B′ B of equation (11) is symmetric and that the trivial solution is an Eigen vector. Therefore, p= q= a, say, and Ga = λ a, an application of equation (17) to an Eigen equation yields the following formula for eliminating the effect of solution from an Eigen equation(5)

$$G_1 = G - \lambda\, \mathrm{a}(a'\ \mathrm{a})^{-1}\ a' \quad \dots (18)$$

If we normalize a to a^* by $a^* = \frac{a}{\sqrt{a'a}}$ that is $a^{*\prime}a^*$ =1 then

$$G_1 = G - \lambda\, a^* a^{*\prime} \quad \dots (19)$$

This formula is used to extract the contribution of each successive solution in dual scaling and as such formula (19) is important to remember.
Applying formula (18) to our problem of eliminating the second trivial solution from the equation, we catenulate the residual matrix, indicated by C_1 , noting a in (18) is $D_c^{\frac{1}{2}}$I .as

$$C_1 = B'B - \frac{D_c^{\frac{1}{2}} 1 I\ \ D_c^{\frac{1}{2}}}{f_t} \quad \dots (20)$$

Note that in deriving formula (20) we also used the relations $\eta^2 = 1$ and
$W'\,W = 1 D_c^{\frac{1}{2}}\, D_c^{\frac{1}{2}} 1 = f_t$ The last term of the right – hand side of formula (20) in the current example is given as :

(8,9)

$$C_1 = \begin{pmatrix} 0.1552 & 0.0006 & -0.1742 \\ 0.0006 & 0.0207 & -0.0250 \\ -0.1742 & -0.0250 & 0.2241 \end{pmatrix}$$

Let us use [1 0 − 1] *as an initial* vector b_0 then

$$C_1 b_0 = \begin{pmatrix} 0.3294 \\ 0.0256 \\ -0.3983 \end{pmatrix} = b_1$$

The largest obsolete value in b_1 is 0.3983 . Thus max (b_1) = k_1= 0.3983 and

$$b_1^* = \frac{\mathrm{b}_1}{0.3983} = \begin{pmatrix} 0.8270 \\ 0.0643 \\ -1.000 \end{pmatrix} \quad \text{There fore}$$

[7]Equation (6) the vector b_5^* must be rescaled

$$C_1 b_1^* = \begin{pmatrix} 0.3026 \\ 0.0268 \\ -0.3698 \end{pmatrix} = b_2 \quad , K_2 = 0.3698$$

We repeat …..

$$C_1 b_1^* = b_{j+1} \quad , \quad \frac{b_{j+1}}{|K_{j+1}|} = b_{j+1}^*$$

Until we reach

$$b_j^* = b_{j+1}^*$$

(10)

Since $b_4 = b_5$ the iterative process has converged to the solution and
$K_5 = \eta_1^2 = 0.3683$, That K_5 is the largest Eigen value to obtain the corresponding weight vector that satisfies
Equation (6) states that $X'\,D_C X = f_t$ But $X'\,D_C X = W'W$, and $f_t = 29$ There fore

$$\mathrm{W} = \sqrt{\frac{29}{b_5^{*\prime} b_5^*}} \cdot b_5^* = \begin{pmatrix} 3.4029 \\ 0.3051 \\ -4.1626 \end{pmatrix}$$

[8]The optimal weight vector $\mathcal{X}_1$ is given by:

$$\mathcal{X}_1 = D_c^{-\frac{1}{2}}\mathrm{w} = \begin{pmatrix} \frac{-1}{\sqrt{10}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{11}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{8}} \end{pmatrix} \begin{pmatrix} 3.4029 \\ 0.3051 \\ -4.1626 \end{pmatrix}$$

$$= \begin{pmatrix} 1.0761 \\ 0.0220 \\ -1.4717 \end{pmatrix}$$

(11,12)

rafal.suood.99@gmail.com

Therefore, optimal Weights for categories good, average, and poor are 1.0761, 0.0920, and –104717.

We can start in same way as we have shown, except that we look at y for x and the between-column sum of squares for the between- row sum of squares .The squared correlation ratio η^2 is now defined as ratio of the between- column sum of squares to the total sum of squares

And this ratio is maximized with respect to y under the constraints f' y=0 and y' $D_r y = f_t$

Following the same procedure as before we obtain

$$(BB' - \eta^2 I)\, D_r^{\frac{1}{2}} y = 0 \qquad \dots \quad (21)$$

Let η be the positive square root of η^2 . Then

$$y = \left(\frac{1}{\eta}\right) D_r^{-1}\, F\, \boldsymbol{x} \qquad \dots \quad (22)$$

There are two trivial solution as before, $D_r^{\frac{1}{2}}$y =0 and y=1 .The second trivial vector always yields $\eta^2 = 1$.

So every process is patrolled to the previous problem. Indeed, there exists complete symmetry or duality between the two problems .Thus, rather than repeating the same process let us present an important relation between x and y .

$$\mathcal{X} = \left(\frac{1}{\eta}\right) D_C^{-1}\, F^{-1}\, y \qquad \dots \quad (23)$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \frac{1}{\sqrt{0.3683}} \begin{pmatrix} \frac{1}{10} & 0 & 0 \\ 0 & \frac{1}{10} & 0 \\ 0 & 0 & \frac{1}{9} \end{pmatrix} \begin{pmatrix} 1 & 3 & 6 \\ 3 & 5 & 2 \\ 6 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1.071 \\ 0.0920 \\ -104717 \end{pmatrix}$$

$$= \begin{pmatrix} -1.2322 \\ 0.1227 \\ 1.2326 \end{pmatrix}$$

Conclusion:

We can notice that $\mathcal{X}_1 < \mathcal{X}_2 < \mathcal{X}_3$ and $y_1 < y_2 < y_3$ i.e. we find that the best teach is the third with the best optimal weight for categories …………e.t

From the outset of dual scaling research the mean orientation has been to extend its applicability to a wider variety of categorical data .As arousal data (i.e. the main data types for correspondence analysis) , but also purid comparison rank order , successive categories , sorting and multiway data matrices .

In addition to this wider applicability an analogue of discriminant analysis for categorical data has been developed under the name forced classification and the procedure of generalized forced classification is also available. Other computational procedures were also developed Including the method of successive data modification (S D M) for quantification of ordered categories , partially optimal scaling for data with preassigned weight or mixed (categorical plus continuous) data, and the piecewise method of reciprocal average for handling a large number of multiple- choice items .The introduction and discussion of many of these procedures would certainly be enough for one book- another reason for staying within the familiar territory of dual scaling .

Consequently there may be an excessive number of references to those studies conducted in Toronto.

REFERENCES

Ahn,H.(1991) .Effects of missing responses in multiple choice data on dual scaling . unpublished doctoral dissertation ,University of Toronto ,Ganda .

Benzecri,J.P.et al (1973) .LAnalyse does Donnees .II.LAnalyse des correspondences [data

rafal.suood.99@gmail.com

https://doi.org/10.26682/sjuod.2021.24.1.6

analysis II. Correspondence analysis]. Paris France : Dunod .

Coombs, C.H.(1964).Atheory of data .New York ,Joha Wiley and Jons,Ine

Frank, G.R.(1985)Evaluating measure through data quantification : Applying dual scaling to an advertising copy test .Journal of Business Research, 13,61-69

Halpine,s.(1991) , Phase angle differences in visual evoked potentials : Do they affect estimate of the signal ? unpublished doctor distribution ,university of Toronto –Ganda .

Nishisato ,S.(1986c) .Multi-dimensional analysis of successive categories .InJ.de Leeuw, W.Heiser ,J.Meulman ,& F. Gritchley (Eds), Multi-dimensional data analysis (pp.249-250) .Leiden ,the Nehter lands, DSWO Press .

Rao,c.R. (1965)Linear statistical inference and the application , New York : wiley

Shizuhiko, Nishisatio (1994) Elements of dual sealing “ An introduction t j

parctical data Analysis .

Shizuhiko, NishisatioElements of dual sealing An introduction parctical data Analysis .

Sachs ,J.(1984) .Approximatiay an anderlying continuous variable with dual scaling unpublished masters theis, unversity of Toronto ,Ganda .

Tversky , A. (1964) . on the opimal number of alternative Sat achoice point .

Journal of Mathematical Psychology ,I,(386-93)

Watanabe ,S.(1969) .Knowing and guessing : Aquantitative study of In ference and information . New York :Wiley.

rafal.suood.99@gmail.com