RV Coefficient
and
Congruence Coefficient

Hervé Abdi[1]

1 Overview

The congruence coefficient was first introduced by Burt (1948) under the name of *unadjusted correlation* as a measure of the similarity of two factorial configurations. The name *congruence* coefficient was later tailored by Tucker (1951, see also Harman, 1976). The congruence coefficient is also sometimes called a *monotonicity* coefficient (Borg & Groenen, 1997, p. 203). The congruence coefficient takes values between −1 and +1.

The R_V coefficient was introduced by Escoufier (1973, see also Robert & Escoufier, 1976) as a measure of similarity between squared symmetric matrices (specifically: positive semi-definite matrices) and as a theoretical tool to analyze multivariate techniques. The R_V coefficient is used in several statistical techniques such as STATIS and DISTATIS (see corresponding entries and Abdi, 2003). In order to compare rectangular matrices using the R_V coefficient the first step is to transform them into square matrices. The R_V

[1]In: Neil Salkind (Ed.) (2007). *Encyclopedia of Measurement and Statistics.* Thousand Oaks (CA): Sage.
Address correspondence to: Hervé Abdi
Program in Cognition and Neurosciences, MS: Gr.4.1,
The University of Texas at Dallas,
Richardson, TX 75083–0688, USA
E-mail: herve@utdallas.edu http://www.utd.edu/~herve

coefficient takes values between 0 and +1 (because it is used with positive semi-definite matrices, see Appendix for a proof).

These coefficients are similar to the correlation coefficient and are sometimes called *vector* or *matrix* correlation coefficient. This is a potentially misleading appellation since these coefficients are *not*, correlation coefficients because, contrary to the correlation coefficient, the mean of the observations is not subtracted prior to the computation.

The computational formulas of these coefficients are identical but their usage and theoretical foundations differ. Also, their sampling distributions differ because of the types of matrices with which they are used.

2 Notations and computational formulas

Let **X** be an I by J matrix and **Y** be an I by K matrix. The vec operation transforms a matrix into a vector whose entries are the elements of the matrix. The trace operation applies to square matrices and gives the sum of the diagonal elements.

The congruence coefficient is defined when both matrices have the same number of rows and columns (*i.e.,* $J = K$). These matrices can represent factor loadings (*i.e.,* factors by items) or factor projections (*i.e.,* observations by factors). The congruence coefficient is denoted φ or sometimes r_c, and it can be computed with three different equivalent formulas:

$$\varphi = r_c = \frac{\sum_{i,j} x_{i,j} y_{i,j}}{\sqrt{\left(\sum_{i,j} x_{i,j}^2\right)\left(\sum_{i,j} y_{i,j}^2\right)}} \tag{1}$$

$$= \frac{\mathrm{vec}\{\mathbf{X}\}^\mathsf{T}\,\mathrm{vec}\{\mathbf{Y}\}}{\sqrt{\left(\mathrm{vec}\{\mathbf{X}\}^\mathsf{T}\,\mathrm{vec}\{\mathbf{X}\}\right)\left(\mathrm{vec}\{\mathbf{Y}\}^\mathsf{T}\,\mathrm{vec}\{\mathbf{Y}\}\right)}} \tag{2}$$

$$= \frac{\mathrm{trace}\left\{\mathbf{X}\mathbf{Y}^\mathsf{T}\right\}}{\sqrt{\left(\mathrm{trace}\left\{\mathbf{X}\mathbf{X}^\mathsf{T}\right\}\right)\left(\mathrm{trace}\left\{\mathbf{Y}\mathbf{Y}^\mathsf{T}\right\}\right)}}\,. \tag{3}$$

The R_V coefficient was defined by Escoufier (1973) as a similarity coefficient between positive semi-definite matrices. Escoufier and Robert and Escoufier (1976) pointed out that the R_V coefficient had important mathematical properties because it can be shown that most multivariate analysis techniques amount to maximizing this coefficient with suitable constraints. Recall, at this point, that a matrix **S** is called *positive semi-definite* when it can be obtained as the product of a matrix by its transpose. Formally, we say that **S** is positive semi-definite when there exists a matrix **X** such that:

$$\mathbf{S} = \mathbf{X}\mathbf{X}^\mathsf{T}\,. \tag{4}$$

Note that as a consequence of the definition, positive semi-definite matrices are square, symmetric, and that their diagonal elements are always larger or equal to zero.

If we denote by **S** and **T** two positive semi-definite matrices of same dimensions, the R_V coefficient between them is defined as

$$R_V = \frac{\mathrm{trace}\left\{\mathbf{S}^\mathsf{T}\mathbf{T}\right\}}{\sqrt{\left(\mathrm{trace}\left\{\mathbf{S}^\mathsf{T}\mathbf{S}\right\}\right) \times \left(\mathrm{trace}\left\{\mathbf{T}^\mathsf{T}\mathbf{T}\right\}\right)}}\,. \tag{5}$$

This formula is computationally equivalent to

$$R_V = \frac{\operatorname{vec}\{\mathbf{S}\}^\mathsf{T} \operatorname{vec}\{\mathbf{T}\}}{\sqrt{\left(\operatorname{vec}\{\mathbf{S}\}^\mathsf{T} \operatorname{vec}\{\mathbf{S}\}\right)\left(\operatorname{vec}\{\mathbf{T}\}^\mathsf{T} \operatorname{vec}\{\mathbf{T}\}\right)}} \tag{6}$$

$$= \frac{\sum_i^I \sum_j^I s_{i,j} t_{i,j}}{\sqrt{\left(\sum_i^I \sum_j^I s_{i,j}^2\right)\left(\sum_i^I \sum_j^I t_{i,j}^2\right)}} \,. \tag{7}$$

For rectangular matrices, the first step is to transform the matrices into positive semi-definite matrices by multiplying each matrix by its transpose. So, in order to compute the value of the R_V coefficient between the I by J matrix $\mathbf{X}$ and the I by K matrix $\mathbf{Y}$, the first step it to compute

$$\mathbf{S} = \mathbf{X}\mathbf{X}^\mathsf{T} \text{ and } \mathbf{T} = \mathbf{Y}\mathbf{Y}^\mathsf{T} \,. \tag{8}$$

If we combine Equation 5 and 8, we find that the R_V coefficient between these two rectangular matrices is equal to

$$R_V = \frac{\operatorname{trace}\left\{\mathbf{X}\mathbf{X}^\mathsf{T}\mathbf{Y}\mathbf{Y}^\mathsf{T}\right\}}{\sqrt{\left(\operatorname{trace}\left\{\mathbf{X}\mathbf{X}^\mathsf{T}\mathbf{X}\mathbf{X}^\mathsf{T}\right\}\right) \times \left(\operatorname{trace}\left\{\mathbf{Y}\mathbf{Y}^\mathsf{T}\mathbf{Y}\mathbf{Y}^\mathsf{T}\right\}\right)}} \,. \tag{9}$$

The comparison of Equation 3 and 9 shows that the congruence and the R_V coefficients are equivalent only in the case of positive semi-definite matrices.

From a linear algebra point of view, the numerator of the R_V coefficient corresponds to a scalar product between positive semi-definite matrices, and therefore gives to this set of matrices the structure of a vector space. Within this framework, the denominator of the R_V coefficient is called the *Frobenius*, or *Schur*, or *Hilbert-Schmidt* matrix scalar product (see *e.g.*, Horn & Johnson, 1985, p. 291), and the R_V coefficient is a *cosine* between matrices. This vector space structure is responsible of the mathematical properties of the R_V coefficient.

3 Sampling distributions

The congruence and the R_V coefficients quantify the similarity between two matrices. An obvious practical problem is to be able to perform statistical testing on the value of a given coefficient. In particular it is often important to be able to decide if a value of coefficient could have been obtained by chance alone. To perform such statistical tests, we need to derive the sampling distribution of the coefficient under the null hypothesis (*i.e.,* in order to test if the population coefficient is null). More sophisticated testing requires to derive the sampling distribution for different values of the population parameters. So far, analytical methods have failed to characterize such distributions, but computational approaches have been used with some success. Because the congruence and the R_V coefficients are used with different types of matrices, their sampling distributions are likely to differ and so, work done with each type of coefficient has been carried independently of the other.

3.1 Congruence coefficient

Recognizing that analytical methods were unsuccessful, Korth and Tucker (1976) decided to use Monte Carlo simulations to gain some insights into the sampling distribution of the congruence coefficient. Their work was completed by Broadbooks and Elmore (1987, see also Bedeian, Armenakis & Randolph, 1988). From this work, it seems that the sampling distribution of the congruence coefficient depends upon several parameters including the original factorial structure and the intensity of the population coefficient and therefore no simple picture emerges, but some approximations can be used. In particular, for testing that a congruence coefficient is null in the population, an approximate conservative test is to use Fisher Z-transform and to treat the congruence coefficient like a coefficient of correlation. Broadbooks and Elmore provide tables for population values different from zero. With the availability of fast computers, these tables can easily be extended to accommodate specific cases. The meaningfulness of this index is explored by Lorenzo-Seva and ten Berge (2006).

3.2 R_V coefficient

Statistical approaches for the R_V coefficient have focused on permutation tests. In this framework, the permutations are performed on the entries of each column of the rectangular matrices **X** and **Y** used to create the matrices **S** and **T**. Interestingly, work by Kazi-Aoual *et al.*, (1995, see also Schlich, 1996) has shown that the mean and the variance of the permutation test distribution can be computed directly from **S** and **T**.

The first step is to derive an index of the dimensionality or rank of the matrices. This index denoted $\beta_{\mathbf{S}}$ (for matrix $\mathbf{S} = \mathbf{X}\mathbf{X}^{\mathsf{T}}$) is also known as ν in the brain imaging literature where it is called a *sphericity* index and is used as an estimation of the number of degrees of freedom for multivariate tests of the general linear model (see *e.g.,* Worsley and Friston, 1995). This index depends upon the eigenvalues (see entry) of the **S** matrix denoted ${}_{\mathbf{S}}\lambda_\ell$ and it is defined as:

$$\beta_{\mathbf{S}} = \frac{\left(\sum_{\ell}^{L} {}_{\mathbf{S}}\lambda_\ell\right)^2}{\sum_{\ell}^{L} {}_{\mathbf{S}}\lambda_\ell^2} = \frac{\operatorname{trace}\{\mathbf{S}\}^2}{\operatorname{trace}\{\mathbf{SS}\}} . \tag{10}$$

The mean of the set of permutated coefficients between matrices **S** and **T** is then equal to

$$E\left(R_V\right) = \frac{\sqrt{\beta_{\mathbf{S}}\beta_{\mathbf{T}}}}{I-1} . \tag{11}$$

The case of the variance is more complex and involves computing three preliminary quantities for each matrix. The first quantity is denoted $\delta_{\mathbf{S}}$ (for matrix **S**), it is equal to:

$$\delta_{\mathbf{S}} = \frac{\sum_{i}^{I} s_{i,i}^2}{\sum_{\ell}^{L} {}_{\mathbf{S}}\lambda_\ell^2} \tag{12}$$

The second one is denoted $\alpha_{\mathbf{S}}$ for matrix **S**, and is defined as:

$$\alpha_{\mathbf{S}} = I - 1 - \beta_{\mathbf{S}} , \tag{13}$$

The third one is denoted $C_{\mathbf{S}}$ (for matrix **S**) and is defined as:

$$C_{\mathbf{S}} = \frac{(I-1)\left[I(I+1)\delta_{\mathbf{S}} - (I-1)\left(\beta_{\mathbf{S}}+2\right)\right]}{\alpha_{\mathbf{S}}(I-3)} . \tag{14}$$

With these notations, the variance of the permuted coefficients is obtained as

$$V(R_V) = \alpha_{\mathbf{S}}\alpha_{\mathbf{T}} \times \frac{2I(I-1) + (I-3)\,C_{\mathbf{S}}C_{\mathbf{T}}}{I(I+1)(I-2)(I-1)^3} \tag{15}$$

The sampling distribution of the permutated coefficients is relatively similar to a normal distribution (even though it is, in general, *not* normal) and therefore we can use a Z criterion to perform null hypothesis testing or to compute confidence intervals. For example, the criterion

$$Z_{R_V} = \frac{R_V - E(R_V)}{\sqrt{V(R_V)}} , \tag{16}$$

can be used to test the null hypothesis that the observed value of R_V was due to chance.

3.3 An example

As an example, we will use two scalar product matrices from the STATIS example entry (Experts 1 and 3). These matrices are listed below:

$$\mathbf{S} = \begin{bmatrix} 29.56 & -8.78 & -20.78 & -20.11 & 12.89 & 7.22 \\ -8.78 & 2.89 & 5.89 & 5.56 & -3.44 & -2.11 \\ -20.78 & 5.89 & 14.89 & 14.56 & -9.44 & -5.11 \\ -20.11 & 5.56 & 14.56 & 16.22 & -10.78 & -5.44 \\ 12.89 & -3.44 & -9.44 & -10.78 & 7.22 & 3.56 \\ 7.22 & -2.11 & -5.11 & -5.44 & 3.56 & 1.89 \end{bmatrix} .$$

and

$$\mathbf{T} = \begin{bmatrix} 11.81 & -3.69 & -15.19 & -9.69 & 8.97 & 7.81 \\ -3.69 & 1.81 & 7.31 & 1.81 & -3.53 & -3.69 \\ -15.19 & 7.31 & 34.81 & 9.31 & -16.03 & -20.19 \\ -9.69 & 1.81 & 9.31 & 10.81 & -6.53 & -5.69 \\ 8.97 & -3.53 & -16.03 & -6.53 & 8.14 & 8.97 \\ 7.81 & -3.69 & -20.19 & -5.69 & 8.97 & 12.81 \end{bmatrix} .$$

We find the following value for the R_V coefficient:

$$\begin{aligned} R_V &= \frac{\sum_i^I \sum_j^I s_{i,j} t_{i,j}}{\sqrt{\left(\sum_i^I \sum_j^I s_{i,j}^2\right)\left(\sum_i^I \sum_j^I t_{i,j}^2\right)}} \\ &= \frac{(29.56 \times 11.81) + (-8.78 \times -3.69) + \cdots + (1.89 \times 12.81)}{\sqrt{\left[(29.56)^2 + (-8.78)^2 + \cdots + (1.89)^2\right]\left[(11.81)^2 + (-3.69)^2 + \cdots + (12.81)^2\right]}} \\ &= .79 . \end{aligned} \tag{17}$$

To test the significance of a value of $R_V = .79$, we first compute the following quantities

$$\begin{aligned} &\beta_{\mathbf{S}} = 1.0954 \quad \alpha_{\mathbf{S}} = 3.9046 \quad \delta_{\mathbf{S}} = 0.2951 \quad C_{\mathbf{S}} - 1.3162 \\ &\beta_{\mathbf{T}} = 1.3851 \quad \alpha_{\mathbf{T}} = 3.6149 \quad \delta_{\mathbf{T}} = 0.3666 \quad C_{\mathbf{T}} = -0.7045 \end{aligned} \tag{18}$$

Plugging these values into Equations 11 15, and 16, we find

$$E(R_V) = 0.2464, \quad V(R_V) = 0.0422 \text{ and } Z_{R_V} = 2.66 . \tag{19}$$

Assuming a normal distribution for the Z_{R_V} gives a p value of .0077, which would allow for the rejection of the null hypothesis for the observed value of the R_V coefficient.

Appendix: The RV coefficient takes values between 0 and 1 for positive semi definite matrices

Let **S** and **T** be two positive semi definite matrices. From the Cauchy-Schwartz inequality, we know that the absolute value of the numerator is always smaller or equal to the denominator (and so, R_V is smaller than 1), therefore we only need to prove that the numerator of the R_V coefficient is positive or null. This amounts to showing that

$$\text{trace}\{\mathbf{ST}\} \geq 0. \tag{A1}$$

Because **T** is positive semi definite, the left part of Equation A1 can be rewritten as

$$\text{trace}\left\{\mathbf{T}^{\frac{1}{2}}\mathbf{S}\mathbf{T}^{\frac{1}{2}}\right\} . \tag{A2}$$

Because **S** is positive semi definite, the matrix $\mathbf{T}^{\frac{1}{2}}\mathbf{S}\mathbf{T}^{\frac{1}{2}}$ is also positive semi definite, and therefore all its eigenvalues are positive or null. Thus, its trace being equal to the sum of its eigenvalues is also positive or null and this completes the proof.

References

[1] Abdi, H. (2003b). Multivariate analysis. In M. Lewis-Beck, A. Bryman, & T. Futing (Eds): *Encyclopedia for research methods for the social sciences.* Thousand Oaks (CA): Sage.

[2] Bedeian, A.G., Armenakis, A.A., & Randolph, W.A. (1988). The significance of congruence coefficients: A comment and statistical test. *Journal of Management*, **14**, 559–566.

[3] Borg, I., & Groenen, P., (1997). *Modern multidimensional scaling.* New York: Springer Verlag.

[4] Broadbooks, W.J., & Elmore, P.B. (1987). A monte Carlo study of the sampling distribution of the congruence coefficient. *Educational and Psychological Measurement*, **47**, 1–11.

[5] Burt., C., (1948). factor analysis and canonical correlations. *British Journal of Psychology, Statistical Section,* **1**, 95–106.

[6] Escoufier, Y., (1973). Le traitement des variables vectorielles. *Biometrics,* **29**, 751–760.

[7] Harman, H.H. (1976). *Modern factor analysis (3rd. edition revised)*. Chicago: Chicago University Press.

[8] Horn, R.A., & Johnson, C.R. (1985) *Matrix Analysis.* Cambridge: Cambridge University Press.

[9] Holmes, S. (to appear). Multivariate Analysis: The French Way, *in Festschrift for David Freedman.* Stanford: IMS Lecture Notes.

[10] Kazi-Aoual, F., Hitier S., Sabatier R., & Lebreton J.-D. (1995). Refined approximations to permutation tests for multivariate inference. *Computational Statistics and Data Analysis,* **20**, 643–656.

[11] Korth, B.A., & Tucker, L.R. (1976). The distribution of chance congruence coefficients from simulated data. *Psychometrika,* **40**, 361–372.

[12] Lorenzo-Seva, U., & ten Berge, J.M.F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology,* **2**, 57–64.

[13] Robert, P. & Escoufier, Y. (1976) A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Applied Statistics,* **25**, 257–265.

[14] Schlich, P. (1996). Defining and validating assossor compromises about product distances and attribute correlations. In T, Næs, & E. Risvik (Eds.), *Multivariate analysis of data in sensory sciences.* New York: Elsevier.

[15] Tucker L.R. (1951). *A method for the synthesis of factor analysis studies* (Personnel Research Section Report No. 984). Washington: Department of the Army.

[16] Worsley K.J., & Friston, K.J (1995). Analysis of *f*MRI time-series revisited—Again. *NeuroImage,* **2**, 173-181.