Comparison between Principal Component Analysis and Formal Concept Analysis of Repertory-Grids

Norbert Spangenberg
Zentrum für psychosomatische Medizin
der Justus-Liebig-Universität Gießen, FRG

Karl Erich Wolff
FB Mathematik und Naturwissenschaften
der Fachhochschule Darmstadt
und
Forschungsgruppe Begriffsanalyse
der Technischen Hochschule Darmstadt, FRG

1. Introduction

We give a first comparison between the *Principal Component Analysis PCA* ("one of the oldest and best known techniques of multivariate analysis" cf. JOLLIFFE [6]) and a very young algebraic technique for the visualization of data namely the *Formal Concept Analysis FCA* (WILLE [11]) by applying both methods to *Repertory-Grid-Tests* (SLATER [8]). We demonstrate the efficiency of these methods using the following grid of a 24 years old patient suffering from bulimia nervosa.

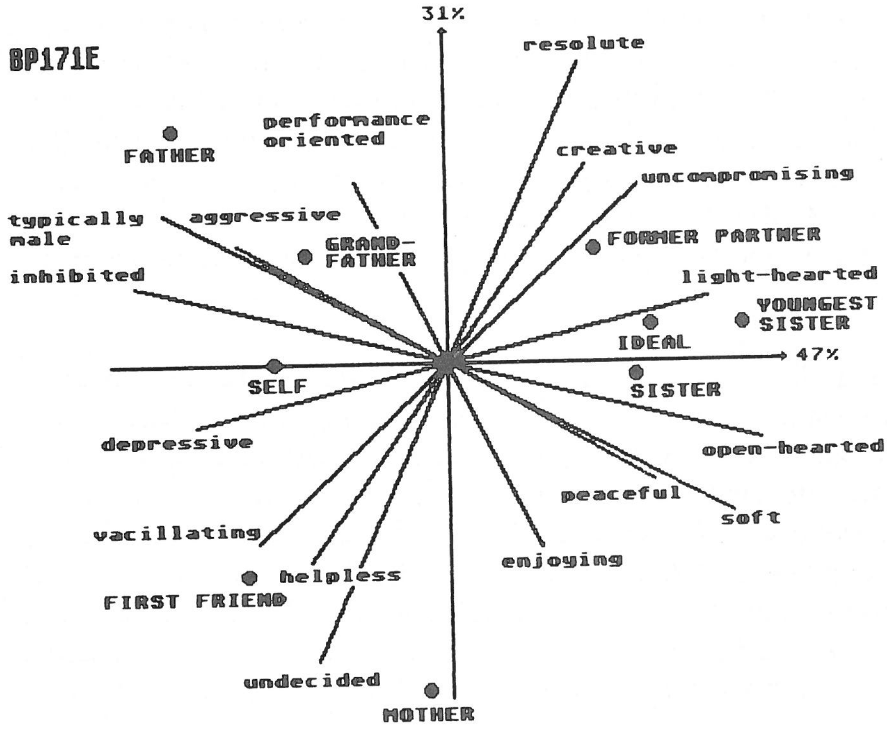
GRID 171E	SE	ID	FP	MO	FA	SI	GF	YS	FF
aggressive - peaceful	2	5	5	4	1	4	1	4	4
vacillating - uncompromising	5	5	6	2	4	5	4	6	2
perform.oriented - enjoying	3	4	1	4	1	5	4	6	6
light-hearted - depressive	4	1	2	5	5	2	4	1	4
typically male - soft	3	5	4	6	1	5	1	6	2
helpless - creative	3	6	4	2	5	5	5	6	2
resolute - undecided	4	2	1	6	1	2	2	1	5
inhibited - open-hearted	2	5	6	6	1	6	4	6	2

In this test the patient judged her most important persons, namely

her SELF (SE), IDEAL (ID), FORMER PARTNER (FP), MOTHER (MO), FATHER (FA), SISTER (SI), GRANDFATHER (GF), YOUNGEST SISTER (YS) and her FIRST FRIEND (FF), using marks from 1 to 6 with respect to self-choosen pairs of constructs.
A reading example: The patient judged her MOTHER with respect to "light-hearted - depressive" with the mark 5, i.e. clearly as "depressive", her IDEAL as "very light-hearted".

2. Principal Component Analysis and Biplots

In the evaluation of grids the "reduction of dimensionality" is usually done by the construction of certain kinds of "biplots" (introduced by GABRIEL [1,2], cf. also JOLLIFFE [6], SLATER [8]).
In order to give a clear comparison between the biplots of PCA and the line-diagrams of FCA we shortly describe the construction of the following "SLATER-biplot" of our grid.



This biplot was constructed from the grid by the following procedure:
(1) The grid is mathematically described as a *real* $n\times p$-matrix.
(2) Subtraction of the row means results in the $n\times p$-dispersion matrix X (of deviations from construct means).
(3) Via singular value decomposition (see GOLUB-REINSCH [5], GABRIEL [1,2], JOLLIFFE [6]) the dispersion matrix is factorized as

(3.1) $X = ULA'$,

where U and A are columnorthonormalized $n\times r$- resp. $p\times r$-matrices, r is the rank of X, A' is the transposed matrix of A and L is a $r\times r$-diagonal matrix, whose diagonalelements $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ are the positive singular values of XX', which are the square roots of the r positive eigenvalues of XX' (counted with their multiplicity).
(4) From the singular value decomposition we obtain

(4.1) $XX' = ULA'ALU' = (UL)(UL)'$ and

(4.2) $X'X = ALU'ULA' = (AL)(AL)'$.

By (4.1) [resp.(4.2)] the rows Q_i of UL [resp. P_j of AL] have the same inner products as the corresponding rows [columns] of X. From UL = XA [resp. AL = X'U] we obtain, that $Q_{ik} = (XA)_{ik}$ [$P_{jk} = (X'U)_{jk}$] is the coordinate of the projection of the i-th row [j-th column] of X onto the (normalized) k-th column a_k of A [u_k of U]: The Q_i's and P_j's represent as coordinate vectors the rows [columns] of X "in the right lengths, angles and euclidean distances".
(5) If r = 2 the rows of UL [AL] represent the rows (pairs of constructs) [columns (persons)] of X in the plane $\mathbb{R}^2$.
If r > 2 the matrix X = ULA' is approximated by it's rank-2 HOUSEHOLDER-YOUNG-approximation

$$X_{[2]} = U_{(2)}L_2 A_{(2)}' = u_1\sigma_1 a_1' + u_2\sigma_2 a_2'$$

(where $U_{(2)}$ [$A_{(2)}$] is the matrix of the first two columns of U [A], $L_2 = \mathrm{diag}(\sigma_1,\sigma_2)$). The i-th row of $U_{(2)}L_2 = X_{[2]}A_{(2)}$ [resp. j-th row of $A_{(2)}L_2 = X_{[2]}'U_{(2)}$], namely $Q_{i[2]} = (u_{i1}\sigma_1, u_{i2}\sigma_2)$ [resp. $P_{j[2]} = (a_{j1}\sigma_1, a_{j2}\sigma_2)$] contains just the first two coordinates of Q_i [resp.P_j] and represents therefore the i-th row [j-th column] of X in that plane, which is spanned by u_1, u_2 [resp. a_1, a_2]. In the biplot above the $Q_{i[2]}$'s and the $-Q_{i[2]}$'s (representing pairs of constructs) are drawn as line segments, the $P_{j[2]}$'s as points.

The following observation, that the inner products between the Q_i's and the P_j's are given by $(UL)(AL)' = UL^2A'$, which is different from $ULA' = X$, leads us to the biplots introduced by GABRIEL [1], who uses factorizations of $X = ULA' = GH'$ with suitable matrices G, H. For $G = UL$, $H = A$ the GABRIEL-biplot contains the vectors $g_i = Q_{i[2]}$ and the "points" $h_j = (a_{j1}, a_{j2}) = P_{j[2]} L_2^{-1}$. $GH' = X$ explains the "projection rule": $g_i h_j' \approx x_{ij}$ for the interpretation of biplots. Finally we mention, that the k-th axis of the biplot is usually labeled with the quotient $\sigma_k^2 / (\sigma_1^2 + \ldots + \sigma_r^2)$, $(k = 1,2)$, given in percent. From this information we can easily obtain the slope (a_{j2}/a_{j1}) of the vector h_j from the slope of the vector $P_{j[2]}$ by multiplication with $(\sigma_1/\sigma_2) = (47/31)^{1/2}$, hence this GABRIEL-biplot is very similar to the SLATER-biplot.

3. Formal Concept Analysis applied to grids

The application of Formal Concept Analysis to grids was explained by the authors in [9,10]. Therefore we restate only the main ideas of this application: From a "many-valued context", e.g. the grid, we generate a table of crosses, formally a *context* $\mathbb{K}$. The hierarchy of *concepts* of $\mathbb{K}$ can be represented in *line-diagrams* of the *concept lattice* of $\mathbb{K}$, such that the *context* $\mathbb{K}$ *can be reconstructed* from the line-diagram. At first we demonstrate how to *condense* the information of the grid by a suitable scaling: We scale the complete grid with the "threshold-scale 2-5", so that for example the first line

	SE	ID	FP	MO	FA	SI	GF	YS	FF
aggressive - peaceful	2	5	5	4	1	4	1	4	4

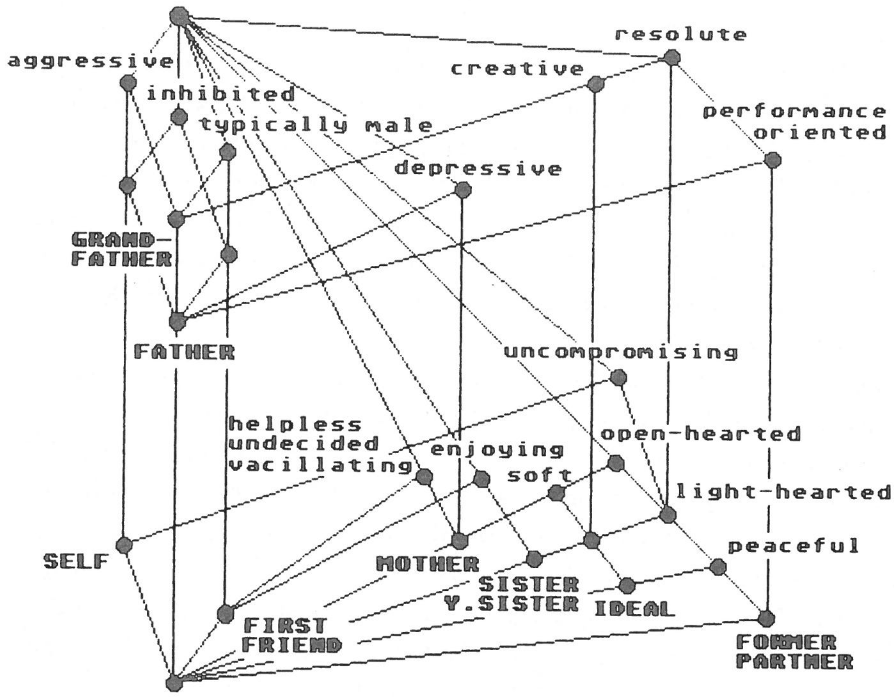
of the grid splits into two lines:

	SE	ID	FP	MO	FA	SI	GF	YS	FF
aggressive	X				X		X		
peaceful		X	X						

Reading example: The IDEAL has the attribute "peaceful", since it has a mark 5 or 6 in the line "aggressive-peaceful".

This leads to a context with 9 objects and 16 attributes (= constructs) with the following line-diagram:

171E2V



How to read such a diagram?
An object g (e.g. the IDEAL) has the attribute m (e.g. resolute) in the given context if and only if there is an upwards leading path from the point named g to the point named m. Examples: The "aggressive" persons of this context are SELF, FATHER and GRANDFATHER, since they are the only ones with marks 1 or 2 in the line "aggressive-peaceful" of the grid. Each "light-hearted" person is "open-hearted", "uncompromising" and "resolute". The only person, which is "helpless" and "typically male" is the FIRST FRIEND. In this context the SISTER and the Y.SISTER have exactly the same attributes. Let's compare this line-diagram and the biplot above:
Both of them show the same partition of persons, indeed an extent-partition (cf. WILLE [12]), consisting of the aggressive, the helpless (undecided, vacillating) and the light-hearted persons. From the line-diagram we see, that the SELF is uncompromising as well as all the light-hearted persons, in contrast to the biplot-im-

pression : "The SELF is vacillating", resulting from the projection rule in the "weak form" :
"$Q_{i[2]}P_{j[2]}' \gg 0 \Rightarrow$ the 2. part of construct i applies to person j",
"$Q_{i[2]}P_{j[2]}' \ll 0 \Rightarrow$ the 1. part of construct i applies to person j",
which is mainly used (since often fulfilled) in the SLATER-biplots. The projection rule also suggests a large difference in the values for MOTHER and GRANDFATHER with respect to "performance oriented - enjoying", but these values coincide. The length of the vector "Y. SISTER" indicates her highly extreme values (which are at each construct more extreme than those of the SISTER), a fact not recognizable from this line-diagram.

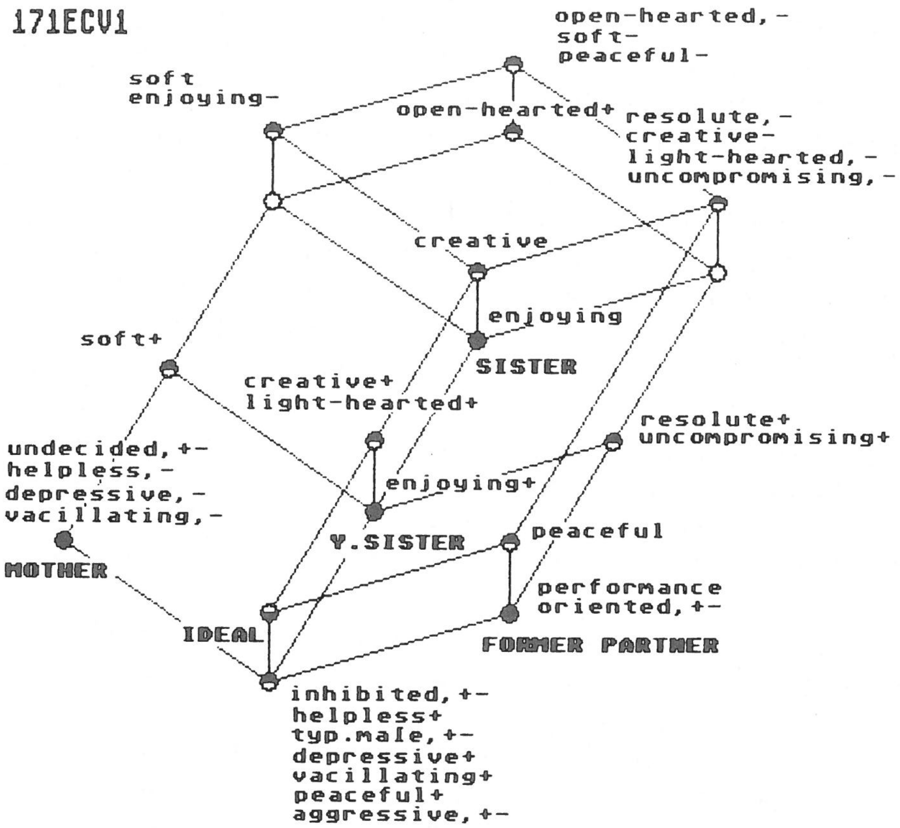
If we wish to see the whole information of the grid, we choose the biordinal scale O_{3+3} (cf. GANTER, WILLE [4]) to express the "biordinal meaning" of the marks 1 to 6 for a pair of opposite constructs. This leads to a line-diagram with 55 concepts, which we don't show here. Instead we look "through a magnifier" at the five "open-hearted" persons of the line-diagram 171E2V and generate from their 8×5-subgrid using the biordinal scale O_{3+3} a line-diagram, which contains the whole information of this subgrid. The first line

	ID	FP	MO	SI	YS
aggressive - peaceful	5	5	4	4	4

of this "open-hearted" 8×5-subgrid generates the following first six lines of a context with 8×6 rows and 5 columns:

	ID	FP	MO	SI	YS
aggressive+ := aggressive - peaceful 1					
aggressive := aggressive - peaceful 2					
aggressive- := aggressive - peaceful 3					
peaceful- := aggressive - peaceful 4	X	X	X	X	X
peaceful := aggressive - peaceful 5	X	X			
peaceful+ := aggressive - peaceful 6					

Reading example: The IDEAL has mark 5, hence we say "The IDEAL is peaceful and weak peaceful (= peaceful-)", since we wish to express the order "peaceful+ ⇒ peaceful ⇒ peaceful-". The following line-diagram of (the dual of) this 48×5-context shows, that each attribute of the SISTER applies also to the Y.SISTER, e.g. the SISTER is "creative" as the Y.SISTER, who is even "creative+".



4. Conclusion

1. *Reduction*: a) In the construction of biplots the "reduction of dimensionality" via a projection from $\mathbb{R}^r$ to $\mathbb{R}^2$ causes the misleading biplot-information demonstrated for example in the biplot above. b) In contrast to the biplots (and many other methods of multivariate analysis) the line-diagrams of FCA reach the aim of an *injective* representation, which permits to reproduce the original data exactly. To condense the information of large data sets in a self-choosen and interpretable way one can use suitable scales (e.g. nominal, ordinal, biordinal, interordinal and threshold-scales) (cf. GANTER, WILLE [4], WILLE [12,13]).

2. *Scaling*: The most important difference between both methods with respect to measurement theory (cf. ROBERTS [7]) lies in the scaling of the data by a numerical metric scale in PCA and by discrete scales in FCA (cf. GANTER, STAHL, WILLE [3], GANTER, WILLE [4]).

5. Literature

[1] K.R.Gabriel: The biplot graphic display of matrices with application to principal component analysis. Biometrica 58, (1971), 453-467.

[2] K.R.Gabriel: Biplot display of multivariate matrices for inspection of data and diagnosis. In: V.Barnett (ed.): Interpreting multivariate data. Wiley, Chichester 1981, 147-173.

[3] B.Ganter, J.Stahl, R.Wille: Conceptual measurement and many-valued contexts. In: W.Gaul, M.Schader (eds.): Classification as a tool of research. North-Holland, Amsterdam 1986, 169-176.

[4] B.Ganter, R.Wille: Conceptual Scaling. FB4-Preprint Nr. 1174, TH Darmstadt 1988.

[5] G.H.Golub, C.Reinsch: Singular Value Decomposition and Least Squares Solutions. In: J.H.Wilkinson, C.Reinsch: Linear Algebra, Springer Verlag, Berlin, Heidelberg, New York 1971, 134-151.

[6] I.T.Jolliffe: Principal Component Analysis. Springer-Verlag, New York 1986.

[7] F.S.Roberts: Measurement theory. Addison-Wesley, Reading 1979.

[8] P.Slater: The measurement of interpersonal space by Grid Technique. Vol. I and II, Wiley, New York, 1977.

[9] N.Spangenberg, K.E.Wolff: Conceptual grid evaluation. In: H.H.Bock (ed.): Classification and related methods of data analysis. Proceedings of the First Conference of the International Federation of Classification Societies, Technical University of Aachen/FRG, 1987, North Holland, Amsterdam 1988, 577-580.

[10] N.Spangenberg, K.E.Wolff: Formal Concept Analysis of Repertory Grids: A case study of a patient with Anorexia nervosa. To appear in the Proceedings of the 17h European Conference on Psychosomatic Research, Marburg, 1988.

[11] R.Wille: Restructuring lattice theory: an approach based on hierarchies of concepts. In: Ordered Sets (ed. I.Rival). Reidel, Dordrecht-Boston 1982, 445-470.

[12] R.Wille: Bedeutungen von Begriffsverbänden. In: B.Ganter, R.Wille, K.E.Wolff (Hrsg.): Beiträge zur Begriffsanalyse. B.I.-Wissenschaftsverlag, Mannheim/Wien/Zürich 1987, 161-211.

[13] R.Wille: Lattices in data analysis: How to draw them with a computer. In: I.Rival (ed.): Algorithms and order. Kluver Academic Publishers 1989, 33-58.