DOCUMENT RESUME

ED 367 711 TM 021 242

AUTHOR Blankmeyer, Eric
TITLE Principal Components and Scale Dependence.
PUB DATE 94
NOTE 9p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Change; *Factor Analysis; Foreign Countries; Higher Education; Medical Students; *Regression (Statistics); *Research Methodology; Research Problems; *Scaling; Scores; Statistical Analysis; *Statistical Studies
IDENTIFIERS Invariance; Netherlands; *Renormalization; *Scale Dependence

ABSTRACT

A limitation of the principal components method is its scale dependence. This note shows that the method is scale invariant if the normalization is modified in an obvious way. Then the effect of a change in units is as transparent as in linear regression, and principal components can be used without apology. Most researchers who use multivariate methods take for granted the invariance property of linear regression, and they can now feel the same assurance about principal components. It is not likely that the researcher would ever need to compute the renormalization proposed. The adjustment is hypothetical, but it demonstrates that the choice of units is inconsequential. (Contains eight references.) (SLD)

TM

Principal Components and Scale Dependence

Eric Blankmeyer

Department of Finance and Economics

Southwest Texas State University

San Marcos, TX 78666

Tel. 512-245-2547

Abstract. A limitation of the principal component method is its scale dependence. This note shows that the method is scale invariant if the normalization is modified in an obvious way.

Principal Components and Scale Dependence

Introduction

The principal component method is an important analytical technique. It computes scaled scores from multivariate data (Morrison 1990, chapter 8) and achieves parsimony in linear regression (Johnston 1984, pp. 481-482, 542-544). As orthogonal regression, the method is applicable when the independent variables are subject to random errors of measurement (Johnston 1984, pp. 428-435; Malinvaud 1980, chapter 10). However, there are concerns about the scale dependence of principal components:

"This dependence on the unit of measurement is obviously a weakness of the principal component technique....If a variable is measured in such small units that its numerical values dominate those of the other...variables, the first principal component will reflect the behavior of this particular variable rather closely...." (Theil 1971, p. 55).

"Clearly, principal components are not invariant under linear transformation, including separate scaling, of the original coordinates. Thus the principal components of the covariance matrix are not the same as those of the correlation matrix or of some other scaling according to measures of 'importance'..... There does not seem to be any general elementary rationale to motivate the choice of scaling of the variables as a preliminary to principal components analysis on the resulting covariance matrix" (Gnanadesikan 1977, pp. 11-12).

"Changing the unit of measurement of some variables will in general lead to different values of those log-likelihood ratio statistics that are associated with the [common principal

component] model...." (Flury 1988, p. 158).

"The components obtained from [the covariance matrix and the correlation matrix] are not in general the same, nor is it possible to pass from one solution to another by a simple scaling of the coefficients....Furthermore...the sampling theory of components extracted from correlation matrices is exceedingly more complex than that of covariance-matrix components" (Morrison 1990, p. 314).

"A tacit assumption when using covariance input is that the variables should not have grossly different variances. If they do...then the first few principal components will be pulled toward those variables with the larger variances" (Dillon and Goldstein 1984, pp. 33-35).

However, the situation is more favorable than these comments suggest. Malinvaud (1980, pp. 39-42) shows how a linear transformation of the variables changes the principal components. In this note, Malinvaud's theorem is reformulated to address the issue of scale dependence. We argue that the principal component method has in fact a kind of invariance when the normalization is modified in an obvious way. Then the effect of a change in units is as transparent as in linear regression, and principal components can be used without apology.

A renormalization

Suppose that N joint observations on K variables are arrayed in a NxK matrix $\underline{X}$. If each variable is measured in deviations from its sample mean, then the covariance matrix is $\underline{S} = (1/N)\underline{X}'\underline{X}$. Concretely, we consider the following hypothetical data matrix $\underline{X}$, where N = 15 and K = 3:

Table 1. A Data Matrix

Variables		
1	2	3
.9260	.5004	.4169
1.6620	1.9718	1.4486
-.3018	-1.0531	-.4306
-.2614	.7813	-.0460
1.3143	1.7030	-1.4224
.1500	-.0766	-.2453
-.8937	-1.0059	-.6061
.8243	1.1856	.1651
-1.2773	-1.9401	-1.5147
-.0776	.6863	1.4251
-1.5846	-1.5090	-.9391
1.3235	.1756	.6682
-.6865	-.7981	-.6070
-1.4029	-1.0746	.1414
.2855	.4534	1.5459

We have $\underline{S}$ =

1.0286	1.0127	.4375
1.0127	1.3148	.5579
.4375	.5579	.8913

The first principal component of $\underline{S}$ is $\underline{b}$, a column vector of K elements. As usual, we choose $\underline{b}$ and the Lagrange multiplier L to maximize

$$\underline{b}'\underline{S}\underline{b} - L(\underline{b}'\underline{b} - 1) \ . \qquad (1)$$

The first-order conditions for a maximum are

$$(\underline{S} - L\underline{I})\underline{b} = \underline{0} \ , \qquad (2)$$

where $\underline{I}$ is the identity matrix of order K. A nontrivial solution exists if and only if the determinant

$$\det(\underline{S} - L\underline{I}) = 0 \ . \qquad (3)$$

The largest root of (3) is of course the largest eigenvalue of $\underline{S}$. When this value is substituted into (2), $\underline{b}$ is found to be the corresponding eigenvector. For our hypothetical data, equations (2) and (3) are satisfied by L = 2.5052 and $\underline{b}$ = (.5958, .6954, .4018)'. From these results one can also compute 15 scaled "scores" $\underline{Xb}$.

To investigate scale dependence, let us suppose that each observation on the first variable is multiplied by a positive constant c; that is, the first column of $\underline{X}$ is multiplied by c. As a result, the first row and the first column of $\underline{S}$ are multiplied by c. We have emphasized that this change of units alters L and $\underline{b}$ in a complicated way.

For example, if the scale factor c = 10, then the covariance matrix of our hypothetical data becomes

102.8566	10.1270	4.3755
10.1270	1.3148	.5579
4.3755	.5579	.8913

Its largest eigenvalue equals 104.0486, and the corresponding eigenvector is (.9942, .0982, .0427)'. This eigenvalue and its eigenvector are not related in a simple way to the values obtained before the first variable was rescaled. The change of units makes the rescaled variable dominate the first principal component. The 15 scores $\underline{Xb}$ likewise undergo a complicated transformation.

Suppose, however, that the normalization is also modified. In

$\underline{b}'\underline{b} = 1$, the first term is to be replaced by $(cb_1)^2$. The situation is now that the first row and the first column of $(\underline{S} - L\underline{I})$ have been multiplied by c. According to a well-known theorem on determinants (Morrison 1990, p. 45), (3) becomes

$$c^2 \det(\underline{S} - L\underline{I}) = 0 \ . \qquad (4)$$

A comparison of (3) and (4) shows that the largest eigenvalue of $\underline{S}$ is a solution to both equations. The change of units and the renormalization leave the eigenvalue unaltered.

How is the eigenvector affected by these adjustments ? We have just remarked that $(\underline{S} - L\underline{I})$ is multiplied by c in its first row and column. In equations (2), let us divide the first equation by c. The effect of the change of units and the renormalization is then simple: the first column of $(\underline{S} - L\underline{I})$ has been multiplied by c. That column contains the coefficients of b_1. It follows that the new eigenvector is identical to the old eigenvector except that b_1 is replaced by b_1/c. Moreover, the N scores $\underline{Xb}$ are unaltered by the change in units.

We return to our example, where the first variable has been multiplied by 10 and the normalization is now $(10b_1)^2 + b_2^2 + b_3^2 = 1$. Dividing the first equation in (2) by 10, we have

$$\begin{aligned} (10.2857 - 10L)b_1 + 1.0127b_2 + .4375b_3 &= 0 \\ 10.1270b_1 + (1.3148 - L)b_2 + .5579b_3 &= 0 \\ 4.3755b_1 + .5579b_2 + (.8913 - L)b_3 &= 0 \end{aligned}$$

These equations have the solution $\underline{b} = (.0596, .6954, .4018)'$. Compared to the original solution, b_1 has been divided by 10, while b_2, b_3 and the scores $\underline{Xb}$ are unchanged.

Conclusions

This result suggests four comments. First, the proposed renormalization should not be disconcerting. After all, the constraint $\underline{b}'\underline{b} = 1$ is imposed merely to avoid the trivial solution $\underline{b} = \underline{0}$. Nothing of substance is involved in the choice of a normalization; it is a matter of convenience. Writing on the multidimensional analysis of preferences, Srinivasan and Shocker (1973, p. 341) remark, "It appears more rational to believe that the different dimensions have to be differentially weighted. The weights take account both of the units in which each dimension is scaled (scale factors) and the relative importance (or salience) of each attribute....Normally the units in which each attribute is measured will not be identical....Scale factors permit such different units to be meaningfully combined into a single preference measure."

Second, our proposal makes principal components behave exactly like linear regression under a change in the units of an independent variable; that is, b_1 becomes b_1/c, and no other coefficient is affected (Morrison 1990, pp. 96-97). Most researchers who use multivariate methods take for granted this invariance property of linear regression; they can now feel the same assurance about principal components.

Third, our argument also applies to the K-1 smaller principal components as long as cb_1 replaces b_1 everywhere --in the orthogonality constraints as well as in the normalization. Since our procedure leaves all the eigenvalues unchanged, their sum equals the trace of the original covariance matrix rather than the trace of the covariance matrix after a change of units.

Fourth, it is unlikely that a researcher would ever need to compute the renormalization we propose. The adjustment is hypothetical. It merely demonstrates that the choice of units is inconsequential, as indeed it should be. In particular, suppose that principal components are extracted from a correlation matrix under the constraint $\Sigma(b_i^2/s_{ii}) = 1$, where s_{ii} is the variance of variable i. Then the resulting scores $\underline{Xb}$ are identical to the scores obtained from the covariance matrix when $\Sigma b_i^2 = 1$.

References

Dillon, W. and Goldstein, M. 1984. Multivariate Analysis: Methods and Applications. New York: Wiley.

Flury, B. 1988. Common Principal Components and Related Multivariate Models. New York: Wiley.

Gnanadesikan, R. 1977. Methods for Statistical Data Analysis of Multivariate Observations. New York: Wiley.

Johnston, J. 1984. Econometric Methods. New York: McGraw-Hill.

Malinvaud, E. 1980. Statistical Methods of Econometrics. Amsterdam: North-Holland.

Morrison, D. 1990. Multivariate Statistical Methods. New York: McGraw-Hill.

Srinivasan, V. and Shocker, A. 1973. "Linear programming techniques for multidimensional analysis of preferences," Psychometrika, vol. 38, no. 3, pp. 337-369.

Theil, H. 1971. Principles of Econometrics. New York: Wiley.