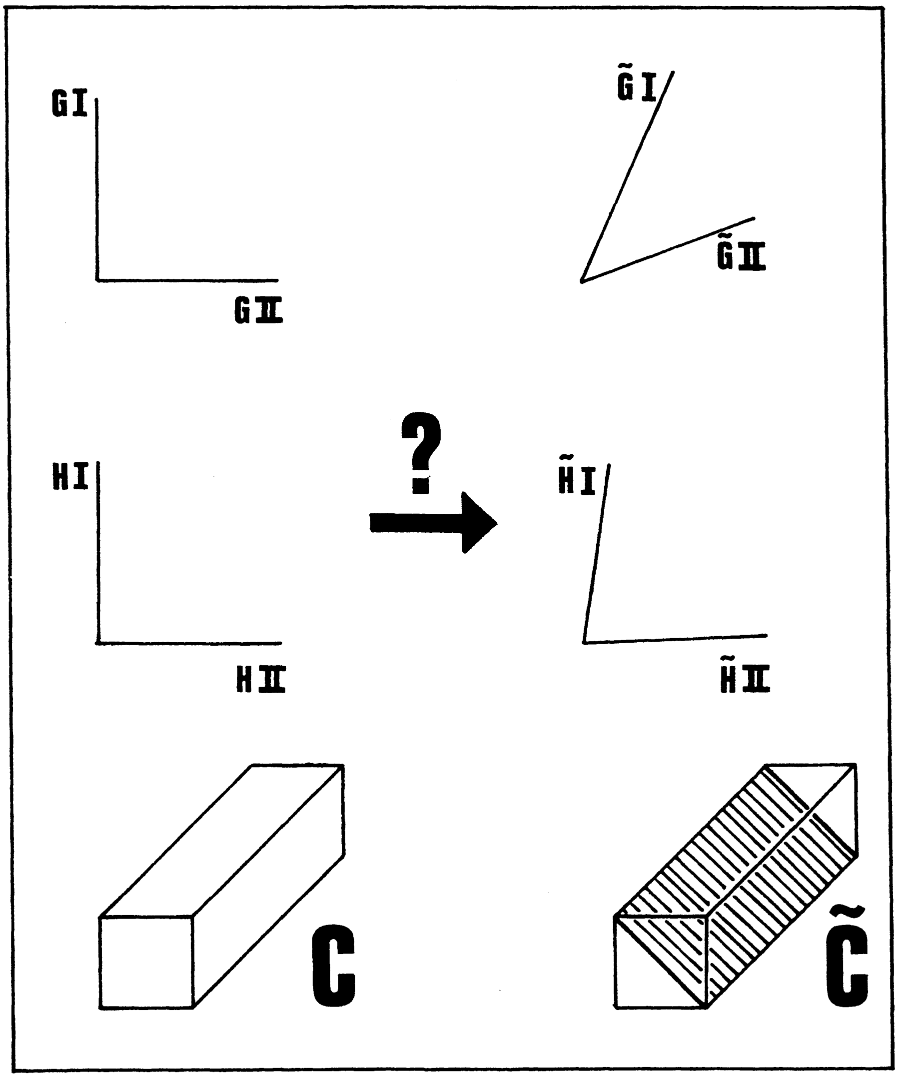
TRANSFORMATIONS OF CORE MATRICES

5



5.1 INTRODUCTION

In this chapter we will discuss the problem of diagonality of the extended core matrix of the Tucker2 model. As the subject is still being explored a definitive treatment cannot yet be given. In section 3.2 we indicated that the orthonormal restrictions put on the Tucker2 and Tucker3 model were necessary to identify the equations for solving the minimization problem, but that these restrictions could be made without loss of generality. In fact, we use the transformational freedom to choose such component matrices that a convenient and efficient algorithm could be devised. Once a solution is found we can drop the restrictions, and use any (non-singular) transformation we like on the components or core matrix, provided the appropriate counter-rotations are performed as well.

One way to use the transformational freedom in three-mode principal component models is to search for transformations that create "simple structures" in the component matrices. Another approach is to search for transformations that induce a simple structure into the core matrix. In this chapter we will look at transformation procedures which attempt the latter by searching for those transformations which diagonalize the core matrix. By doing this we will restrict ourselves to situations where the number of components is equal for the first and second mode (i.e. s=t). This is in itself no restriction as the procedures to be described are such that s and t will become equal in any event.

The procedures outlined in this chapter can be applied to the frontal core planes from both Tucker2 (T2) and Tucker3 (T3) models. In section 3.4 we used the idea of transforming the T3 core matrix to diagonality to show that in specific cases the Tucker3 model has the same form as the PARAFAC1/CANDECOMP model. In this chapter we will not treat these and other issues connected with the Tucker3 model, but leave them for another study. We will concentrate on the diagonality of the extended core matrix of the Tucker2 model.

We will discuss two transformation procedures to optimize diagonality: one using orthonormal transformations, and the other using non-singular transformations. One of the purposes of the transformations to diagonality is to investigate whether the data fit more restricted models such as INDSCAL, which *require* the extended core matrix to be diagonal. We will, however, not present detailed comparisons of solutions from TUCKALS2 plus transformations, and solutions from INDSCAL, as for instance MacCallum (1976a,b) did. In section 11.2 such comparisons are made for data from the *Cola study* (Schiffman, Reynolds, & Young, 1981), but the example is not very demanding.

Another related aim of the transformations is to simplify the interpretation, as a diagonal core matrix with its zero off-diagonal elements displays a typical 'simple structure'. In the case of non-singular transformations, this simplicity is bought by the non-orthogonality of the components in the first and second mode. At present, it is not clear which of the two properties is more desirable in specific situations, the more so because especially the non-singular transformation still poses a number of unsolved interpretational problems.

An extended three-mode core matrix $C = (C_1,\ldots,C_n)$* is defined to be *diagonal* if for each k (k=1,..,n) $c_{pqk} = 0$ for all $p \neq q$. Note that we do not require c_{ppk} to be unequal to zero, but not all c_{ppk} for all k and fixed p may be zero at the same time, because then we would have ended up with one component less in the first and second mode. From now on we will assume that for each p there is always a k, for which c_{ppk} is not zero.

The procedures outlined in section 5.2 and 5.3, and compared in section 5.4, will be applied to two examples in section 5.5.

The discussion of the theory and application of the transformations will be rather incomplete, primarily because the experience with these methods is still very limited.

---

* For convenience, we will write in this chapter C instead of $\tilde{C}$ for the extended core matrix.

5.2 ORTHONORMAL TRANSFORMATIONS

Problem and solution. In this section we will outline a procedure to transform an extended core matrix to optimal diagonality - in a least squares sense - by using two orthonormal transformations. The procedure was first presented in Kroonenberg & De Leeuw (1977, Appendix One).

Diagonality problem (ON)

Let $C = (C_1, C_2, \ldots, C_n)$ with $C_k \varepsilon R^{s\times s}$ $(k=1,\ldots,n)$ be given. Find the $(s\times s)$ orthonormal matrices K and L and $D = (D_1, D_2, \ldots, D_n)$ with D_k diagonal $(k=1,\ldots,n)$, such that

$$\sigma(K,L,D) = \sum_{k=1}^{n} \text{tr } (D_k - KC_kL')'(D_k - KC_kL') \tag{5.1}$$

is as small as possible.

Theorem 5.1

Let $C = (C_1, C_2, \ldots, C_n)$ with $C_k \varepsilon R^{s\times s}$ $(k=1,\ldots,n)$, and the problem ON be given. Then

$$\hat{K} = \hat{U}(\hat{U}'\hat{U})^{-\frac{1}{2}} \text{ with } \hat{U} = \sum_{k=1}^{n} \hat{D}_k \hat{L} C_k',$$

$$\hat{L} = \hat{V}(\hat{V}'\hat{V})^{-\frac{1}{2}} \text{ with } \hat{V} = \sum_{k=1}^{n} \hat{D}_k \hat{K} C_k, \text{ and}$$

$$\hat{D}_k = \text{diag } (\hat{K}C_k\hat{L}') \quad (k=1,\ldots,n)$$

solve the diagonality problem ON.

Proof:

The solution of problem ON is equivalent to the minimization of

$$\tilde{\sigma}(K,L,D,M,N) = \sum_{k=1}^{n} \text{tr } (D_k - KC_kL')'(D_k - KC_kL') -$$

$$-\tfrac{1}{2} \text{ tr } M(K'K - I_s) - \tfrac{1}{2} \text{ tr } N(L'L - I_s) \tag{5.2}$$

where M and N are symmetric matrices of Lagrange multipliers. Let σ be rewritten in the two following ways

$$1.\ \sigma(K,L,D) = \sum_{k=1}^{n} \text{tr } D_k'D_k - 2\sum_{k=1}^{n} \text{tr } K'D_kLC_k' + \sum_{k=1}^{n} \text{tr } C_k'C_k$$

$$= \sum_{k=1}^{n} \text{tr } D_k'D_k - 2 \text{ tr } K' \sum_{k=1}^{n} D_kLC_k' + \sum_{k=1}^{n} \text{tr } C_k'C_k$$

$$= \sum_{k=1}^{n} \text{tr } D_k'D_k - 2 \text{ tr } K'U + \sum_{k=1}^{n} \text{tr } C_k'C_k$$

$$2.\ \sigma(K,L,D) = \sum_{k=1}^{n} \text{tr } D_k'D_k - 2 \text{ tr } L' \sum_{k=1}^{n} D_kKC_k + \sum_{k=1}^{n} \text{tr } C_k'C_k$$

$$= \sum_{k=1}^{n} \text{tr } D_k'D_k - 2 \text{ tr } L'V + \sum_{k=1}^{n} \text{tr } C_k'C_k.$$

Substituting these expressions succesively into (5.2) and differentiating $\tilde{\sigma}$ with respect to K, M, L, N, D, and equating these partial derivatives to zero, we obtain the following set of equation from the stationary equations:

$$\hat{U} = \hat{K}\hat{M} \text{ and } K'K = I_s \tag{5.3}$$
$$\hat{K}'\hat{K} = I_s \tag{5.4}$$
$$\hat{V} = \hat{L}\hat{N} \tag{5.5}$$
$$\hat{L}'\hat{L} = I_s \tag{5.6}$$
$$\hat{D}_k = \text{diag } (\hat{K}C_k\hat{L}') \quad (k=1,\ldots,n) \tag{5.7}$$

Premultiplying (5.3) with its transpose, and using (5.4) we get

$$\hat{U}'\hat{U} = \hat{M}\hat{K}'\hat{K}\hat{M} = \hat{M}'\hat{M} = \hat{M}^2 \implies \hat{M} = (\hat{U}'\hat{U})^{\frac{1}{2}} \tag{5.8}$$

Substituting (5.8) into (5.3), and isolating $\hat{K}$:

$$\hat{K} = \hat{U}(\hat{U}'\hat{U})^{-\frac{1}{2}}\ .$$

Analoguously

$$\hat{L} = \hat{V}(\hat{V}'\hat{V})^{-\frac{1}{2}}.$$

Substituting these $\hat{K}$ and $\hat{L}$ into (5.7) is sufficient to find $\hat{D}_k$ for k=1,...,n. ▼▼

Algorithm. From the above theorem a computational procedure can easily be derived, resulting in an alternating least squares algorithm similar to the TUCKALS2 algorithm. A main iteration step of the ON-algorithm will be defined as:

K substep

$$U_a = \sum_{k=1}^{n} D_k^{(a)} L_a C_k'$$

$$K_{a+1} = U_a (U_a' U_a)^{-\frac{1}{2}}$$

L substep

$$V_a = \sum_{k=1}^{n} D_k^{(a)} K_{a+1} C_k$$

$$L_{a+1} = V_a (V_a' V_a)^{-\frac{1}{2}}$$

D substep

$$D_k^{(a+1)} = \text{diag}\left[K_{a+1} C_k L_{a+1}'\right] \quad (k=1,\ldots,n)$$

Both $(U_a'U_a)^{-\frac{1}{2}}$ and $(V_a'V_a)^{-\frac{1}{2}}$ can be computed in the same manner as in the equivalent expression in the TUCKALS2 algorithm, i.e. by solving the eigenproblem of $U_a'U_a$ and $V_a'V_a$, and taking the inverse square root of the eigenvalues. Problems of non-uniqueness occur here too, in the case of singularities in $U_a'U_a$ and $V_a'V_a$, but these can be overcome in the same manner as in the TUCKALS2 algorithm.

For proof of the convergence one can adapt the proof for the TUCKALS algorithms given in section 4.4 and in Kroonenberg & De Leeuw (1980).

5.3 NON-SINGULAR TRANSFORMATIONS

Problem and solution. The procedure presented in this section is, in fact, nothing but the CANDECOMP procedure, as outlined in Carroll & Chang (1970). As set forward in section 3.3 the CANDECOMP model

$$z_{ijk} = \sum_{p=1}^{n} g_{ip} h_{jp} c_{ppk}$$

is used to find the best approximate decomposition for certain scalar-product data. CANDECOMP can be used for non-singular transformation of an extended core matrix with different first and second mode; a similar proposal has been made by Cohen (1974, 1975) to use INDSCAL on the extended core matrix from three-mode scaling.

The problem of achieving optimal diagonality of the extended core matric using non-singular transformations can be formulated as the

Diagonality problem (NS)

Let $C = (C_1, C_2, \ldots, C_n)$ with $C_k \in R^{s\times s}$ $(k=1,\ldots,n)$ be given. Find the non-singular $A \in R^{s\times s}$ and $B \in R^{s\times s}$, and $D = (D_1, D_2, \ldots, D_n)$ with D_k is diagonal $(k=1,\ldots,n)$, such that

$$\tau(A,B,D) = \sum_{k=1}^{n} \operatorname{tr} (C_k - AD_kB')'(C_k - AD_kB') \tag{5.9}$$

is as small as possible.

Theorem 5.2

Let $C_1 = (C_1, C_2, \ldots, C_n)$ with $C_k \in R^{s\times s}$ $(k=1,\ldots,n)$, and the diagonality problem NS be given. Then

$\hat{A} = \hat{U}_A'\hat{V}_A^{-1}$ with $\hat{U}_A = \sum_{k=1}^{n} \hat{D}_k\hat{B}'C_k'$ and $\hat{V}_A = \sum_{k=1}^{n} \hat{D}_k\hat{B}'\hat{B}\hat{D}_k$,

$\hat{B} = \hat{U}_B'V_B^{-1}$ with $\hat{U}_B = \sum_{k=1}^{n} \hat{D}_k\hat{A}'C_k$ and $\hat{V}_B = \sum_{k=1}^{n} \hat{D}_k\hat{A}'\hat{A}\hat{D}_k$, and

$\hat{D}_k = (\hat{d}_1^k, \ldots, \hat{d}_s^k)$ with $\hat{d}_p^k = \sum_{q=1}^{s} (\hat{B}'\hat{C}_k'A)_{qq} \; (\hat{A}'\hat{A}x\hat{B}'\hat{B})^{-1}_{pq}$,

where "x" is the element-wise product of $(p=1,\ldots,s$; $k=1,\ldots,n)$ two matrices, solve the diagonality problem NS, and the minimum is equal to

$$\tau(\hat{A},\hat{B},\hat{D}) = \sum_{k=1}^{n} \operatorname{tr} C_k'C_k - \operatorname{tr} \hat{B}\hat{U}_B' = \sum_{k=1}^{n} \operatorname{tr} C_k'C_k - \operatorname{tr} \hat{A}\hat{U}_A'.$$

Proof:

$\tau(A,B,D) = \sum_{k=1}^{n} \text{tr } (C_k - AD_kB')'(C_k - AD_kB')$ can be written in two ways:

1. $$\tau(A,B,D) = \sum_{k=1}^{n} \text{tr } C_k'C_k - 2\sum_{k=1}^{n} \text{tr } BD_kA'C_k + \sum_{k=1}^{n} \text{tr } BD_kA'AD_kB'$$
$$= \sum_{k=1}^{n} \text{tr } C_k'C_k - 2 \text{ tr } B\{\sum_{k=1}^{n} D_kA'C_k\} + \text{tr } B\{\sum_{k=1}^{n} D_kA'AD_k\}B'$$
$$= \sum_{k=1}^{n} \text{tr } C_k'C_k - 2\text{tr } BU_B + \text{tr } BV_BB' \qquad (5.10)$$
with $U_B = \sum_{k=1}^{n} D_kA'C_k$, and
$V_B = \sum_{k=1}^{n} D_kA'AD_k$

2. $$\tau(A,B,D) = \sum_{k=1}^{n} \text{tr } C_k'C_k - 2 \text{ tr } AU_A + \text{tr } AV_AA' \qquad (5.11)$$
with $U_A = \sum_{k=1}^{n} D_kB'C_k'$, and
$V_A = \sum_{k=1}^{n} D_kB'BD_k$

Differentiating τ with respect to A, B, and D leads to

$$\frac{\delta}{\delta A} \tau(A,B,D)\Big|_{\hat{A}} = -2U_A' + 2AV_A\Big|_{\hat{A}} = 0$$

$$\frac{\delta}{\delta B} \tau(A,B,D)\Big|_{\hat{B}} = -2U_B' + 2BV_B\Big|_{\hat{B}} = 0$$

$$\frac{\delta}{\delta D_k} \tau(A,B,D)\Big|_{\hat{D}_k} = -2\text{diag } (B'C_k'A) + 2 \text{ diag } (A'AD_kB'B)\Big|_{\hat{D}_k} = 0$$

$(k=1,\ldots,n)$

which gives as solution of the stationary equations

$$\hat{A} = \hat{U}_A' \hat{V}_A^{-1} \qquad (5.12)$$

$$\hat{B} = \hat{U}_B' \hat{V}_B^{-1}, \text{ and} \qquad (5.13)$$

$$d_p^k = \sum_{q=1}^{s} (B'C_k'A)_{qq} (A'A \times B'B)_{pq}^{-1} \quad (p=1,\ldots,s;k=1,\ldots,n) \qquad (5.14)$$

with "x" the elementwise matrix product.

To obtain the value of the minimum substitute $\hat{A}, \hat{B}$, and $\hat{D}$ into (5.10):

$$\tau(\hat{A},\hat{B},\hat{D}) = \sum_{k=1}^{n} \text{tr } C_k'C_k - 2\,\text{tr } \hat{B}\hat{U}_B + \text{tr } \hat{B}\hat{V}_B\hat{B}'$$

using (5.13) this gives

$$\tau(\hat{A},\hat{B},\hat{D}) = \sum_{k=1}^{n} \text{tr } C_k'C_k - 2\,\text{tr } \hat{B}\hat{U}_B + \text{tr } \hat{B}\hat{V}_B\hat{V}_B^{-1}\hat{U}_B =$$

$$= \sum_{k=1}^{n} \text{tr } C_k'C_k - \text{tr } \hat{B}\hat{U}_B,$$

which with the analoguous result for (5.11) and (5.12) gives the desired result. ▼▼

Standardization of transformation matrices

Given a solution (A,B,D) has been found, any

$A^* = A\Delta$, $B^* = B\tilde{\Delta}$ and $D^* = (D_1^*, D_2^*, \ldots, D_n^*)$ with

$D_k^* = \Delta^{-1} D_k \tilde{\Delta}^{-1}$ and $\Delta, \tilde{\Delta}$ full-rank diagonal matrices also constitute a solution, as

$$A^*D_k^*B^{*\prime} = (A\Delta)\ (\Delta^{-1} D_k \tilde{\Delta}^{-1})\ \tilde{\Delta}B' = AD_kB'$$

Without restricting the generality we may, therefore, fix the arbitrary scaling constants δ, and $\tilde{\delta}$ in such a way that the columns of A and B have unit length (see also Carroll & Chang, 1970, p.288-289). Some such choice is necessary to identify the stationary equations, and this particular choice has the advantage that the orthonormal transformation procedure from the previous section is a special case of the non-singular one.

As mentioned before, the solution given above is the same as that given by Carroll & Chang (1970), only the subject weights d_p^k are here treated per plane. In other words, we present CANDECOMP here as a procedure for a component model with two reduced modes with a diagonal extended core matrix, rather than a procedure for a component model with three reduced modes with body diagonal core matrix (see the discussion of these models in Chapter 3).

Algorithm. *Theorem 5.2* can be used to construct an alternating least squares algorithm to find the non-singular transformation A and B. One main iteration step of the NS-algorithm will be defined as:

A-substep

1. $UA_a = \sum_{k=1}^{n} D_k^{(a)} B_a' C_k'$

2. $VA_a = \sum_{k=1}^{n} D_k^{(a)} B_a' B_a D_k^{(a)}$

3. $\tilde{A} = UA_a' VA_a^{-1}$

4. $\Delta = \begin{pmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_s \end{pmatrix}$ with $\delta_p = \sqrt{\sum_{q=1}^{s} \tilde{a}_{pq}}$ (p=1,...,s)

5. $A_{a+1} = \tilde{A}\Delta^{-1}$

B-substep

1. $UB_a = \sum_{k=1}^{n} D_k^{(a)} A_{a+1}' C_k$

2. $VB_a = \sum_{k=1}^{n} D_k^{(a)} A_{a+1}' A_{a+1} D_k^{(a)}$

3. $\tilde{B} = UB_a' VB_a^{-1}$

4. $\tilde{\Delta} = \begin{pmatrix} \tilde{\delta}_1 & & \\ & \ddots & \\ & & \tilde{\delta}_s \end{pmatrix}$ with $\tilde{\delta}_p = \sqrt{\sum_{q=1}^{s} \tilde{b}_{pq}}$ (p=1,...,s)

5. $B_{a+1} = \tilde{B}\tilde{\Delta}^{-1}$

D-substep

2. $VD_a = (A_a' A_a) \times (B_a' B_a)$

$UD_a = B_a' C_k' A_a$ (k=1,...,n)

3. $d_{pk}^{(a+1)} = \sum_{q=1}^{s} (UD_a)_{qq} (VD_a^{-1})_{pq}$ (p=1,...,s; k=1,...,n)

5.4 COMPARISON OF TRANSFORMATION PROCEDURES

Once the optimal transformation matrices have been found, they can be applied to the component matrices with the inverse transformations applied to the core matrix. Thus, after having found the optimal orthonormal transformations $\hat{K}$ and $\hat{L}$, $\hat{Z}_k = GC_kH'$ may be decomposed as $\hat{Z}_k = G^*C_k^*H^{*\prime}$ with $G^* = G\hat{K}'$, $H^*=H\hat{L}'$ and $C_k^*=\hat{K}C_k\hat{L}'$. Similarly, after having found the optimal non-singular transformations $\hat{A}$ and $\hat{B}$, $\hat{Z}_k$ may be decomposed as $\hat{Z}_k = G^*C_k^*H^{*\prime}$ with $G^* = G\hat{A}$, $H^*=H\hat{B}$ and $C^*_k = \hat{A}^{-1}C_k(\hat{B}')^{-1}$. By using the transformations this way no additional loss is incurred and the transformed solution is just as adequate from the fitting point of view, as the original principal component solution.

In comparing the two transformation procedures it should be realized that the two loss functions have a different character, and that this leads to a number of differences in behaviour. The orthonormal loss function is

$$\sigma(K,L,D) = [[D_k - KC_kL']]^2 \quad , \text{ K, L orthonormal}$$

which leads to $\hat{D}_k = \text{diag } [\hat{K}C_k\hat{L}']$, thus $\sigma(\hat{K},\hat{L},\hat{D})$ is the sum of squares of the off-diagonal elements of the transformed $\hat{C}_k^* = \hat{K}C_k\hat{L}'$, and because of that, ON is a true diagonalization procedure. The analoguous loss function in the non-singular case would be something like

$$\tilde{\tau}(A,B,D) = [[D_k - AC_kB']]^2 \quad , \text{ A, B non-singular with unit length columns.}$$

The problem with this loss function is that the restrictions do not identify the minimization problem. A stricter requirement would be that the determinants of A and B are equal to one, but this could lead to rather complicated algorithms, which still have to be investigated.

The CANDECOMP loss function and solution described in section 5.3 was chosen for our preliminary investigations into non-singular transformations of the core matrix

$$\tau(A,B,D) = [[C_k - AD_kB']]^2 \quad \text{A, B non-singular with unit length columns.}$$

Properly speaking this is not a diagonalization procedure, but a decomposition of the core planes into the transformation matrices and diagonal matrices. $\hat{D}_k$ is not the diagonal of $C_k^* = A^{-1}C_k(B')^{-1}=$

when there is no exact solution, but the difference will become smaller when the loss becomes smaller. The parallel of the non-singular loss function for the orthonormal case would be the loss function for an *orthonormal INDSCAL* model, and the difference with the orthonormal transformation procedure in section 5.2 is that the definition of U and V is slightly different

$$U = \sum_{k=1}^{n} C_k L' D_k ,$$

and

$$V = \sum_{k=1}^{n} C_k' K' D_k .$$

For the non-singular case the difference in loss function implies that there is a difference between (1) the results (in terms of sums of squares) from the transformation procedure to find A and B, and (2) the results from applying A and B to the core matrix:

$$(1) \quad NS : \quad Z_k \cong GC_kH' \cong G\{AD_kB'\}H' = (GA)D_k(HB)'$$
$$(2) \quad NS : \quad Z_k \cong GC_kH' = (GA)\{A^{-1}C_k(B')^{-1}\}(HB)'$$

For comparison in the orthonormal case these decompositions are:

$$(1) \quad ON : \quad Z_k \cong GC_kH' \cong G\{K'D_kL\}H' = (GK')D_k(HL')'$$
$$(2) \quad ON : \quad Z_k \cong GC_kH' = (GK')\{KC_kL'\}(HL').$$

We will illustrate the differences between the two kinds of results for the non-singular procedure in section 5.5.

For interpretational purposes, there is a distinct disadvantage connected with the non-singular transformation when there is no exact solution to the diagonalization procedure NS. If there is no exact solution, $A^{-1}C_k(B')^{-1}$ is not diagonal, and the off-diagonal elements are no longer the sole expression of the relationships between components. Part of these relationships has been transferred to the non-orthogonality of the components themselves. No such complications occur with orthonormal transformations. In other words, the strength of the relationships between the components is now divided over two quantities, and this poses as yet unsolved interpretational complications. Especially in those cases where the non-singular transformations become nearly singular as, for instance, in the *Perceived reality study* (see section 5.5).

If the solution of the non-singular transformation is not exact, and one wants to avoid the complications of this splitting up of dependencies between components, one could settle for the extra loss from the transformation procedure, and use the values of D_k as the saliences or subject weights in combination with the correlated components $G^* = GA$, and $H^* = HB$, which are the same for (1) and (2) anyway. The scalar products $G^{*\prime}G^*$ and $H^{*\prime}H^*$ which are equal to $A'A$ and $B'B$ respectively then indicate the covariations of the components, or correlations if the component matrix is centred; the singular values of scalar product matrices will indicate the degree of non-singularity.

In section 5.5 we present some results of transforming the core matrix both by orthonormal and non-singular transformations.

5.5 ILLUSTRATIONS OF TRANSFORMATIONS

The main function of this section is to show numerical results illustrating the theory. Proper interpretation and assessment, especially of the non-singular transformation procedure, requires a more extensive investigation. For the orthonormal transformation procedure the situation is simpler, as it is a true diagonalization procedure, and the orthonormality of the transformation leaves the main characteristics of the TUCKALS2 solution unimpaired, and, therefore, poses no additional interpretational problems.

Four ability-factor study. (Meyers, Dingman, Orpet, Sitkei, & Watts, 1964; see Chapter 12) From Tables 5.1 and 5.2, which show the various core matrices, it follows that the orthonormal transformation of the core matrix does not improve the diagonality very much. Similar small improvements of diagonality can be observed in many other data sets, especially in those with symmetric frontal planes in the data block, like correlation and (dis)similarity matrices.

The non-singular transformation procedure succeeds rather well with a mere increase in the standardized loss of .0064 compared to the TUCKALS2 loss. This means that it is possible to decompose the

Table 5.1 *Four ability-factor study: results of transformation procedures (4x4 solution)*

Standardized sums of squares TUCKALS2 algorithm

Total sum of squares	-SS(Total)	1.0000
Fitted sum of squares (=sum of squares of core matrix)	-SS(Core)	.9394
Sum of squares of diagonal elements of core matrix	-SS(Dia)	.9247
Sum of squares of off-diagonal elements	-SS(Off)	.0147

Standardized sums of squares of transformation procedure

		ON	NS
Fitted sum of squares of the transformation procedure	-SS(Proc-Fit)	.9250	.9333
Sum of squares of the diagonal matrices D_k	-SS(Dia)	.9250	.6652

Standardized sums of squares after applying transformations

		ON	NS
Sum of squares of core matrix	-SS(Core)	.9394	.6922
Sum of squares of diagonal elements of core matrix	-SS(Dia)	.9250	.6749
Sum of squares of off-diagonal elements of core matrix	-SS(Off)	.0143	.0172

Transformation matrices

Orthonormal case (K=L)				non-singular case (A=B)			
1.000	.002	-.002	-.002	.972	-.189	.464	.841
-.001	.996	.091	-.003	-.012	.926	-.349	.238
.002	-.090	.983	-.161	-.207	.060	.646	.330
.002	-.012	.161	.987	-.109	-.321	-.496	.357
Singular values of non-singular transformation				1.423	1.009	.814	.528

Average core matrices

three-mode analysis				orthonormal				non-singular			
10.79	-.04	.10	.03	10.79	-.04	.11	.06	7.99	-.01	.01	-.11
-.04	2.45	.03	.01	-.04	2.45	-.04	.01	-.01	2.32	.03	-.01
.10	.03	1.65	-.03	.11	-.04	1.66	.02	.01	.03	2.04	.00
.03	.02	-.03	1.37	.06	.01	.01	1.37	-.11	-.01	.00	4.08

frontal planes C_k of the core matrix into $\hat{A}\hat{D}_k\hat{B}'$ (or $\hat{A}\hat{D}_k\hat{A}'$ for the Meyers et al. data, as the input frontal planes are symmetric) without any real loss compared to the original TUCKALS2 solution. Thus the diagonal elements (see column 4, Table 5.2) can be interpreted as the weights or saliences which the groups attach to the axes of the common transformed space.

If one, in general, does not want to accept the additional loss, the core plane C_k should be transformed into

$$C_k^* = \hat{A}^{-1}C_k(\hat{B}')^{-1}$$

Table 5.2 *Four ability-factor study: core matrices* (x 10)

	TUCKALS2				Orthonormal				Non-singular				Diagonals from NS transformation procedure			
R2	79				79				65				56			
	-3	33			-3	33			- 4	28				30		
	2	-2	26		2	-3	26		-12	-4	29				24	
	7	3	1	13	7	2	1	13	-10	3	2	48				38
R4	135				135				98				103			
	3	20			3	20			-1	15				17		
	-4	-0	15		-4	-1	15		-3	-3	15				15	
	2	3	2	9	2	3	2	10	6	2	1	42				48
R6	116				116				65				75			
	4	32			4	33			3	35				33		
	7	4	8		6	2	7		9	2	11				13	
	6	-5	2	17	7	-4	1	18	7	-2	-4	52				57
N2	136				136				122				112			
	1	10			2	10			-2	10				9		
	-11	1	19		-10	1	19		-3	4	20				19	
	-8	1	-1	11	-10	1	-0	11	-2	-1	3	27				26
N4	84				84				56				55			
	-5	26			-4	26			-3	22				23		
	9	1	16		10	-0	16		8	0	25				29	
	-3	3	-2	15	-2	3	-2	15	-6	2	-1	41				35
N6	98				98				72				71			
	0	26			1	25			6	29				26		
	3	-1	16		3	-2	16		2	3	22				22	
	-2	-4	-2	16	-1	-5	-1	16	-1	5	-0	35				35

as explained in sections 5.3 and 5.4. When the fit is not perfect the C_k^* will not be completely diagonal, and the diagonal elements of C_k^* will not be exactly the same as in D_k (compare columnns 3 and 4 in Table 5.2), but the transformed component space $G^*=GA$ and $H^*=HB$ are the same as was set forth in section 5.4. The transformed components will in general no longer be orthogonal, and from Table 5.3 it can be seen that substantial correlations (and scalar products) may arise.

Table 5.3 shows the transformed components for Meyers et al. data. Whether one prefers the transformed components or the original ones, seems largely a matter of taste. The insight in the spatial arrangement of the tests (see Fig. 12.1) is not greatly enhanced by the non-singular transformation. On the other hand, the comparisons between the various weights the groups attach to the components is somewhat simpler due to the (near)diagonality. One

Table 5.3 *Four ability-factor study: transformed component space* (x 100)

Tests		1	2	3	4
A	1	32	35	7	14
Hand-Eye	2	30	38	-10	23
Psychomotor	3	32	41	-14	15
B	4	22	3	49	40
Perceptual	5	22	-1	48	43
Speed	6	22	-0	44	40
C	7	35	-39	25	3
Linguistic	8	32	-37	33	-3
Ability	9	34	-30	18	-3
D	10	22	-29	- 9	41
Figural	11	29	-30	-20	32
Reasoning	12	28	- 8	-15	34

component correlations

100			
- 3	100		
-27	-16	100	
-91	18	0	100

simpler due to the (near)diagonality. One of the reasons for the relatively small differences is that the core matrix was already reasonably diagonal to start with.

Perceived reality study. Non-singular transformations of the core matrix of this study (Van der Voort, 1982), discussed in detail in section 7.5, show an entirely different picture (Table 5.4).

Although the fit of the non-singular transformation procedure was quite good (the additional loss was only .0050), the results are far from attractive. The smallest singular values of the transformation matrices A and B are getting rather small, indicating that A and B are approaching singularity. Also noteworthy is the very large sum of squares of the core matrix. Note that the sum of squares of the core elements no longer adds up to the TUCKALS SS(Fit), because of correlations between components. The higher values in the core matrix after the non-singular transformation are the immediate consequence of these high correlations. It is, by the way, possible to scale the sum of squares of the transformed

Table 5.4 *Perceived reality study: results of transformation procedures*

(3x3 solution)

Standardized sums of squares

TUCKALS2		Transformation procedures	ON	NS	Application of transformations	ON	NS
SS(Tot)	1.0000	SS(Proc.Fit)	.8538	.8966	SS(Core)	.9016	5.1259
SS(Core)	.9016	SS(Dia)	.8538	4.6836	SS(Dia)	.8538	4.8062
SS(Dia)	.8468				SS(Off)	.0478	.3197
SS(Off)	.0548	iterations	22	> 200			

Transformation matrices

Orthonormal K			Orthonormal L			Non-singular A			Non-singular B		
.990	-.142	-.012	.996	.087	.006	.995	.490	-.942	.445	-.087	.030
.142	.990	.011	-.087	.992	.089	-.093	.727	.108	.828	.995	.884
.010	-.012	1.000	.002	-.090	.996	-.031	.482	.318	-.341	.052	.467
			singular values			1.46	.92	.16	1.63	.48	.31

Average core matrices

TUCKALS2			Orthonormal			Non-singular		
-7.27	.17	.05	-7.25	.23	.01	-15.14	.22	.28
.32	4.09	.31	- .36	4.13	-.04	.02	5.56	-.05
.09	-.27	1.77	- .00	-.16	1.78	.01	.21	12.17

core matrix down to its original size, but only by multiplying the components with the reciprocal scaling constants. Inspection of the transformed core matrix (not shown) indicates that large off-diagonal elements exist, thus pointing to non-diagonality. On the other hand, the good fit of the procedure shows that the D_k can be used as saliences for the transformed components. An adequate way to deal with this seeming contradiction still has to be developed.

5.6 CONCLUDING REMARKS

This chapter has been concerned with the problem of diagonality of the extended core matrix. Especially the non-singular transformation procedure is still problematic both technically and interpretationally. From a technical point of view it is not clear if the CANDECOMP procedure is the most adequate procedure to use for the purpose, and how diagonality should be measured in the presence of near-singularity of the transformation matrices. On the interpretational side the problem exists how to deal with large off-diagonal elements in a situation of good fit, and how to use the diagonality to its utmost advantage. The interpretation of highly related components is also somewhat difficult to deal with. The problem of interpreting 'oblique' components is, of course, partly conceptual and has been discussed extensively in the context of standard principal component analysis and factor analysis, a discussion we do not go into here.

Further insight into the behaviour of the solutions and further evaluation of the results may be obtained by a direct comparison with the results of, for instance, an INDSCAL analysis on the data of the *Four ability-factor study*, and a CANDECOMP analysis on the data of the *Perceived reality study*. This, however, merits another study.

II

THEORY FOR APPLICATIONS

SUMMARY

As Part I, *Part II* deals with theoretical issues, but now the focus is on those theoretical problems which arise out of applying three-mode principal component analysis to real data sets. Three main issues are tackled: preprocessing of input, postprocessing of output, and the analysis of the not-fitted part of the data.

The first part of *Chapter 6* reviews proposals which have been put forward to scale input data, such that they are fit for a (three-mode) principal component analysis. Procedures for handling means and variances are discussed. To this end a distinction is made between (un)interpretable and (in)comparable means and variances. A large variety of models exist for dealing with interpretable means, which generally consist of additive terms for the means (in an analysis-of-variance fashion), and multiplicative or product terms for the components. Some such models are discussed and evaluated. The problem of iterative standardization in three-mode models is discussed briefly.

The second part of Chapter 6 deals with the interpretation of output, and ways to improve the interpretability of the results. Within this context the scaling of components and core matrices as well as their interpretation, joint plots, and component scores are treated in some detail.

In *Chapter 7* the focus is on that part of the data which is not accommodated by the three-mode model. After a general discussion of residuals from principal component analysis, detailed recommendations and procedures are provided (and applied) for three-mode residuals, such as analysis of variance of the squared residuals, sums-of-squares plots, and the use of normal probability plots for the residuals.