JOURNAL OF CHEMOMETRICS
J. Chemometrics 2000; **14**: 285–299

Construction and analysis of degenerate PARAFAC models

Pentti Paatero*†

Department of Physics, University of Helsinki, Box 9, FIN-00014 Helsinki, Finland

SUMMARY

A mathematical framework is presented for constructing degenerate CANDECOMP/PARAFAC models. It is possible to construct degenerate arrays which can be approximated by two-factor models to arbitrary precision but which do not possess an exact two-factor representation. Equivalence of different degenerate presentations is demonstrated. By using this model, tasks are constructed where the straight path from the specified starting point to the best-fit solution will pass through a degenerate area. Swamp behavior is observed when such tasks are solved by various algorithms. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: alternating least squares; CANDECOMP; Kruskal polynomial; positive matrix factorization; swamps; three-way arrays

1. INTRODUCTION

According to the CANDECOMP/PARAFAC (CP) decomposition [1–3] a three-way array $\underline{X}$ is expressed as a trilinear expression of three factor matrices **A, B** and **C**:

$$x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} \quad (i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K) \tag{1}$$

Often the model is used for finding a least squares approximation for a given three-way array $\underline{X}^0$:

$$x^0_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} + e_{ijk} \tag{2}$$

where the factor matrices **A, B** and **C** are to be determined so that a norm of the difference array $\underline{E}$ is minimized.

Work with Equation (2) has long been hampered by the occurrence of so-called degenerate arrays. Best-fit solutions of such arrays approach infinity in such a way that the fitted array remains finite although some (or all) of the factor elements approach plus or minus infinity.

In this work a mathematical model is presented which allows one to construct arrays whose two-factor representations contain arbitrarily large factor elements. It is also possible to construct degenerate arrays which can be approximated by two-factor models to arbitrary precision but which

* Correspondence to: P. Paatero, Department of Physics, University of Helsinki, Box 9, FIN-00014 Helsinki, Finland.
† E-mail: pentti.paatero@helsinki.fi.
Contract/grant sponsor: University of Helsinki

 Accepted 30 November 1999

do not possess an exact two-factor representation.

When discussing the possible arrays, it is assumed tacitly in this work that the three dimensions of the array $\underline{\mathrm{X}}$ have been fixed to some chosen values. Generally it is also assumed that the number of factors has been fixed to two; thus also the dimensions of the three factor matrices are determined.

1.1. Notation

The notation suggested by Kiers [4] is used. In addition, the triple product is defined according to Kruskal [5] as follows. Equation (1) is written as $\underline{\mathrm{X}} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]$. Similarly, $\underline{\mathrm{X}} = [\mathbf{t}, \mathbf{u}, \mathbf{v}]$, where **t**, **u** and **v** are vectors of correct dimensions, specifies the equations $x_{ijk} = t_i u_j v_k$, $i = 1,\ldots,I$, $j = 1,\ldots,J$, $k = 1,\ldots,K$. Three-way arrays are displayed so that slices corresponding to successive values of the last index are printed side by side. Thus the indices of elements of a displayed array of dimensions $2 \times 2 \times 2$ are as follows:

$$\left(\begin{array}{cc|cc} 111 & 121 & 112 & 122 \\ 211 & 221 & 212 & 222 \end{array}\right) \tag{3}$$

When discussing the three-way least squares problem, the 'true solution' is denoted by $\underline{\mathrm{X}}^0$. Then the task is to determine an array $\underline{\mathrm{X}}$ so that the expression $Q = \|\underline{\mathrm{X}} - \underline{\mathrm{X}}^0\|_{\mathrm{F}}^2$ is minimized.

1.2. Terminology

The concept of degeneracy is used in two meanings. A *degenerate array* (in the exact sense of the word) is an array having a certain rank (e.g. rank = 3) which may be approximated arbitrarily well by a factorization of lower rank (e.g. rank = 2). In such approximating factorizations there are large positive and negative contributions which mostly cancel each other. These factorizations are called *degenerate* in the qualitative sense of the word. The 'degree of degeneracy' means the degree of cancellation of positive and negative contributions. The word *loadings* is used for all three aspects of the factors; this is in contrast to two-way terminology where the words 'loadings' and 'scores' differentiate between the two aspects.

1.3. Previous work by Kruskal and Ten Berge

It is well known that Kruskal laid the foundations for understanding degenerate arrays [5–8]. After the main results of the present work had been formulated, the author became aware of prior unpublished work by Kruskal along these lines. Kruskal kindly sent his notes of this earlier work. He derives Equation (5) for a degenerate $2 \times 2 \times 2$ array by a constructive process. Essentially he works in the opposite direction from the present work: starting from the properties of a degenerate solution, he constructs Equation (5) by a technique which guarantees the property of degeneracy.

Kruskal also discusses the use of a diagnostic polynomial D of the elements of a $2 \times 2 \times 2$ array for determining the rank of the array (see Reference [7], p. 10). Positive and negative values of D signify that the array is of rank 2 or 3 respectively. If $D = 0$, however, the rank may be either 0, 1, 2 or 3. Ten Berge presents the expression of the polynomial based on determinants over pairs of fibers of the array [9]. He also sharpens the definition of the diagnostic properties of D: only such arrays are considered where at least one slice $\mathbf{X}_{::1}$ is non-singular and the other slice $\mathbf{X}_{::2}$ is non-zero and not proportional to $\mathbf{X}_{::1}$. In this class of arrays, having $D > 0$ is the necessary and sufficient condition for rank = 2. If $D \leq 0$, rank = 3. The arrays with $D = 0$ are special: although their rank is 3, they may be approximated arbitrarily well by arrays whose rank is 2.

2. DEGENERATE ARRAYS AS LIMITING POINTS OF SEQUENCES OF SOLUTIONS WITH INCREASING DEGREES OF DEGENERACY

The columns of the factor matrices **A**, **B** and **C** contain the loadings of the factors. The first column of matrix **A** contains the 'A-mode loadings' for the first factor, etc.

The following two-factor model was constructed as a tool for exploring degeneracy:

$$\underline{\mathrm{X}} = [\mathbf{A}, \mathbf{B}, \mathbf{C}] \tag{4}$$

$$= \left[\left(\frac{1}{\epsilon}\mathbf{a} + \frac{\epsilon^2}{2}\boldsymbol{\alpha} \;\middle|\; -\frac{1}{\epsilon}\mathbf{a} + \frac{\epsilon^2}{2}\boldsymbol{\alpha} \right), \left(\frac{1}{\epsilon}\mathbf{b} + \frac{\epsilon^2}{2}\boldsymbol{\beta} \;\middle|\; -\frac{1}{\epsilon}\mathbf{b} + \frac{\epsilon^2}{2}\boldsymbol{\beta} \right), \left(\frac{1}{\epsilon}\mathbf{c} + \frac{\epsilon^2}{2}\boldsymbol{\gamma} \;\middle|\; -\frac{1}{\epsilon}\mathbf{c} + \frac{\epsilon^2}{2}\boldsymbol{\gamma} \right) \right] \tag{5}$$

The symbols **a**, **b**, **c** and $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ represent arbitrary column vectors of correct dimensions. These are called the *generating vectors* of the model. The generating vectors form two sets of three vectors. The sets have different roles.

This model was created by an intuitive trial-and-error process. Thus it may be impossible to 'derive' the model from simpler concepts. The properties of the model are verified in the following. It is assumed that the vectors **a** and $\boldsymbol{\alpha}$ form a linearly independent set, and similarly {**b**, $\boldsymbol{\beta}$} and {**c**, $\boldsymbol{\gamma}$}.

2.1. Discussion of the degenerate model

The parameter ϵ controls the degree of degeneracy. When ϵ approaches zero, the loadings of all three modes of the two factors approach infinity proportionally to $1/\epsilon$ in such a way that the loadings of the first factor grow proportionally to the vectors **a**, **b** and **c** whereas the loadings of the second factor grow proportionally to the vectors $-\mathbf{a}$, $-\mathbf{b}$ and $-\mathbf{c}$. The vectors of the two factors become quickly more and more linearly dependent as the contributions due to the vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ decrease proportionally to the square of ϵ.

The model has been carefully constructed so that the negative powers of ϵ cancel out. This assures that the array $\underline{\mathrm{X}}$ remains finite although the loadings approach infinity when ϵ approaches zero. The following expression is obtained for the array $\underline{\mathrm{X}}$ of Equation (5):

$$\underline{\mathrm{X}} = [\boldsymbol{\alpha}, \mathbf{b}, \mathbf{c}] + [\mathbf{a}, \boldsymbol{\beta}, \mathbf{c}] + [\mathbf{a}, \mathbf{b}, \boldsymbol{\gamma}] + \frac{\epsilon^6}{4}[\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}] \tag{6}$$

For a non-zero ϵ the rank of this array is 2. When $\epsilon \to 0$, the loadings of the two-factor presentation of this array approach infinity, as explained for the model (5). The array $\underline{\mathrm{X}}$ approaches arbitrarily close to $\underline{\mathrm{X}}^*$ defined by

$$\underline{\mathrm{X}}^* = [\boldsymbol{\alpha}, \mathbf{b}, \mathbf{c}] + [\mathbf{a}, \boldsymbol{\beta}, \mathbf{c}] + [\mathbf{a}, \mathbf{b}, \boldsymbol{\gamma}] \tag{7}$$

The elements of the difference $\underline{\mathrm{X}} - \underline{\mathrm{X}}^*$ decrease proportionally to the *sixth* power of ϵ. The properties of the array $\underline{\mathrm{X}}^*$ are interesting. The array is defined by an expression whose rank is 3. It may be approximated by rank = 2 arrays to arbitrary precision. Kruskal *et al.* [8] describe arrays having similar properties. According to them, it appears that $\underline{\mathrm{X}}^*$ belongs to a boundary between two different domains in the space of all possible $\underline{\mathrm{X}}$. On the low-rank side of the boundary there are arrays whose rank is 2. On the other side the rank is 3. As will be shown later, an array on the high-rank side of the boundary is given by

$$\underline{\mathrm{X}} = [\boldsymbol{\alpha}, \mathbf{b}, \mathbf{c}] + [\mathbf{a}, \boldsymbol{\beta}, \mathbf{c}] + [\mathbf{a}, \mathbf{b}, \boldsymbol{\gamma}] + h[\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}] \tag{8}$$

where the scalar h is negative.

2.2. *Alternative forms of the first degenerate model*

The model (5) may be written in the following alternative forms where the array $\underline{\mathrm{X}}$ remains the same for all values of ϵ although the factor matrices **A**, **B** and **C** are changed:

$$\begin{aligned}\underline{\mathrm{X}} &= [\mathbf{A}, \mathbf{B}, \mathbf{C}] \\ &= \left[\left(\frac{1}{\epsilon}\mathbf{a} + \frac{\epsilon^2}{2}\boldsymbol{\alpha}\middle| -\frac{1}{\epsilon}\mathbf{a} + \frac{\epsilon^2}{2}\boldsymbol{\alpha}\right), \left(\frac{1}{\epsilon}\mathbf{b} + \frac{\epsilon^2}{2}\boldsymbol{\beta}\middle|\frac{1}{\epsilon}\mathbf{b} - \frac{\epsilon^2}{2}\boldsymbol{\beta}\right), \left(\frac{1}{\epsilon}\mathbf{c} + \frac{\epsilon^2}{2}\boldsymbol{\gamma}\middle|\frac{1}{\epsilon}\mathbf{c} - \frac{\epsilon^2}{2}\boldsymbol{\gamma}\right)\right] \qquad (9)\\ &= \left[\left(\frac{1}{\epsilon^3}\mathbf{a} + \frac{1}{2}\boldsymbol{\alpha}\middle| -\frac{1}{\epsilon^3}\mathbf{a} + \frac{1}{2}\boldsymbol{\alpha}\right), \left(\mathbf{b} + \frac{\epsilon^3}{2}\boldsymbol{\beta}\middle|\mathbf{b} - \frac{\epsilon^3}{2}\boldsymbol{\beta}\right), \left(\mathbf{c} + \frac{\epsilon^3}{2}\boldsymbol{\gamma}\middle|\mathbf{c} - \frac{\epsilon^3}{2}\boldsymbol{\gamma}\right)\right] \qquad (10)\end{aligned}$$

In these models the difference between the B-mode vectors approaches zero, and similarly the difference between the C-mode vectors. In the model (10) the diverging behavior has been concentrated in the loadings of the first mode. Then the loading vectors of the other two modes approach the generating vectors **b** and **c** when ϵ approaches zero.

2.3. *The second degenerate example*

By the trial-and-error process, another example was found that is not so symmetric as the first one. The equations for this model are

$$\begin{aligned}\underline{\mathrm{X}} &= [\mathbf{A}, \mathbf{B}, \mathbf{C}] \\ &= \left[\left(\frac{1}{\epsilon}\mathbf{a}\middle| -\frac{1}{\epsilon}\mathbf{a} + \epsilon^2\boldsymbol{\alpha}\right), \left(\frac{1}{\epsilon}\mathbf{b}\middle| -\frac{1}{\epsilon}\mathbf{b} + \epsilon^2\boldsymbol{\beta}\right), \left(\frac{1}{\epsilon}\mathbf{c} + \epsilon^2\boldsymbol{\gamma}\middle| -\frac{1}{\epsilon}\mathbf{c}\right)\right] \qquad (11)\end{aligned}$$

Performing the computations gives

$$\underline{\mathrm{X}} = [\boldsymbol{\alpha}, \mathbf{b}, \mathbf{c}] + [\mathbf{a}, \boldsymbol{\beta}, \mathbf{c}] + [\mathbf{a}, \mathbf{b}, \boldsymbol{\gamma}] - \epsilon^3[\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{c}] \qquad (12)$$

Comparison with Equation (6) shows that when $\epsilon \to 0$, the array (12) approaches the same limiting array $\underline{\mathrm{X}}^*$ as the array (6). However, the difference term has a different form. Furthermore, now the difference decreases proportionally to the *third* power of ϵ. Both the positive and negative values of ϵ correspond to arrays whose rank is 2. The value $\epsilon = 0$ corresponds to a degenerate array.

The second degenerate example has special historical significance. When discussing his unpublished work leading to Equation (5), Kruskal mentioned that the existence of sequences of models as defined by Equation (11) prevented him from formulating a satisfactory proof of his results and thus prevented him from publishing that work.

3. CONNECTIONS WITH THE TUCKER3 MODEL

In order to avoid ambiguity with the model (1), the well-known Tucker3 (T3) model [10] is here written as

$$x_{ijk} = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} \overline{a}_{ip}\overline{b}_{jq}\overline{c}_{kr}g_{pqr} \quad (i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K) \qquad (13)$$

for the case where the core array $\underline{\mathrm{G}}$ is of dimensions $P \times Q \times R$. Setting $P = Q = R = 2$, the factor matrices $\overline{\mathbf{A}}$, $\overline{\mathbf{B}}$ and $\overline{\mathbf{C}}$are defined as follows:

$$\overline{\mathbf{A}} = (\mathbf{a}|\boldsymbol{\alpha}), \qquad \overline{\mathbf{B}} = (\mathbf{b}|\boldsymbol{\beta}), \qquad \overline{\mathbf{C}} = (\mathbf{c}|\boldsymbol{\gamma}) \tag{14}$$

It is easy to see that the degenerate model (7) may be interpreted as a T3 model where the core array is

$$\underline{\mathrm{G}} = \left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right) \quad \left[\begin{array}{c}\text{meaning of} \\ \underline{\mathrm{G}}\ \text{elements}\end{array}\right] : \left(\begin{array}{cc|cc} \mathbf{abc} & \mathbf{a}\boldsymbol{\beta}\mathbf{c} & \mathbf{ab}\boldsymbol{\gamma} & \mathbf{a}\boldsymbol{\beta}\boldsymbol{\gamma} \\ \boldsymbol{\alpha}\mathbf{bc} & \boldsymbol{\alpha}\boldsymbol{\beta}\mathbf{c} & \boldsymbol{\alpha}\mathbf{b}\boldsymbol{\gamma} & \boldsymbol{\alpha}\boldsymbol{\beta}\boldsymbol{\gamma} \end{array}\right) \tag{15}$$

The general model (8) is similarly interpreted as T3 with

$$\underline{\mathrm{G}} = \left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & h \end{array}\right) \tag{16}$$

If h is positive, there is a two-factor CP model corresponding to the T3 model of Equation (16). However, when $h \to 0$, the factors of the corresponding CANDECOMP/PARAFAC model grow without limit, and with $h \leq 0$ there is no exactly corresponding two-factor CP model any more. As a description of a given (empirical) degenerate array, Equations (13)–(15) might be considered more desirable than the corresponding CP model (5) where factor elements approach infinity.

3.1. *Same rank for* $\underline{\mathrm{X}}$ *and* $\underline{\mathrm{G}}$

It is possible to state a connection between the ranks of $\underline{\mathrm{X}}$ and $\underline{\mathrm{G}}$ in Equation (13). When considering the equation

$$x_{ijk} = \sum_{p=1}^{P} \overline{a}_{ip} w_{pjk} \quad (i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K) \tag{17}$$

Kruskal (see Reference [5], p. 104) shows that if the number of linearly independent slices is the same in the sets $\mathbf{X}_{i::}$ (i = 1,…,I) and $\mathbf{W}_{p::}$ (p = 1,…,P), then rank($\underline{\mathrm{X}}$) = rank($\underline{\mathrm{W}}$). On the other hand, the numbers of linearly independent slices are the same if the matrix ($\overline{\mathbf{A}}$) is of full column rank, i.e. if rank ($\overline{\mathbf{A}}$) = P. By applying this result three times, we get the rule that $\underline{\mathrm{X}}$ and $\underline{\mathrm{G}}$ in Equation (13) are of same rank if all three matrices $\overline{\mathbf{A}}$, $\overline{\mathbf{B}}$ and $\overline{\mathbf{C}}$ are of full column rank, i.e. if rank ($\overline{\mathbf{A}}$) = P, rank($\overline{\mathbf{B}}$) = Q and rank($\overline{\mathbf{C}}$) = R. In this work the full rank of $\overline{\mathbf{A}}$, $\overline{\mathbf{B}}$ and $\overline{\mathbf{C}}$ is always assumed. Thus we may decide the rank of $\underline{\mathrm{X}}$ by inspecting the corresponding core array $\underline{\mathrm{G}}$.

4. ANALYSIS OF THE 2 × 2 × 2 ARRAY

The Kruskal polynomial D has already been discussed in Section 1·3. It is useful to derive for D an expression that is directly based on the elements of the array. In order to simplify notation, the elements of the array $\underline{\mathrm{X}}$ are here denoted as

$$\underline{\mathrm{X}} = \left(\begin{array}{cc|cc} a & b & e & f \\ c & d & g & h \end{array}\right) \tag{18}$$

and the factors as

$$\mathbf{A} = \begin{pmatrix} o & u \\ p & v \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} q & w \\ r & x \end{pmatrix}, \qquad \mathbf{C} = \begin{pmatrix} s & y \\ t & z \end{pmatrix} \tag{19}$$

By starting from the definition of D as given by Ten Berge [9], one obtains through a tedious but straightforward calculation the following presentation for the Kruskal polynomial D (one should note that although this D looks quite different from the D shown by Ten Berge, the two representations describe exactly the same polynomial of the elements of the array $\underline{\mathrm{X}}$):

$$\begin{aligned} D(a,b,\ldots,h) = {} & aahh + bbgg + ccff + ddee \\ & - 2(acfh + bdeg + abgh + cdef + adeh + bcfg) \\ & + 4(adfg + bceh) \end{aligned} \tag{20}$$

If $D(\underline{\mathrm{X}}) > 0$, then the array is of rank 2. Similarly, $D(\underline{\mathrm{X}}) < 0$ implies that rank($\underline{\mathrm{X}}$) = 3. On *ordinary points* (i.e. on those points that satisfy the criteria of Ten Berge) of the hypersurface defined by the equation $D(\underline{\mathrm{X}}) = 0$, the rank is 3. At the *special points* rejected by the criteria of Ten Berge, the rank is 0, 1 or 2. At ordinary points of the hypersurface the array is degenerate: at such points the exact representation of $\underline{\mathrm{X}}$ is of rank 3, but an arbitrarily good approximation of $\underline{\mathrm{X}}$ is obtained with rank = 2.

In order to simplify the problem, the following arrays are considered:

$$\underline{\mathrm{X}} = \left(\begin{array}{cc|cc} 0 & 1 & e & 0 \\ 1 & d & 0 & h \end{array}\right) \tag{21}$$

where e is regarded as a constant, typically $e = 30$. For arrays defined by (21), Equation (20) simplifies to $D(d,h) = e^2 d^2 + 4e\,h$. Considering only one of the variables d and h at a time, the earlier results are confirmed: if $e > 0$ and $d = 0$, the sign of h decides if rank = 2 ($h > 0$) or 3 ($h < 0$); if, on the other hand, $e \neq 0$ and $h = 0$, then the array is degenerate at $d = 0$ but has rank 2 with all other values of d.

Considering the two variables d and h together gives a more detailed picture. The parabola $h = -e\,d^2/4$, consisting of degenerate border points, is shown in Figure 1 (See Section 6). All points below the parabola have rank 3, while all points above have rank 2, when considering the array $\underline{\mathrm{X}}(d,h)$ specified by Equation (21).

The CANDECOMP/PARAFAC decomposition of $\underline{\mathrm{X}}$ defined by Equation (21) is expressed in the following form:

$$\underline{\mathrm{X}} = \left(\begin{array}{cc|cc} 0 & 1 & e & 0 \\ 1 & d & 0 & h \end{array}\right) \tag{22}$$

$$= [\mathbf{A}, \mathbf{B}, \mathbf{C}] \tag{23}$$

$$= \left[\begin{pmatrix} o & u \\ p & v \end{pmatrix}, \begin{pmatrix} q & w \\ r & x \end{pmatrix}, \begin{pmatrix} s & y \\ t & z \end{pmatrix}\right] \tag{24}$$

$$= \left[\begin{pmatrix} 1/\epsilon & -1/\epsilon \\ p & v \end{pmatrix}, \begin{pmatrix} 1/\epsilon & -1/\epsilon \\ r & x \end{pmatrix}, \begin{pmatrix} 1/\epsilon & -1/\epsilon \\ t & z \end{pmatrix}\right] \tag{25}$$

The form of Equation (25) fixes the four arbitrary normalizations and guarantees that $x_{111} = 0$. The variables p, v, r, x, t and z may be eliminated by lengthy but straightforward calculations, leading to

$$4h = -ed^2 + e\epsilon^6 \tag{26}$$

By letting $\epsilon \to 0$, the equation of the parabola of degeneracy is again obtained: $4h = -e\,d^2$. In this way the result easily obtained using the Kruskal polynomial can be laboriously confirmed in this special case.

4.1. The general case

Examine the shape of the curve $D(d,h) = 0$ with arbitrary values of the six array elements $\{a,b,c,e,f,g\}$. Ignore the case $a = e = 0$; then D does not depend on d or h at all. At first sight it might appear that $D(d, h)$ is a general polynomial of second degree. However, closer inspection shows that $D(d,h) = 0$ in fact always represents a parabola. The orientation and position of this parabola depend on the numerical values of the six array elements. In contrast, the expression $D(d, f)$ is a general second-order polynomial of d and f. Thus in different (d, f) hyperplanes the shape of the degenerate border curve $D(d,f) = 0$ may be a parabola, an ellipse or a hyperbola.

Denote by $\boldsymbol{\xi}$ the vector consisting of the eight elements of the array $\underline{\mathrm{X}}$. The Hessian matrix $\mathbf{H}$ of D is defined as

$$h_{ij} = \frac{\partial^2 D}{\partial \xi_i \partial \xi_j} \tag{27}$$

The eigenvalues of $\mathbf{H}$ were evaluated at the top of the parabola of degeneracy, i.e. at the point defined by $d = h = 0$ in Equation (21). The numbers of positive, zero and negative eigenvalues of $\mathbf{H}$ were found to be three, three and two respectively. It is seen that the 'top of the parabola' is in fact a saddle point of the surface $D(\boldsymbol{\xi}) = 0$ in the full eight-dimensional space of the elements of $\underline{\mathrm{X}}$.

5. DEGENERATE MODELS AND SWAMPS

The word *swamp* has been used to denote such areas of the solution space where the alternating least squares (ALS) algorithm advances extremely slowly, as if dragging its feet in mud [11]. It is well known that such swamps are connected with degenerate models. It is easy to understand that if the array $\underline{\mathrm{X}}^0$ to be approximated by the CP model is degenerate (or in the higher-rank domain), then any algorithm trying to find the best (lower-rank) approximation $\underline{\mathrm{X}}$ is in trouble. The factor matrices become ill-conditioned when the solution approaches the border of the high-rank domain. Larger and larger changes in the factor matrix elements are required in order to achieve smaller and smaller improvements in the object function. The path of the iteration is driven to end in a swamp! Furthermore, in most cases the degenerate solution is not meaningful for the original problem. It will be necessary to modify the problem somehow in order to avoid creating degenerate solutions. This modifying is intimately connected with the real-life situation that is being modeled. Discussion of this problem goes outside of the present paper.

It is also well known that swamp behavior may occur midway between a non-degenerate starting point and a well-defined non-degenerate best-fit solution of a CANDECOMP/PARAFAC problem [12]. It has been observed that the condition numbers of factor matrices increase along the path and then decrease back to 'normal' levels when the best-fit solution is approached. Rayens and Mitchell [12] report that the severity of half-way swamps can be lessened by applying *regularization* (see below) to the model. The experience of Paatero [13] also supports this finding, although he did not explicitly investigate the question: the algorithm PMF3 is normally used with initial regularization so that the regularization is only decreased after a 'reasonably good' approximation has been found. Half-way swamps have never been a problem with PMF3.

By using the synthetic degenerate models, it is possible to illustrate what happens when the iteration encounters a degenerate solution area. In the following section, examples are created where

the parabola-shaped degenerate area extends between the initial and true solutions. Depending on the geometry of the model, either the iteration is just slowed down, or the iteration is unable to reach the true solution because it cannot get around the swamp.

6. NUMERICAL EXPERIMENTS

When discussing numerical results, the final convergence of the ALS algorithm is mentioned repeatedly. Experience shows that the final convergence of ALS is geometric when the array $\underline{\mathbf{X}}^0$ is not degenerate. Denote by $\underline{\mathbf{X}}^1$ the solution of the iterative computations after 'infinitely many' iterations. (For the present non-degenerate $2 \times 2 \times 2$ cases, $\underline{\mathbf{X}}^1 = \underline{\mathbf{X}}^0$.) Each iteration step reduces the remaining difference $\underline{\mathbf{X}} - \underline{\mathbf{X}}^1$ by a factor $r = 1 - k$, where k is a small constant. For difficult cases the value of the convergence decrement k is typically (much) below 0·01.

The numerical experiments serve to illustrate why the computed solution of a CP model sometimes becomes degenerate even when the true array $\underline{\mathbf{X}}^0$ has rank 2. Intuitively, one might expect such behavior whenever the starting point $\underline{\mathbf{X}}$ and the correct solution $\underline{\mathbf{X}}^0$ are situated so that there is a domain of degeneracy between them. Such a situation is most pronounced when the parabola has a peaked shape and $\underline{\mathbf{X}}$ and $\underline{\mathbf{X}}^0$ are on opposite sides of the parabola. The actual value of the constant e determines the shape of the parabola. With small values, such as $e = 1$, the shape is shallow. A peaked shape is obtained with large values of e. The value $e = 30$ is chosen for all the following computations.

For most computations the free parameters of the array $\underline{\mathbf{X}}^0$ defined by Equations (21) and (25) were chosen as $\epsilon = 0{\cdot}54$ and $d = 0{\cdot}25$. For the other parameters one then obtains $h = -0{\cdot}28$, $p = r = 0{\cdot}37$, $v = x = -0{\cdot}086$, $t = -2{\cdot}57$ and $z = 11{\cdot}3$. The true solution is seen as the mark '$\times$' on the right side of the parabola. During the computations the twelve elements of the factor matrices are not constrained in any way. The numerical values shown in Equation (25) are only used when specifying the initial and final arrays $\underline{\mathbf{X}}$ and $\underline{\mathbf{X}}^0$.

The initial solution, i.e. the starting point of the iterations, was placed on the left side of the parabola, at $\epsilon = 0{\cdot}6$, $d = -0{\cdot}23$, $h = -0{\cdot}047$. Figure 1 shows the path of the solution obtained by using the original ALS-based CANDECOMP/PARAFAC algorithm without any acceleration techniques. The starting point is marked by a plus sign. Every tenth iterative solution is marked by an open circle. The solution is free to move in all directions in the twelve-dimensional solution space. The other ten dimensions may be imagined as being perpendicular to the (d, h) plane of the picture. This example has been constructed so that these extra dimensions do not play a decisive role.

Four different stages can be discerned in the progress of the iteration. During the first 50 steps there is a rapid advance towards the degenerate domain. During the next 150 steps the solution advances slowly. The direction of the 'pull' towards the true solution is almost orthogonal to the barrier that prevents the direct approach towards the true solution. The pull causes the advancing solution to approach closer to the parabola of degeneracy, becoming more and more degenerate. The third stage begins when the solution has passed the top of the parabola. Gradually the pull becomes parallel to the allowed path. The steps become longer, until the solution, after 340 steps, comes almost to a standstill between the true solution and the parabola of degeneracy. The fourth stage, not shown in Figure 1, corresponds to the final convergence towards the true solution. This convergence is extremely slow, because the solution still possesses the degeneracy caused during the second stage of the iteration. The initial value of the convergence decrement k is 0·00007 during the fourth phase. Approximately 4400 steps were needed for covering half of the distance towards solution. Another halving of the remaining distance required only 1500 steps, because the solution was gradually leaving the degenerate area. Close to $\underline{\mathbf{X}}^0$ the decrement had the value 0·00066. It is seen that the decreasing degeneracy allows the geometrical convergence rate to improve almost by a factor of ten during the fourth stage.

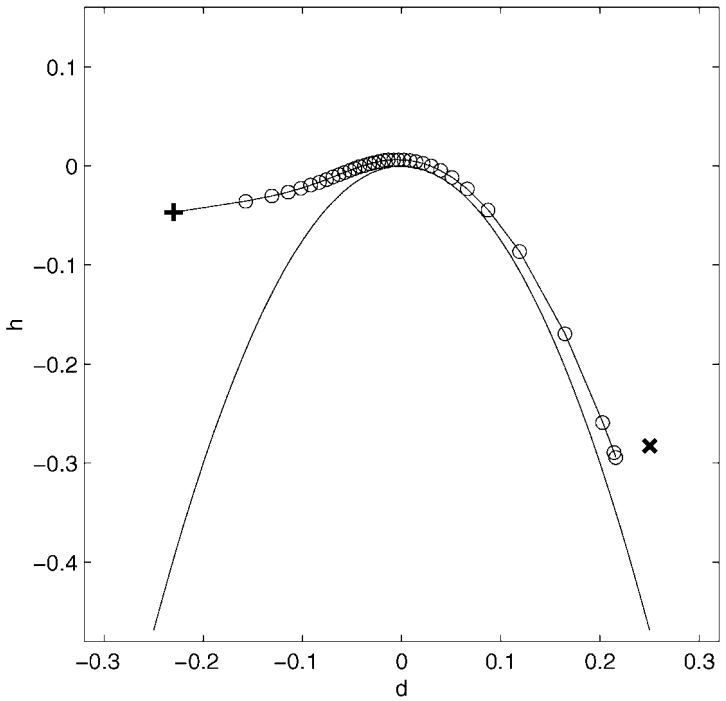

Figure 1. Graphical presentation of the array $\underline{\mathrm{X}}$, as defined by Equation (21), and the 340 first steps of the ALS iteration. The elements of $\underline{\mathrm{X}}$ are designated by letters *a* to *h* according to Equation (18). The abscissa and ordinate correspond to array elements *d* and *h* respectively. The parabola represents the intersection of the degenerate hypersurface with the hyperplane ($a = f = g = 0$, $b = c = 1$, $e = 30$). The points below and above the parabola correspond to rank = 3 and rank = 2 arrays respectively. The starting point (+) and the true solution (×) are in the plane of the figure. The path of the ALS iteration (every tenth solution shown by open circles) is in general *not* in the plane of the figure.

Figure 2 shows the path of the solution when using the program PMF3 based on the Gauss–Newton algorithm [13]. In contrast to the previous figure, each iterative solution is shown here. After 53 steps the remaining distance to the solution corresponds to the result obtained with ALS in 11 000 steps. The four stages can be vaguely discerned in Figure 2 too. There is a marked difference in the second stages of PMF3 and ALS. With PMF3 the solution keeps a certain distance to the parabola of degeneracy. This is caused by the regularization that is always present when using PMF3: the object function *Q*, minimized during the iteration, is of the form

$$Q = \|\underline{\mathrm{X}} - \underline{\mathrm{X}}^0\|_{\mathrm{F}}^2 + \lambda(\|\mathbf{A}\|_{\mathrm{F}}^2 + \|\mathbf{B}\|_{\mathrm{F}}^2 + \|\mathbf{C}\|_{\mathrm{F}}^2) \tag{28}$$

where $\underline{\mathrm{X}}$ is the current approximation of $\underline{\mathrm{X}}^0$ and λ is a user-defined parameter specifying the strength of regularization. Approaching the parabola of degeneracy would increase the numerical values of elements in one or several of the factor matrices **A, B** and **C**, leading to an increase in the second term in *Q*. The solution will stay at a distance where an increase in the second term just balances the decrease in the first term $\|\underline{\mathrm{X}} - \underline{\mathrm{X}}^0\|_{\mathrm{F}}^2$. Thus avoiding the more degenerate solutions is one of the reasons for the fast convergence of PMF3 [14].

The Jacobian matrix **J** of the solution was evaluated at the true solution. Details of this technique

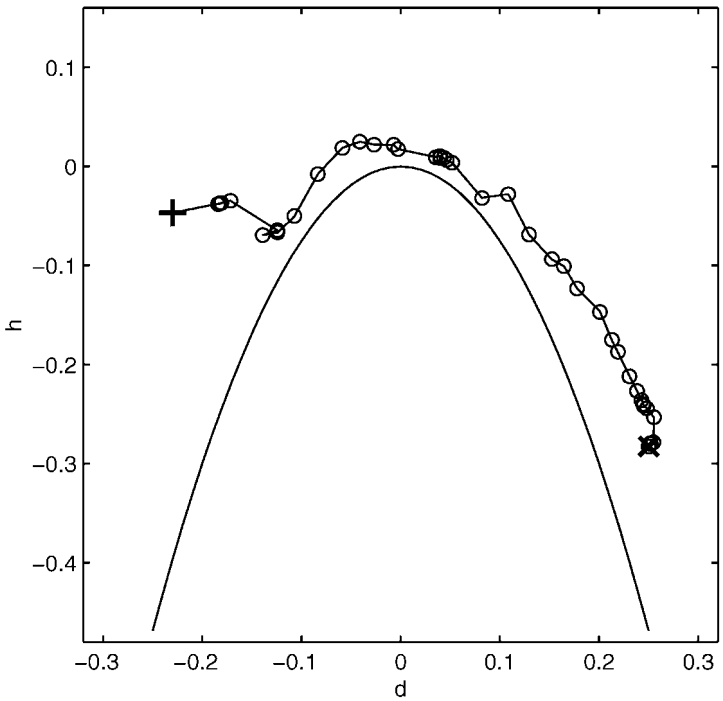

Figure 2. The path of the PMF3 iteration from the starting point (+) to the true solution (×). The 53 iterative solutions are shown by open circles. For other details see caption of Figure 1.

have been described by Paatero [15]. The dimensions of **J** are 8×12, corresponding to eight elements in array $\underline{\mathrm{X}}$ which depend on twelve elements in the three factor matrices **A**, **B** and **C**. There are four 'dead' directions in the space spanned by the twelve factor elements, corresponding to trivial scale changes of the factors. A change in the solution in any one of these dead directions does not change the array at all. Corresponding to the dead directions, four of the singular values of **J** equal zero. These four singular values are ignored in the following. The effective condition number of **J** (the ratio of largest and smallest non-zero singular values) was found to be 1926, meaning that there is one direction in the twelve-dimensional space where the rate of decrease in Q is extremely low. The left and right singular vectors corresponding to the smallest singular value were examined. The left singular vector corresponds to changes in elements of $\underline{\mathrm{X}}$. This vector was found to be practically parallel to the change in $\underline{\mathrm{X}}$ during the fourth stage of ALS. Similarly, the right singular vector corresponds to changes in elements of the three factor matrices **A**, **B** and **C**. It was found that the change in the factors during the fourth stage of ALS iteration was approximately orthogonal to the right singular vectors one through seven. It is seen that the final convergence of $\underline{\mathrm{X}}$ in the CP-ALS algorithm happened in the direction of the least significant singular vectors of **J**. In addition, the factors did also change in the dead directions, represented by the right singular vectors nine to twelve. For comparison, the effective condition number of **J** was also evaluated for the strongly degenerate result of the third phase of ALS. The value of 4756 was obtained, more than twice the value obtained at the true solution.

Different variations of the previous examples were computed as follows. The initial regularization parameter λ was increased for PMF3. This led to slightly faster convergence. However, the overall

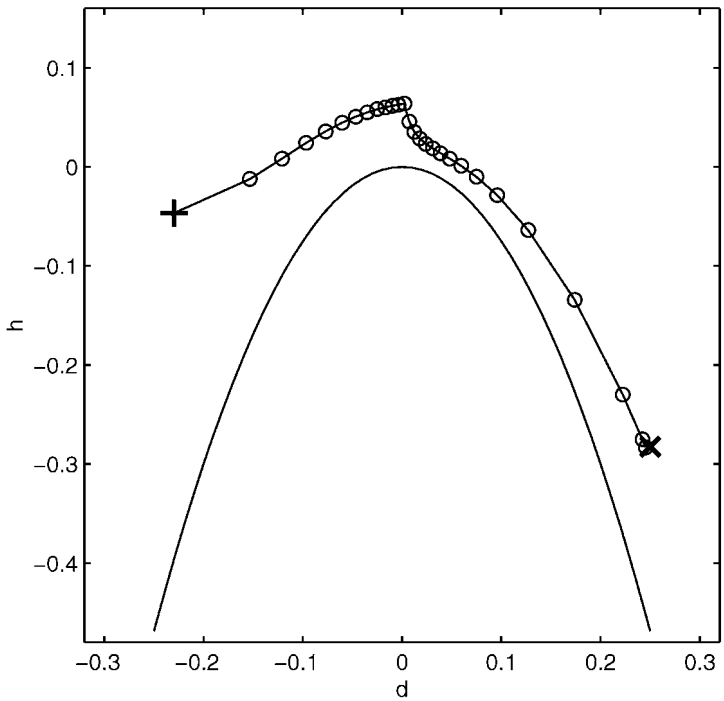

Figure 3. The path of the 270 first steps of the ALS iteration with initial regularization. The regularized form of the object function, shown in Equation (28), was used during the first 120 iteration steps. The basic ALS was used in all the remaining steps. Every tenth solution is shown as an open circle. For other details see caption of Figure 1.

behavior was similar to Figure 2. Also, a special regularized version of the ALS algorithm was tested so that the first 120 steps were computed with regularization, the rest without. The path of the solution is shown in Figure 3. The cusp of the path is at the step when the regularization was removed. The iteration reached the neighborhood of the true solution in 270 steps instead of the original 340 (cf. Figure 1). By chance, the third phase arrived very close to the true solution, so the distance to be covered by the fourth phase (not shown) was quite short. The final convergence decrement was of course identical to the non-regularized ALS run. A regularized step of ALS [12] may be called *ridge regression*. This term is often used when solving ill-conditioned regression problems so that the sum of squares of regression coefficients is included in the object function.

The acceleration mechanism included in the ALS algorithm of Bro was also tested, by using the default settings of the distributed algorithm [16]. A clear improvement was observed. The first three phases of the iteration were completed in only 120 steps, one-third of the original 340. At the beginning of the fourth phase the convergence decrement was $k = 0{\cdot}0003$, increasing to $k = 0{\cdot}0035$ towards the end of the fourth phase. These values represent a fivefold increase with respect to the non-accelerated tests.

Figure 4 shows another example. The true and initial solutions were defined by $(d, h) = (0{\cdot}26, -0{\cdot}39)$ and $(d, h) = (-0{\cdot}26, -0{\cdot}39)$ respectively. Now the top of the parabola reaches so far up that the iterative solution is unable to get around it. For both ALS and PMF3 the solution was stuck at the left side of the parabola. Figure 4 shows the path of the PMF3 solution.

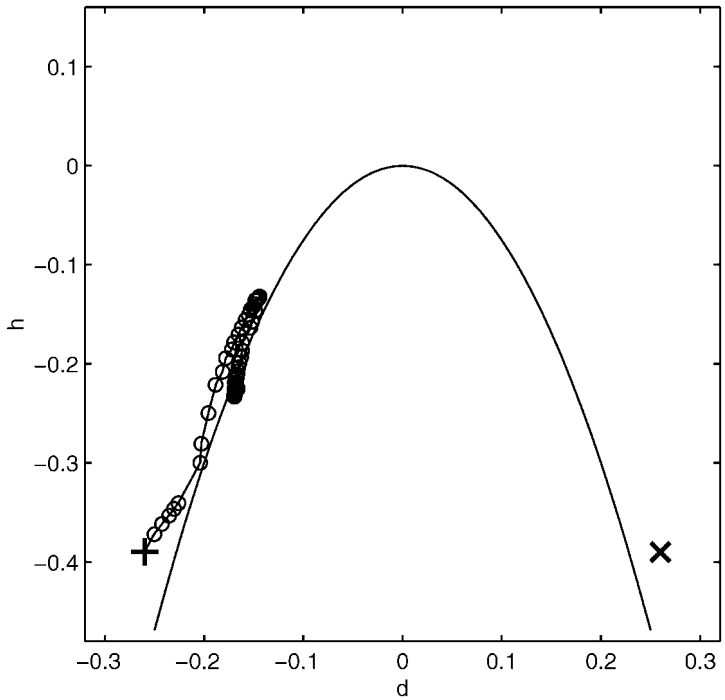

Figure 4. Applying the PMF3 algorithm to a slightly different problem, where the initial (+) and true (×) solutions are situated deeper down in the (*d*,*h*) plane. Other elements of $\underline{\mathrm{X}}$ are the same as in Figure 1. Instead of converging towards the true solution, the iteration plunges into the 'swamp', i.e. towards the degenerate area. The first 103 steps are shown as open circles.

7. HIGHER-ORDER DEGENERACIES

It was possible to construct the following non-trivial example of three-factor degeneracy. In this example all the loadings of the three factors grow without limit so that the array approaches a finite limit. It is assumed that all the sets $\{\mathbf{a},\bar{\mathbf{a}},\boldsymbol{\alpha}\}$ $\{\mathbf{b},\boldsymbol{\beta},\boldsymbol{\mu}\}$ and $\{\mathbf{c},\boldsymbol{\gamma},\boldsymbol{\nu}\}$ are linearly independent. The model is defined by setting

$$\begin{aligned}\underline{\mathrm{X}} &= [\mathbf{A},\mathbf{B},\mathbf{C}]\\ &= \left[\left(-\frac{1}{\epsilon}\mathbf{a}-\frac{1}{\epsilon}\bar{\mathbf{a}}\middle|\frac{1}{\epsilon}\mathbf{a}+\frac{\epsilon^2}{2}\boldsymbol{\alpha}\middle|\frac{1}{\epsilon}\bar{\mathbf{a}}\right),\left(-\frac{1}{\epsilon}\mathbf{b}\middle|\frac{1}{\epsilon}\mathbf{b}+\frac{\epsilon^2}{2}\boldsymbol{\beta}\middle|\frac{1}{\epsilon}\mathbf{b}+\frac{\epsilon^2}{2}\boldsymbol{\mu}\right),\right.\\ &\quad\left.\left(-\frac{1}{\epsilon}\mathbf{c}\middle|\frac{1}{\epsilon}\mathbf{c}+\frac{\epsilon^2}{2}\boldsymbol{\gamma}\middle|\frac{1}{\epsilon}\mathbf{c}+\frac{\epsilon^2}{2}\boldsymbol{\nu}\right)\right]\end{aligned} \tag{29}$$

With $\epsilon > 0$ the rank of this array $\underline{\mathrm{X}}$ is obviously 3, because all three factor matrices are of full rank. Performing the computations gives

$$\underline{\mathrm{X}} = [\mathbf{a},\mathbf{b},\boldsymbol{\gamma}]+[\mathbf{a},\boldsymbol{\beta},\mathbf{c}]+[\boldsymbol{\alpha},\mathbf{b},\mathbf{c}]+[\bar{\mathbf{a}},\mathbf{b},\boldsymbol{\nu}]+[\bar{\mathbf{a}},\boldsymbol{\mu},\mathbf{c}]+O(\epsilon^3) \tag{30}$$

The limit $\underline{\mathrm{X}}^*$ at $\epsilon \to 0$ is obtained in a five-factor representation by simply omitting the last (sixth)

term in the previous equation. The corresponding core array is

$$\underline{\mathrm{G}}^{*} = \underline{\mathrm{G}}(\epsilon \to 0) = \left(\begin{array}{ccc|ccc|ccc} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{31}$$

As shown in the next subsection, the rank of this array is in fact 5. It is easily seen that the presentation of the limit array $\underline{\mathrm{X}}^{*}$ is not unique: rotations are possible between such pairs of factors which are identical in one mode. An example of a rotated presentation of the limit array is

$$\underline{\mathrm{X}} = [\mathbf{a}, \mathbf{b}, (\boldsymbol{\gamma} + \mathbf{c})] + [\mathbf{a}, (\boldsymbol{\beta} - \mathbf{b}), \mathbf{c}] + [\boldsymbol{\alpha}, \mathbf{b}, \mathbf{c}] + [\overline{\mathbf{a}}, \mathbf{b}, \boldsymbol{\nu}] + [\overline{\mathbf{a}}, \boldsymbol{\mu}, \mathbf{c}] \tag{32}$$

Four-factor degeneracy is most easily constructed with $4 \times 4 \times 4$ arrays as a superposition of two two-factor degeneracies. Construct the array

$$\underline{\mathrm{X}}^{h} = [\boldsymbol{\alpha}, \mathbf{b}, \mathbf{c}] + [\mathbf{a}, \boldsymbol{\beta}, \mathbf{c}] + [\mathbf{a}, \mathbf{b}, \boldsymbol{\gamma}] + h[\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}] \tag{33}$$

so that the generating vectors have the values $\mathbf{a} = \mathbf{b} = \mathbf{c} = (1,0,0,0)$ and $\boldsymbol{\alpha} = \boldsymbol{\beta} = \boldsymbol{\gamma} = (0,1,0,0)$. Similarly construct

$$\underline{\mathrm{X}}^{g} = [\boldsymbol{\alpha}, \mathbf{b}, \mathbf{c}] + [\mathbf{a}, \boldsymbol{\beta}, \mathbf{c}] + [\mathbf{a}, \mathbf{b}, \boldsymbol{\gamma}] + g[\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}] \tag{34}$$

with vectors $\mathbf{a} = \mathbf{b} = \mathbf{c} = (0,0,1,0)$ and $\boldsymbol{\alpha} = \boldsymbol{\beta} = \boldsymbol{\gamma} = (0,0,0,1)$. Depending on the values of the scalar coefficients h and g, the rank of the array $\underline{\mathrm{X}} = \underline{\mathrm{X}}^{h} + \underline{\mathrm{X}}^{g}$ may equal 4, 5 or 6. If both coefficients are zero, the array is doubly degenerate: it has rank 6 but it may be approximated arbitrarily well with arrays having rank 4. If both coefficients are positive, the array has rank 4.

7.1. Rank of the array $\underline{\mathrm{G}}^{*}$

Several tools for estimating ranks of arrays are given in Reference [5]. Corollary 1′ on p. 108 is used for showing that the rank of the array $\underline{\mathrm{G}}^{*}$ in Equation (31) is 5. The corollary estimates the rank of an array with the help of an auxiliary row vector z that may be freely chosen to best suit the task. Denote the mode n rank of an array by rank_n. Assuming that a given array $\underline{\mathrm{Z}}$ is 3-nondegenerate, the corollary provides a lower limit for its rank:

$$\mathrm{rank}(\underline{\mathrm{Z}}) \geq \min_{\{\mathbf{w} | \mathbf{w} \cdot \mathbf{z} \neq 0\}} (\mathrm{rank}(\underline{\mathrm{Z}} \times_3 \mathbf{w})) + \mathrm{rank}_3(\underline{\mathrm{Z}}) - 1 \tag{35}$$

Set $\underline{\mathrm{Z}} = \underline{\mathrm{G}}^{*}$ and choose $\mathbf{z} = (1\,0\,0)$. Then the condition for $\mathbf{w}$ in Equation (35) simplifies to requiring that in $\mathbf{w} = (\alpha,\beta,\gamma)$ the first element α is non-zero while the other two are arbitrary. The product $\underline{\mathrm{G}}^{*} \times_3 \mathbf{w}$ evaluates to

$$\underline{\mathrm{G}}^{*} \times_3 \mathbf{w} = \begin{pmatrix} \beta & \alpha & 0 \\ \gamma & 0 & \alpha \\ \alpha & 0 & 0 \end{pmatrix} \tag{36}$$

The determinant of this matrix is $\alpha^3 \neq 0$, independent of β and γ. Thus the first term in Equation (35) is 3. It is easily seen that $\mathrm{rank}_3(\underline{\mathrm{G}}^{*}) = 3$. Substituting into (35) gives $\mathrm{rank}(\underline{\mathrm{G}}^{*}) \geq 3 + 3 - 1 = 5$, which is the desired result.

8. CONCLUSIONS

The examples shown in this work demonstrate the structure of a degenerate three-way array and facilitate an easy construction of various degenerate examples, e.g. for testing algorithms.

It was seen that some amount of regularization helps the algorithms in keeping distance to the *swamps*, i.e. to the degenerate regions where convergence is slow.

It was demonstrated that sometimes the degenerate regions create 'shadows' in the solution space. Although the true solution is non-degenerate, it cannot be reached from a starting point lying in such a shadow. If the iteration is started so that the initial solution is in the shadow, the algorithms get stuck, progressing deeper and deeper into the swamp. In such situations the only remedy is to repeat the computations starting from different pseudorandom initial solutions.

Earlier it had been reported that slow and fast phases may alternate in the progress of an ALS iteration [12]. In this work the nature of these phases is illustrated by an example displaying four distinct phases with fast, slow, fast and extremely slow convergence respectively. The 'simple' task of fitting a $2 \times 2 \times 2$ array displays a surprising richness of features!

ACKNOWLEDGEMENTS

Helpful discussions with Joseph Kruskal, Richard Harshman, Marg Lundy and Donald Burdick are gratefully acknowledged. This work has been supported by a research grant from the University of Helsinki.

REFERENCES

1. Carroll JD, Chang JJ. Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart–Young' decomposition. *Psychometrika* 1970; **35**: 283–319.
2. Harshman RA. Foundations of the PARAFAC procedure: models and conditions for an 'explanatory' multimode factor analysis. *UCLA Working Papers Phonet*. 1970; **16**: 1–84.
3. Harshman RA, Lundy ME. PARAFAC: parallel factor analysis. *Comput. Statist. Data Anal.* 1994; **18**: 39–72.
4. Kiers HAL. Towards a standardized notation and terminology in multiway analysis. *J. Chemometrics* 2000; **14**: 105–122.
5. Kruskal JB. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Appl.* 1977; **18**: 95–138.
6. Kruskal JB. Multilinear methods. In *Research Methods for Multimode Data Analysis*, Law HG, Snyder Jr CW, Hattie JA, McDonald RP (eds). Praeger: New York, 1984; 36–62.
7. Kruskal JB. Rank, decomposition, and uniqueness for 3-way and N-way arrays. In *Multiway Data Analysis*, Coppi R, Bolasco S (eds). Elsevier: Amsterdam, 1989; 7–18.
8. Kruskal JB, Harshman RA, Lundy ME. How 3-MFA data can cause degenerate PARAFAC solutions, among other relationships. In *Multiway Data Analysis*, Coppi R, Bolasco S (eds). Elsevier: Amsterdam, 1989; 115–122.
9. Ten Berge JMF. Kruskal's polynomial for $2 \times 2 \times 2$ arrays and a generalization to $2 \times n \times n$ arrays. *Psychometrika* 1991; **56**: 631–636.
10. Tucker LR. The extension of factor analysis to three-dimensional matrices. In *Contributions to Mathematical Psychology*, Gulliksen H, Frederiksen N (eds). Holt, Rinehart and Winston: New York, 1964; 110–127.
11. Mitchell BC, Burdick DS. Slowly converging PARAFAC sequences: swamps and two-factor degeneracies. *J. Chemometrics* 1994; **8**: 155–168.
12. Rayens WS, Mitchell BC. Two-factor degeneracies and a stabilization of PARAFAC. *Chemometrics Intell. Lab. Syst.* 1997; **38**: 173–181.
13. Paatero P. A weighted non-negative least squares algorithm for three-way 'PARAFAC' factor analysis. *Chemometrics Intell. Lab. Syst.* 1997; **38**: 223–242.
14. Hopke PK, Paatero P, Jia H, Ross RT, Harshman RH. Three-way (PARAFAC) factor analysis: examination and comparison of alternative computational methods, as applied to ill-conditioned data. *Chemometrics Intell. Lab. Syst.* 1998; **43**: 25–42.

15. Paatero P. The multilinear engine — a table-driven least squares program for solving multilinear problems, including the n-way parallel factor analysis model. *J. Comput. Graph. Statist.* 1999; **8**: 854–888.
16. Bro R. [Internet]. Available: http://newton.foodsci.kvl.dk/srccode.html [1998].

