Three Way PCA: Definitions and Properties

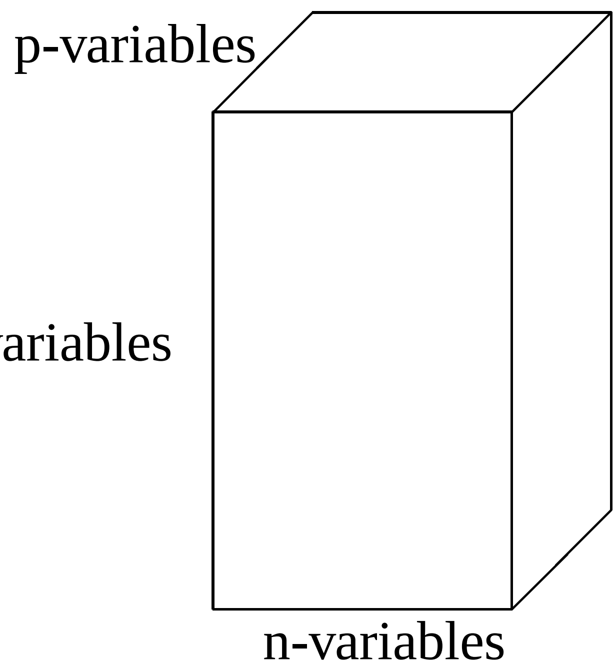
Introduction

- We understand how to decompose and analyze an $n \times p$ matrix X by Singular Value Decomposition (SVD) to define latent factors of variables within the existing structure.

- We would like to extend this well understood methodology to larger data structures, the simplest being a 3-dimensional $n \times m \times p$ array.

- Unfortunately not all of the nice properties of SVD hold. So we need to address what type of information we would like to obtain.
  - The MPCA Model
  - The PARAFAC Model
  - The Tucker3 Model

- The general three way structure is typically pictured as:



- A consideration one must make when analysing a data structure of this type is what sort of relationship exists in the dimensions. Deciding on this will warrant a particular kind of decomposition. Two major classifications are
  - Object Variation
  - System Variation
- Determining which type you have will determine what model you should choose.

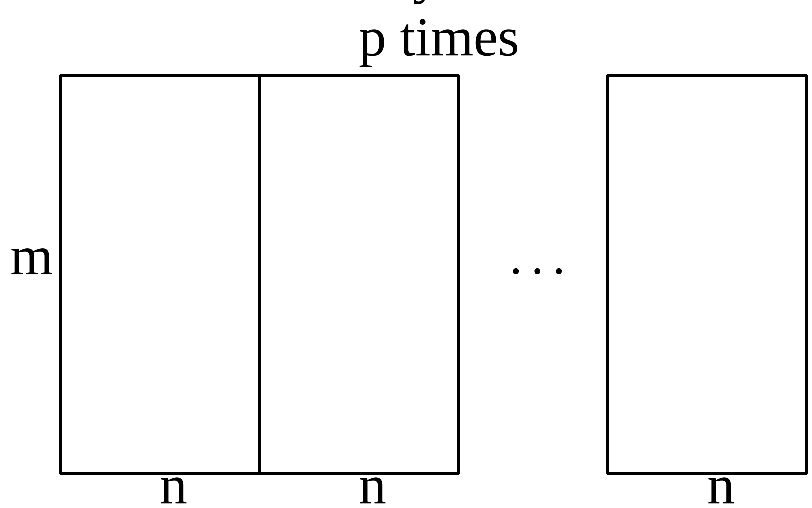
Object Variation and System Variation

- Object variation and system variation are best described by considering the relationship the different dimensions have with each other.
- Object variation supposes that the different dimensions have different effects on their respective factors. This means that for a particular realization of one dimension (e.g. a matrix subset restricted to one element in the third dimension) will have factor structures for the other two dimensions that do not directly relate (i.e. are not proportional to) to any other factor realization.
- For example, suppose the n-variables in the array represent stimuli, the m-variables represent subjects being tested and the p-variables represent different conditions regarding the environment. If we suspect that the conditions will affect the responses on the stimuli (i.e. interact with) then this is an example of object variation.

- System variation would imply that although the factors of stimuli will change under different conditions, these differences will not change the structure of the factors derived by the stimuli. This can also be described by suggesting that the factors of stimuli are proportional to the conditions being administered.
- By suggesting one type of variation over the other will allow for a more comprehensive study to determine the different factors. The type of models considered will then be
- <u>Object Variation</u>
  - MPCA Model
  - Tucker3 Model
- <u>System Variation</u>
  - PARAFAC Model

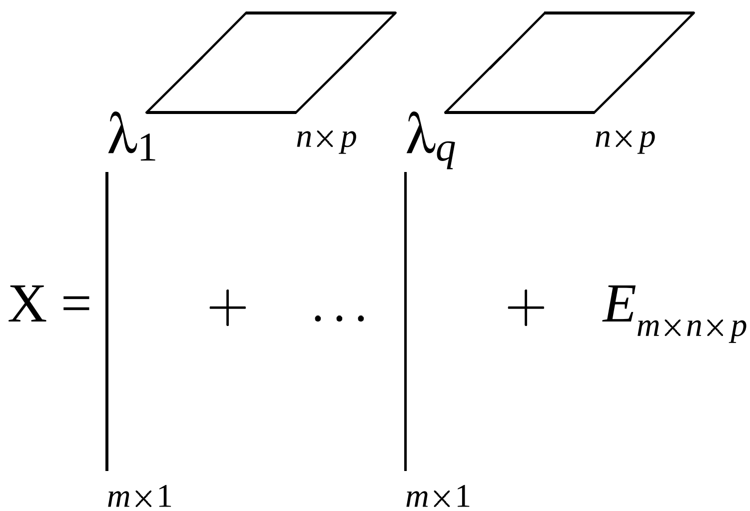
The MPCA Model

- Consider the general three-way set-up from Slide 2. If we break up the array by "unfolding" the p dimension into a large $m \times pn$ matrix we have a realization of an array which can be used for traditional PCA analysis.



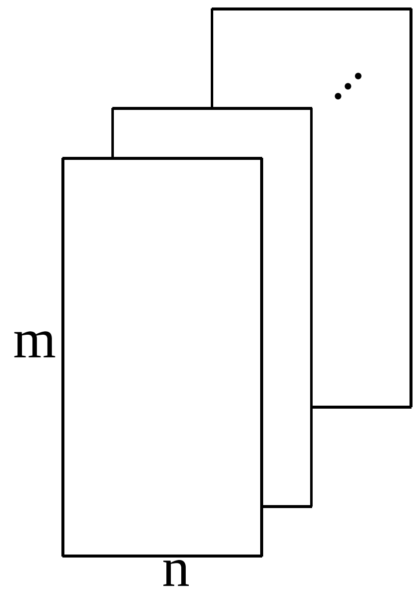
- This representation is not unique. By interchanging the matrix layout (i.e. if we unfolded with respect to n instead of p) will give identical results. This analysis lets us consider the factors regarding the m-variables as they interact with the $n \times p$ combinations.

- To describe the decomposition one usually runs the PCA analysis on the $m \times pn$ matrix. This gives the $m \times 1$ scores and the $pn \times 1$ loadings. The loadings are then refolded back to an $n \times p$ matrix to keep the 3 dimensional representation. Schematically, the decomposition looks like



The PARAFAC Model

- If a system variation is assumed in the factor structure of a data array then the PARAFAC Model can be applied to derive the results. The PARAFAC Model does give a unique solution, but concepts such as orthogonality of the factors must be compromised. Also the PARAFAC model requires that the number of factors considered be constant over the different dimensions.
- By compromising the factor orthogonality one also compromises the idea of nested models. What does this mean ? Basically it says that if I fit a model requiring two factors, then if I change to a three factor model, the first two factors will not necessarily be the same as the smaller model. Hence lack of orthogonality does not break down to nice additive pieces.
- The general set-up is as follows, since we expect the factor structures to remain intact over other variables factors I can describe the data structure in Slide 2 as



- Then each "slice" of the depth array is considered an $m \times n$ matrix which decomposition as (selecting the k^{th} depth X_k):

  $$X_k = AC_kB^t + E_k$$

  Where A is a $m \times q$ common factor matrix of the m variables, C_k a $q \times q$ diagonal matrix containing the q factors relevant to k^{th} depth and B the $n \times q$ common factor matrix of the n variables.

- Then by evaluating all slices we obtain three factor matrices A, B and C which represent the latent contributions from the three dimensions.

- To represent this schematically, one also creates a three dimensional identity "core"matrix $I_{q\times q\times q}$ with 1s on the super diagonal and writes X as

$$
\begin{array}{c}
\qquad\qquad\quad B_{q\times n} \\
\qquad\qquad\quad | \\
X_{m\times n\times p} = A_{m\times q} \overline{\quad\quad} I_{q\times q\times q} \overline{\quad\quad} C_{q\times p} + E_{m\times n\times p}
\end{array}
$$

- Here the factors are global. That is Factor 1 is a combination of $A_{m\times 1}$, $B_{1\times n}$, $C_{p\times 1}$, Factor 2 $A_{m\times 2}$, $B_{2\times n}$, $C_{p\times 2}$, etc.

- When using this type of decomposition one must evaluate the implications of mean centering or standardizing the data, since the basic principle being exploited is system variation.

The Tucker3 Model

- The Tucker3 Model is the most general model considered. It is being mentioned purely for completeness. Some may think of it as relaxation of the PARAFAC model, but it can be considered even more generally than that.
- While the PARAFAC requires both equal numbers of factors amongst dimensions and system variation, the Tucker3 relaxes on both.
- The Tucker3 decomposes $X_{m\times n\times q}$ by any number of factors from each dimension a, b and c. The Tucker3 considered has a schematic representation as

$$\begin{array}{ccccc} & & B_{b\times n} & & \\ & & | & & \\ X_{m\times n\times p} = A_{m\times a} & \text{———} & Z_{a\times b\times c} & \text{———} & C_{c\times p} + \quad E_{m\times n\times p} \end{array}$$

- The major difference between the Tucker3 and the PARAFAC (other than different factors among dimensions) is the introduction of the Z "core" matrix. The PARAFAC model constrained this to be an identity. The Tucker3 relaxes on this assumption thus allowing for factors to interact between dimensions.
- The Tucker3 model does place some constraints on the factors though. Each factor matrix A, B or C is columnwise independent (i.e. $A^t A = I_{a \times a}$). Hence information within the factors of a particular dimension are independent
- A good reference for this type of modelling is: H.G. Law, C.W. Snyder, J.A. Hattie, R.P. MacDonald *Research Methods for Multimode Data Analysis*. ,1984, Praeger Publishers. Evans Call Number: BF 39 R36 1984