



ving



b



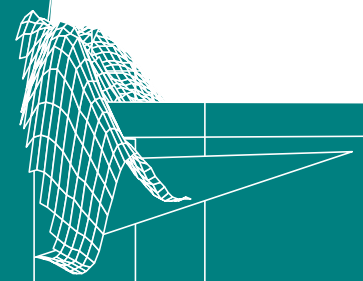
Ras

- Che
- Roya
- rb@kv



Nikos Sid

- Dept. of Ele
- University
- nikos@



Content



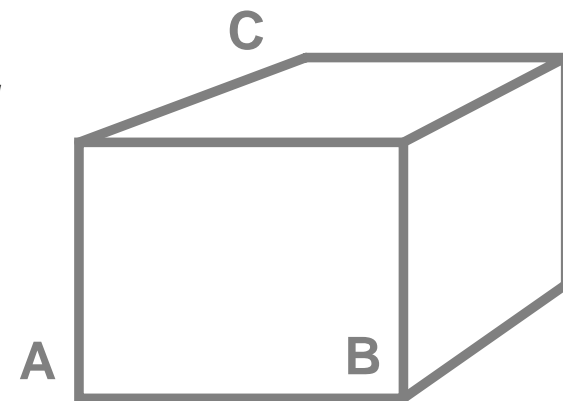
- Introduction
- Models & algorithms
- Applications
- Problems

■ Three-way data?

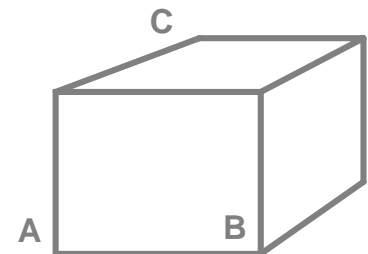
- Simply a set of two-way matrices
- Each mode consist of same basic entities over the other modes
- E.g. same samples measured at different variables, several times
- Instead of a matrix with typical elements x_{ij} , we have an array with elements x_{ijk}

■ Where?

- Sensory analysis, Process analysis, Image analysis, Experimental design, Spectroscopy, Chromatography, Environmental analysis, QSAR, Communication, Medicine, ...



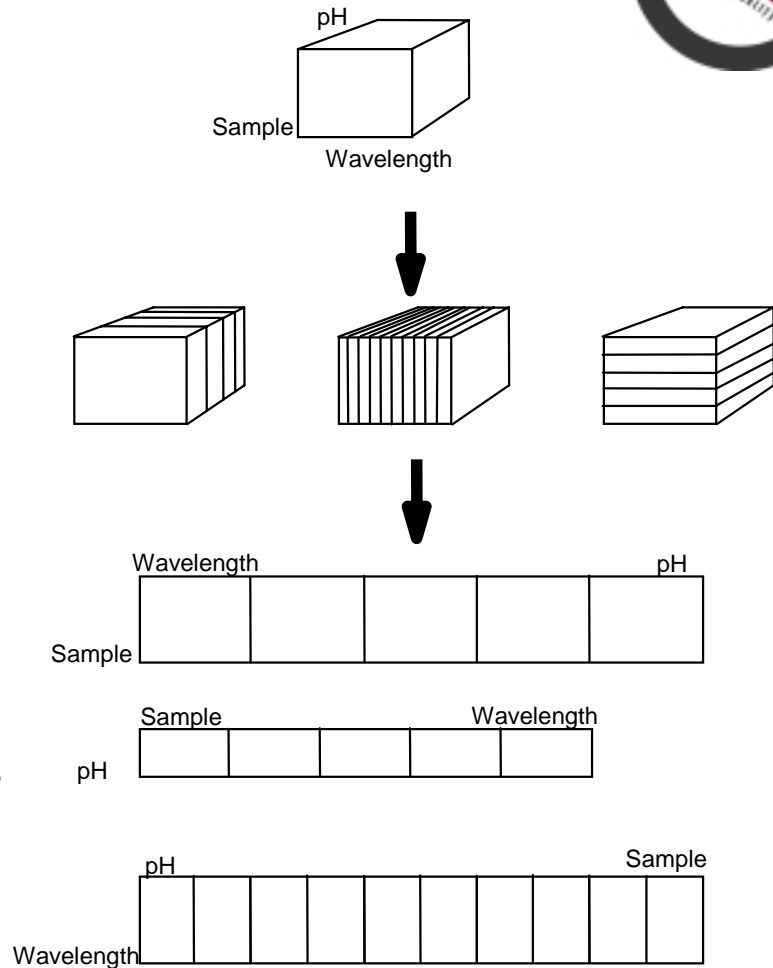
- **Sensory analysis**
 - Score as a function of (Food sample, Judge, Attribute)
- **Process analysis**
 - Measurement as a function of (Batch, Variable, time)
 - Measurement as a function of (Variable, Lag, Location)
- **Image analysis**
 - Pixelvalue as a function of (Sample, Image pixel, Variable)
- **Experimental design**
 - Response as a function of (factor 1, factor2, factor3,..)
- **Spectroscopy**
 - Intensity as a function of (Wavelength, Retention, Sample, Time, Location , Treatment)
- **Environmental analysis**
 - Measurement as a function of (Location, Time, Variable)



Unfolding/matricization



- Traditional approach
 - Unfolding leading to two-way data and analysis
- Three-way models
 - Natural extensions of two-way models
 - PCA leads to PARAFAC or Tucker3 depending on how it is extended
 - Rank-reduced regression, e.g. PLS leads to multilinear PLS (N-PLS)



Unfolding/matricization
Often leads to overfitting
because nature of model \neq
nature of the data

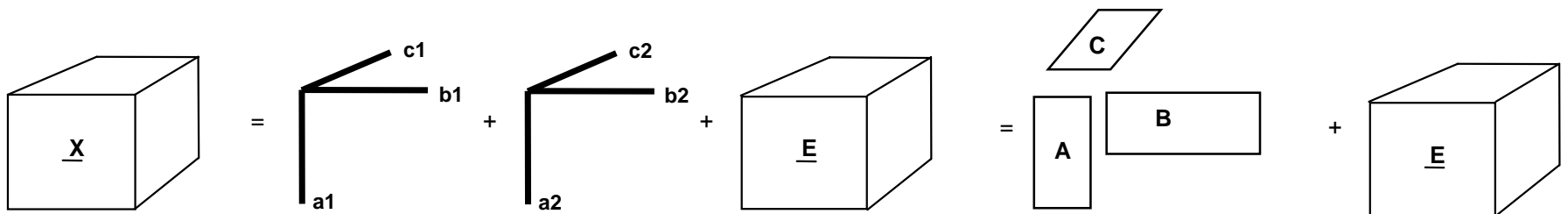
Unfold: Wold, Geladi, Esbensen, Öhman. Principal component- and PLS-analyses generalized to multi-way (multi-order) data arrays. Copenhagen Symposium on Applied Statistics:249-277, 1986.

■ PCA - bilinear model,

$$■ x_{ij} = \sum_f t_{if} p_{jf} + e_{ij}, \quad i=1, \dots, I; \quad j=1, \dots, J$$

■ PARAFAC - trilinear model,

$$■ x_{ijk} = \sum_f a_{if} b_{jf} c_{kf} + e_{ijk}, \quad i=1, \dots, I; \quad j=1, \dots, J; \quad k=1, \dots, K$$



PARAllel FACtor analysis



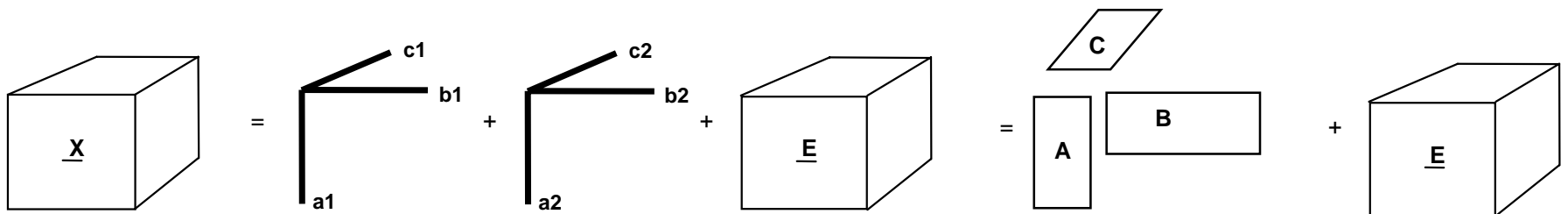
■ PCA - bilinear model,

$$\mathbf{X} = \mathbf{AB}' + \mathbf{E}$$

■ PARAFAC - trilinear model,

$$\mathbf{X}_k = \mathbf{AD}_k\mathbf{B}' + \mathbf{E}_k, k = 1, \dots, K$$

$$\mathbf{D}_k = \text{diag}(\mathbf{C}(k, :))$$

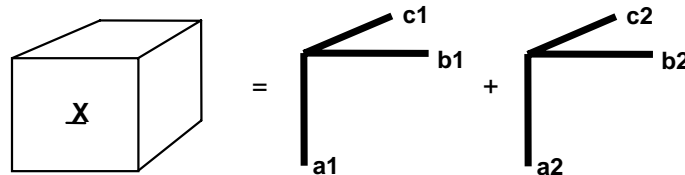


■ Rank of two-way matrix

- Minimum number of bilinear (PCA) components needed to reproduce matrix

■ Rank of three-way array

- Minimum number of trilinear (PARAFAC) components needed to reproduce array

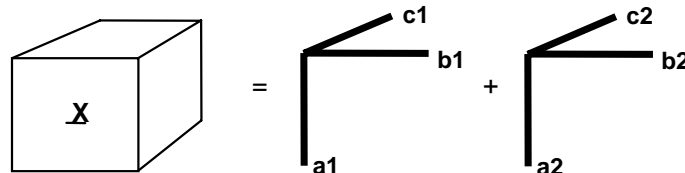


Two-way

- Any random matrix has full rank with probability one
- Row-rank = column-rank
- Simple rules for maximal rank

Three-way

- Random $2 \times 2 \times 2$ is rank 2 (30%) and rank 3 (70%)!
- Row-rank \neq column-rank \neq tube-rank!
- No rules for maximal rank!
 - Except simple cases such as $3 \times 3 \times 3$ is five.

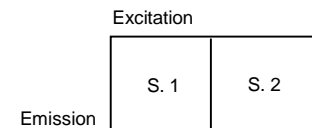
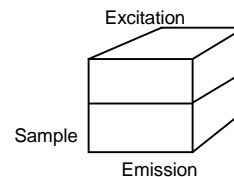
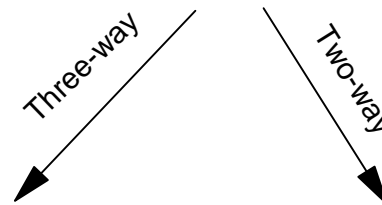
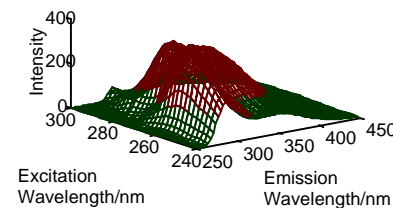
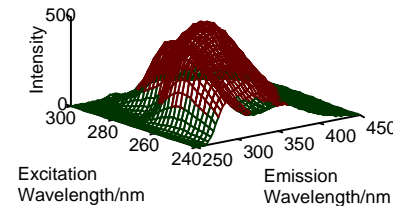


PARAFAC - uniqueness



■ PARAFAC has a unique property: It is unique!!

- Practical example: Fluorescence
- Two samples
- Mixture of three analytes
 - (Trp, Tyr, Phe)



PARAFAC invented in 1970 by Harshman and independently by Carroll & Chang under the name CANDECOMP. Based on a principle of parallel proportional profiles suggested in 1944 by Cattell

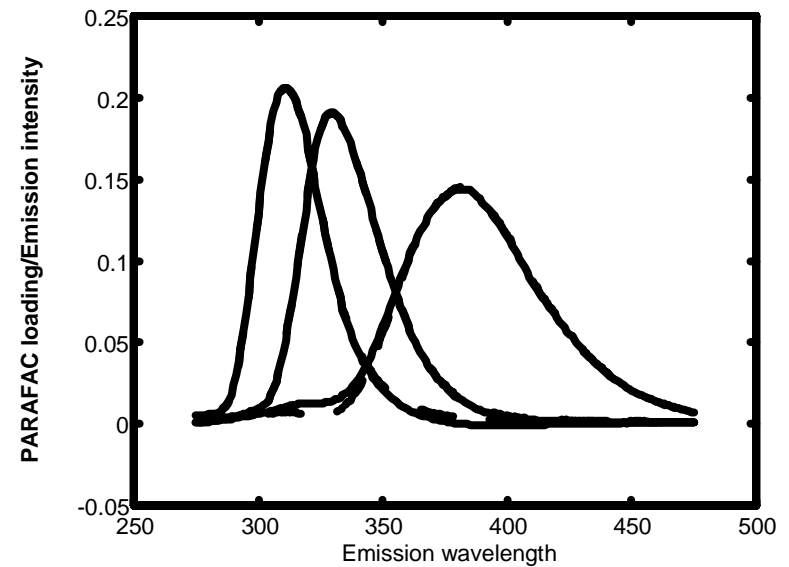
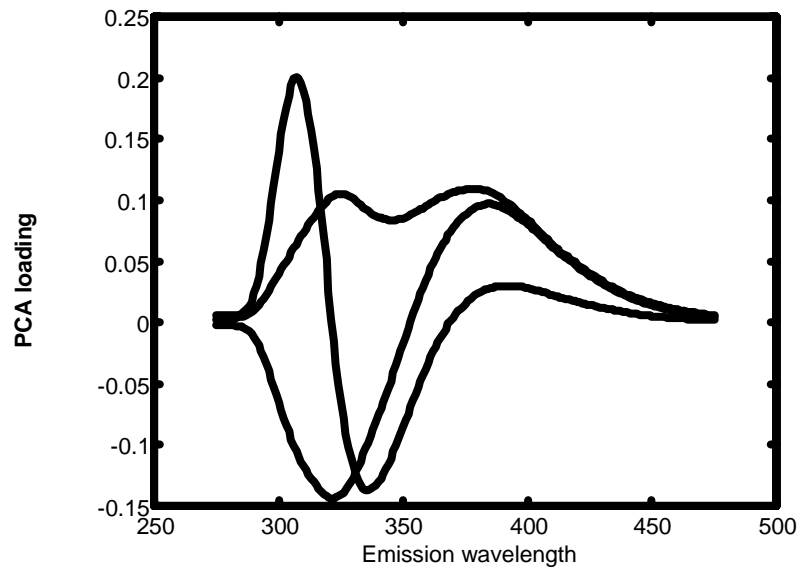
- R. A. Harshman. *UCLA working papers in phonetics* 16:1-84, 1970.
- J. D. Carroll and J. Chang. *Psychometrika* 35:283-319, 1970.
- R. B. Cattell. *Psychometrika* 9:267-283, 1944.

PARAFAC - uniqueness



■ PCA orthogonal loadings

PARAFAC pure spectra!



PARAFAC for CDMA

Blind CDMA

The diagram illustrates the decomposition of a matrix X and a tensor X in the context of Blind CDMA. The top part shows a 2D matrix X (square) equal to the product of a CODE matrix (blue) and a SYMBOL vector (red). The bottom part shows a 3D tensor X (cube) equal to the product of a CODE matrix (blue), a STEERING vector (green), and a SYMBOL vector (red).

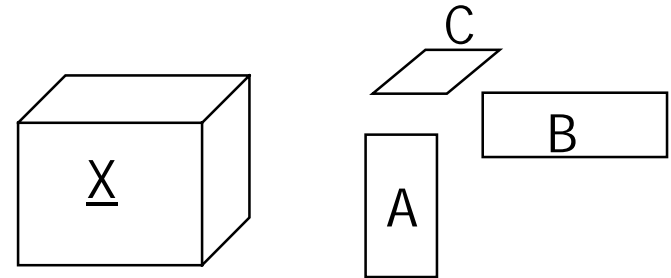
- Fact 1: Low-rank matrix (2-way array) decomposition not unique
- Fact 2: Low-rank 3- and higher-way array decomposition (**PARAFAC**) is unique

■ Uniqueness - conditions

A PARAFAC model is unique when

$$k_A + k_B + k_C \geq 2F + 2$$

F is the number of components and k_A is the k -rank of loading \mathbf{A} = maximal number of randomly chosen columns which will have full rank ($\leq F$)



J. B. Kruskal. *Linear Algebra and its Applications* 18:95-138, 1977.

SidGiaBro, IEEE TSP, Mar 2000,

SidBro. *JChemom* 14 2000,

$N = 3, \mathbb{R}$

any N, \mathbb{C}

$$\sum_{n=1}^N k_n \geq 2F + (N - 1)$$

■ Uniqueness - what does it mean?

- Mixtures of analytes can be separated
- Concentrations can be estimated
- Pure spectra and profiles can be estimated
 - $8 \times 8 \times 8$ array is unique for a 10 component model!!
 - Hence with 8 snapshots, 8 symbols and 8 codes, recovery of 10 sources is possible

■ Examples

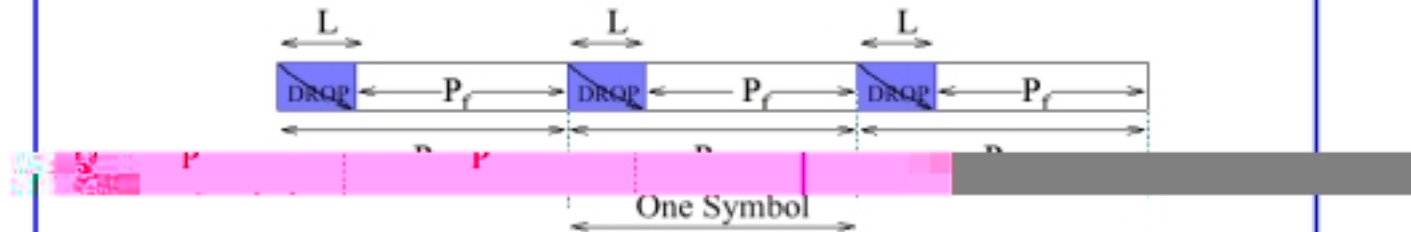
- On-line purity control (Kodak) Windig, *J. Magnetic Resonance* 132:298-306, 1998
- Determination of kinetic parameters Bijlsma, *AIChE Journal* 44:2713-2723, 1998
- Mathematical chromatography Leurgans, *Statistical Science* 7:289-319, 1992
- Localizing cellular phone calls Sidiropoulos & Bro. PARAFAC Techniques for Signal separation. In: *Signal Processing Advances in Communications*, edited by P. Stoica, G. B. Giannakis, Y. Hua, and L. Tong, Prentice-Hall, 2000,

Significance

- ☞ All codes, steering unknown; no need for training
- ☞ LS (cond. ML) → symbols, codes, steering for all users
- ☞ Iterative, $O(FIJK)$
- ☞ Diversity combining: non \perp codes, $F > K, F > I, F > J$
- ☞ Deterministic, modulation-independent
- ☞ Frequency-selective multipath? Performance?

Frequency-Selective Multipath

- ❑ (SS) Symbol-level (but not chip-level) synchronization
- ❑ (FF) Multipath reflectors in the far field
- ❑ (FIR) FIR channels, max order = L (incl. asynchronism)
- ❑ (DP) Spreading gain $K = K_f + L$, and $K_f \geq 2$ ISI-free chips (suffering from ICI only) are available (“discard prefix”):



e.g., [QS-CDMA-Iltis, Bensley-Aazhang, Liu-Xu, Tsatsanis et al]

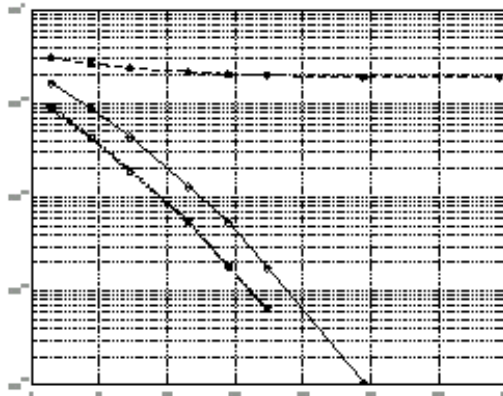
PARAFAC - uniqueness



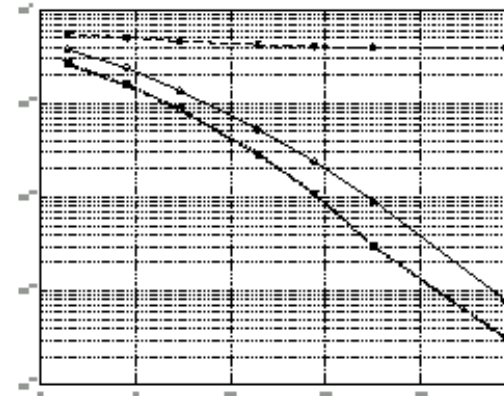
3SPICE-ECE-UMN

Nikos Sidiropoulos / May, 2001

Performance: $F = 4, I = 2, J = 50, K = 6$



DE-BPSK



DE-QPSK

- Blind performance close to non-blind MMSE

■ Alternating least squares algorithm

1. Initialize **B** and **C**

$$2. \mathbf{A} = \left(\sum_{k=1}^K \mathbf{X}_k \mathbf{B} \mathbf{D}_k \right) \{ (\mathbf{B}' \mathbf{B}) * (\mathbf{C}' \mathbf{C}) \}^{-1}$$

$$2. \mathbf{B} = \left(\sum_{k=1}^K \mathbf{X}'_k \mathbf{A} \mathbf{D}_k \right) \{ (\mathbf{A}' \mathbf{A}) * (\mathbf{C}' \mathbf{C}) \}^{-1}$$

$$3. \text{diag} \mathbf{D}_k = \{ (\mathbf{B}' \mathbf{B}) * (\mathbf{A}' \mathbf{A}) \}^{-1} \text{diag} (\mathbf{A}' \mathbf{X}_k \mathbf{B}), k = 1, \dots, K$$

4. Step 2 until relative change in fit is small

■ * Hadamard (elementwise product)

■ **Speed-up**

- Compression of data (LS or Spline-based bases)
- Line-search etc.

■ **Alternative algorithms**

- GRAM/DTLD – direct generalized eigenvalue problem but not LS
- PMF3 – Gauss-Newton: fast but problematic for large arrays
- COMFAC – Uses compression and separation + Gauss-Newton

■ Tucker3 – an extension of SVD

■ PCA formulated as truncated SVD – $\mathbf{X} = \mathbf{USV}' + \mathbf{E}$

■ *Problem:*

Express \mathbf{X} in terms of new unitary bases \mathbf{U} and \mathbf{V}

■ *Solution:*

Regress \mathbf{X} on \mathbf{U} and \mathbf{V} : $\mathbf{S} = \mathbf{U}^+\mathbf{XV}'^+ = \mathbf{U}'\mathbf{XV}$

■ *Hence:*

\mathbf{S} equals (approximates) \mathbf{X} in a new (truncated) coordinate system

■ *Curiosity:*

\mathbf{S} is diagonal, but that's not mandatory

- PCA = SVD can be written (note the off-diagonals)

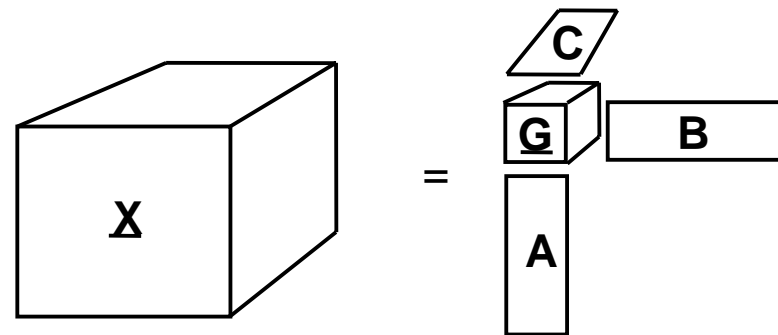
$$\begin{array}{|c|} \hline \mathbf{X} \\ \hline \end{array} = \begin{array}{|c|} \hline \mathbf{S} \\ \hline \end{array} \begin{array}{|c|} \hline \mathbf{B}^T \\ \hline \end{array} = \begin{array}{|c|} \hline \mathbf{A} \\ \hline \end{array} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix}^T$$

$$= \mathbf{a}_1 \mathbf{b}_1^T + 0 \mathbf{a}_2 \mathbf{b}_1^T + 0 \mathbf{a}_1 \mathbf{b}_2^T + 2 \mathbf{a}_2 \mathbf{b}_2^T \equiv \mathbf{a}_1 \mathbf{b}_1^T + 2 \mathbf{a}_2 \mathbf{b}_2^T$$

The Tucker3 model



- For three-way data, three unitary bases, **A**, **B**, and **C**; one for each mode
- Tucker3 is $\mathbf{X} = \mathbf{AG}(\mathbf{C} \otimes \mathbf{B})' + \mathbf{E}$
- Loadings are truncated bases and **G** the representation of **X** in these reduced spaces



•L. R. Tucker. The extension of factor analysis to three-dimensional matrices.
In: *Contributions to Mathematical Psychology*, edited by N. Frederiksen and H. Gulliksen, New York:Holt, Rinehart & Winston, 1964, p. 110-182.

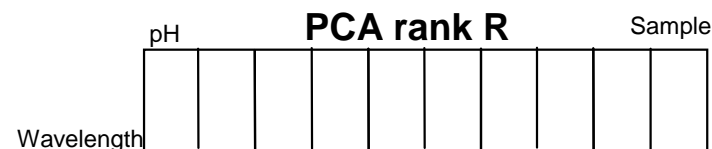
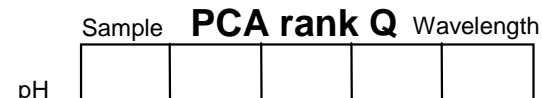
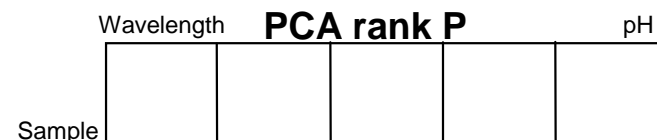
•L. R. Tucker. Some mathematical notes on three-mode factor analysis.
Psychometrika 31:279-311, 1966

■ Important theorem

- *If* the matricized ranks in the three modes are found to be P , Q , and R respectively
- *then* a (P, Q, R) Tucker3 model fit the data perfect

■ Practical value

- If the pseudo-ranks are definitely found to be P , Q , and R respectively
- *then* a (P, Q, R) Tucker3 model fit the data appropriately



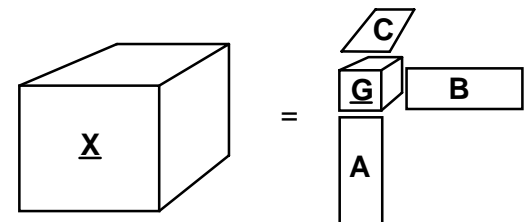
Unfold: Wold, Geladi, Esbensen, Öhman. *Principal component- and PLS-analyses generalized to multi-way (multi-order) data arrays. Copenhagen Symposium on Applied Statistics:249-277, 1986.*

Tucker3 vs PARAFAC



■ Differences from PARAFAC/PCA:

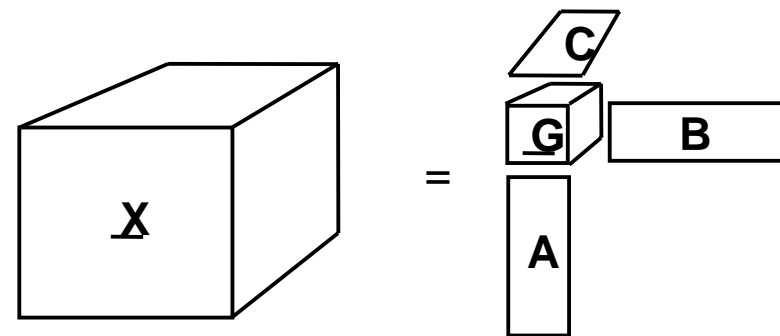
- The number of components can vary in **A**, **B**, and **C**!
- **G** is not superdiagonal
- Tucker loadings not unique (only subspace) = rotational freedom
- Tucker loadings orthogonal => variance-partitioning
- PARAFAC best rank F model, but does not describe the part of data within that subspace!
- Tucker best subspace F model but not rank F !!



Tucker3 - algorithm



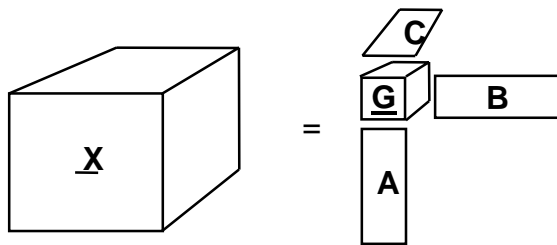
1. Initialize **B** and **C** (e.g. from SVD)
2. **A** equals first P left singular vectors of $\mathbf{X}^{(I \times JK)}(\mathbf{C} \otimes \mathbf{B})$
3. **B** equals first Q left singular vectors of $\mathbf{X}^{(J \times IK)}(\mathbf{C} \otimes \mathbf{A})$
4. **C** equals first R left singular vectors of $\mathbf{X}^{(K \times IJ)}(\mathbf{B} \otimes \mathbf{A})$
5. Go to step 2 until relative changes are small
6. $\mathbf{G} = \mathbf{A}'\mathbf{X}(\mathbf{C} \otimes \mathbf{B})$



Other Tucker models

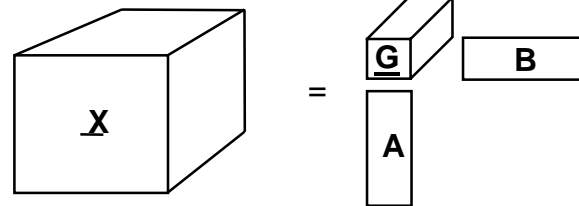


- Tucker3 has the number 3 because three modes are 'reduced'.
- Tucker2 and Tucker1 reduces two and one modes respectively



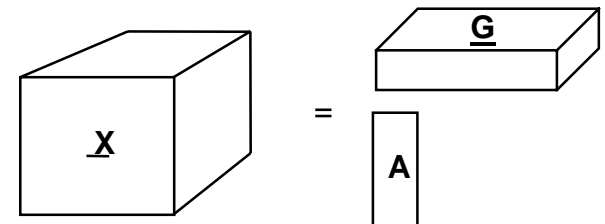
Tucker3

Tucker2 core often called Extended Core Array



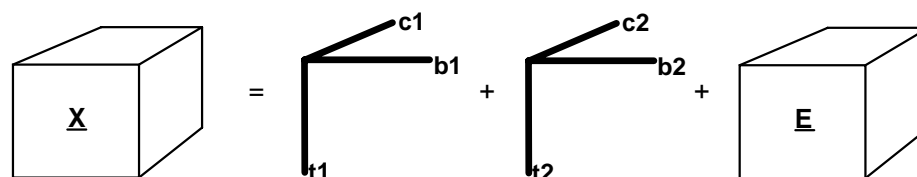
Tucker2

Tucker1 is identical to PCA on matrices X

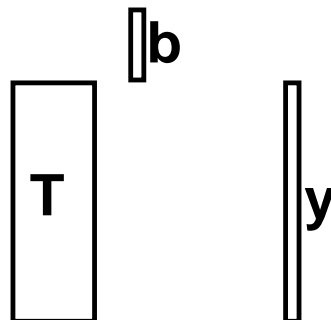


Tucker1

■ N-PLS



- Use a trilinear (PARAFAC-like) model of \underline{X} but such that the scores are predictive of \underline{y} .



Bro. Multiway calibration. Multi-linear PLS. *Journal of Chemometrics* 10:47-61, 1996.

■ Two-way PLS principle

- Find \mathbf{w} ($\|\mathbf{w}\|=1$) and \mathbf{q} ($\|\mathbf{q}\|=1$) such that the one-component models of $\mathbf{X} = \mathbf{t}\mathbf{w}' + \mathbf{E}_X$ and $\mathbf{Y} = \mathbf{u}\mathbf{q}' + \mathbf{E}_Y$ have maximal covariance ($\mathbf{t}'\mathbf{u}$)
- Make regression model to predict \mathbf{u} from \mathbf{t}
- Subtract the model from \mathbf{X} and predictions from \mathbf{Y}
- Proceed with the next component from residuals

■ Two-way PLS: how to do it - one y.

$$\begin{aligned} \max_{\mathbf{w}} \left[\text{cov}(\mathbf{t}, \mathbf{y}) \mid \mathbf{t} = \mathbf{X}\mathbf{w} \right] &= \text{Max covariance} \\ \max_{\mathbf{w}} \left[\mathbf{t}'\mathbf{y} \mid \mathbf{t} = \mathbf{X}\mathbf{w} \right] &= \text{LS model} \\ \max_{\mathbf{w}} \left[\mathbf{y}'\mathbf{X}\mathbf{w} \right] &= \\ \max_{\mathbf{w}} \left[\mathbf{z}'\mathbf{w} \right] &\Rightarrow \\ \mathbf{w} &= \frac{\mathbf{X}'\mathbf{y}}{\|\mathbf{X}'\mathbf{y}\|} \end{aligned}$$

■ Three-way PLS: how to do it - one y.

$$\max_{\mathbf{w}^J, \mathbf{w}^K} \left[\text{cov}(\mathbf{t}, \mathbf{y}) \mid t_i = \sum_{j=1}^J \sum_{k=1}^K x_{ijk} w_j^J w_k^K \right] =$$

$$\max_{\mathbf{w}^J, \mathbf{w}^K} \left[\sum_{i=1}^I t_i y_i \mid t_i = \sum_{j=1}^J \sum_{k=1}^K x_{ijk} w_j^J w_k^K \right] =$$

$$\max_{\mathbf{w}^J, \mathbf{w}^K} \left[\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K y_i x_{ijk} w_j^J w_k^K \right] =$$

$$\max_{\mathbf{w}^J, \mathbf{w}^K} \left[\sum_{j=1}^J \sum_{k=1}^K z_{jk} w_j^J w_k^K \right] =$$

$$\max_{\mathbf{w}^J, \mathbf{w}^K} \left[(\mathbf{w}^J)' \mathbf{Z} \mathbf{w}^K \right] \Rightarrow$$

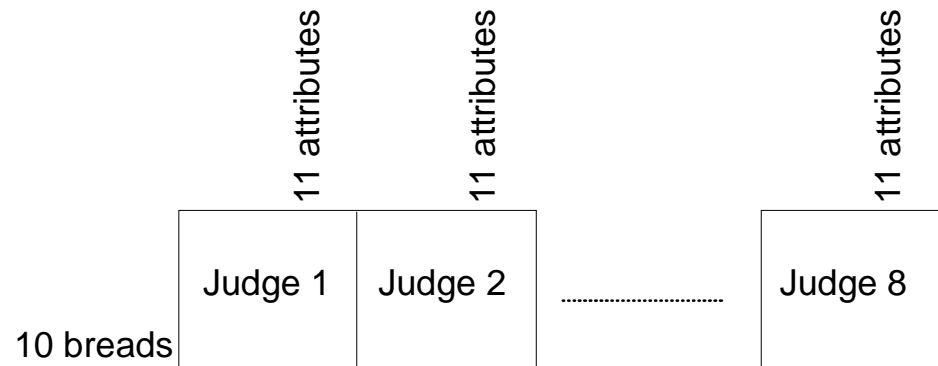
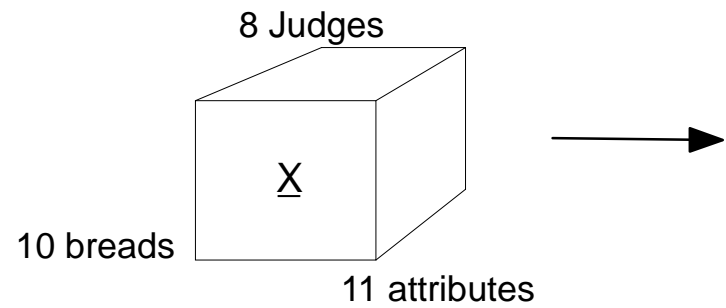
SVD on \mathbf{Z}

Example sensory data



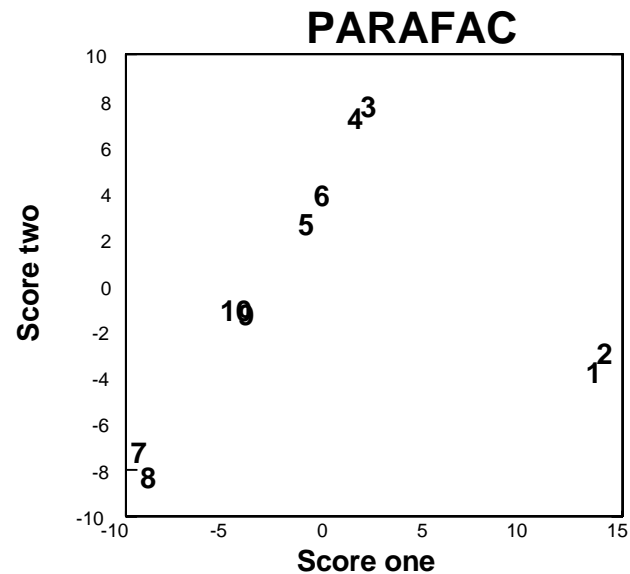
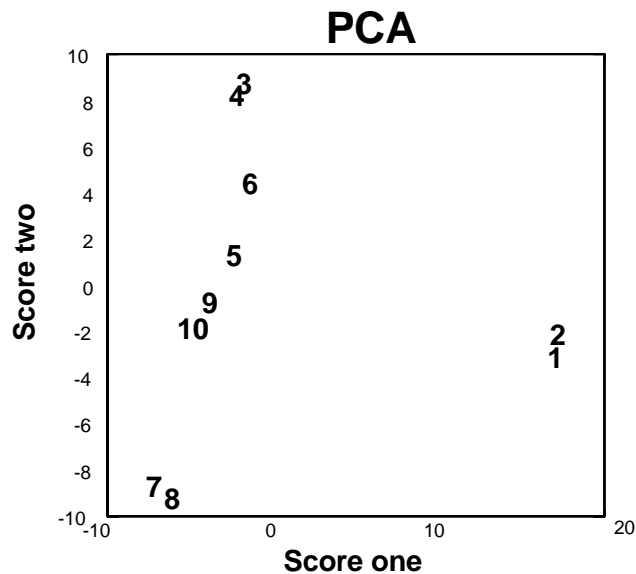
- Three-way, two-way:
- Does it make a difference?

- 5 breads (in replicates) \times 11 attributes \times 8 judges
- Data due to Magni Martens



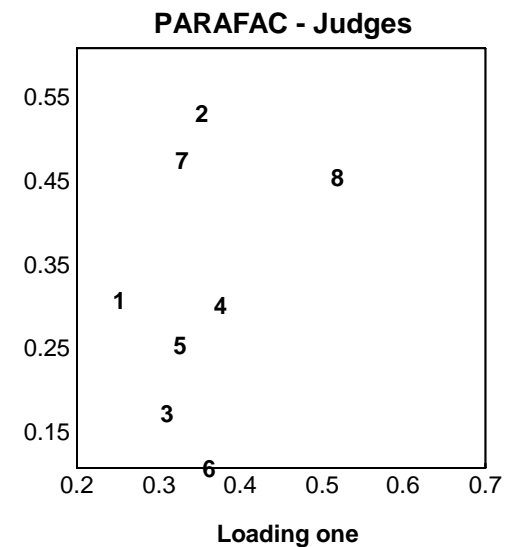
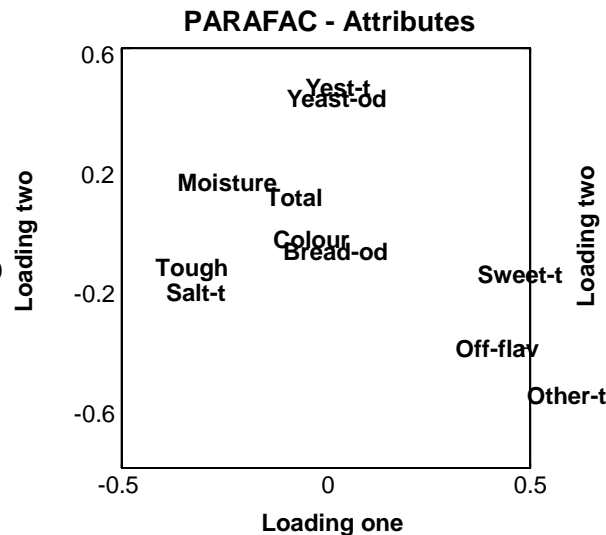
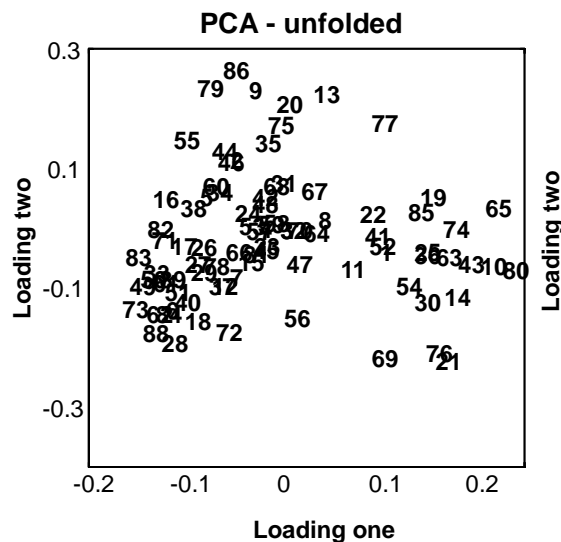
■ Scores from bilinear PCA and trilinear PARAFAC

Three-way more robust because of 'stronger' structural mode



■ Similar but note that replicates are closer for PARAFAC

Loadings from bilinear PCA and trilinear PARAFAC



PARAFAC 19 loading-elements per component
PCA 88 loading-elements per component!

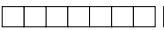

Sensory data



■ Calibration - predict salt content

■ 25% improvement!!

Three-way more predictive because less overfit

| | LV | Variation explained % | | | | RMSE | |
|--|----------|-----------------------|--------|--------|--------|--------|--------|
| | | X cal. | X val. | Y cal. | Y val. | Y cal. | Y val. |
|  unfold-PLS | 1 | 43 | 25 | 80 | 62 | 0.21 | 0.29 |
| | 2 | 61 | 38 | 95 | 76 | 0.10 | 0.23 |
| | 3 | 74 | 49 | 100 | 84 | 0.03 | 0.19 |
| | 4 | 78 | 49 | 100 | 84 | 0.01 | 0.19 |
| | 5 | 84 | 50 | 100 | 84 | 0.00 | 0.19 |
| | 6 | 87 | 52 | 100 | 84 | 0.00 | 0.19 |
|  Trilinear PLS | 1 | 31 | 22 | 75 | 60 | 0.23 | 0.30 |
| | 2 | 46 | 36 | 93 | 82 | 0.12 | 0.20 |
| | 3 | 54 | 44 | 98 | 91 | 0.07 | 0.15 |
| | 4 | 57 | 46 | 100 | 91 | 0.03 | 0.14 |
| | 5 | 60 | 47 | 100 | 90 | 0.02 | 0.15 |
| | 6 | 61 | 47 | 100 | 90 | 0.00 | 0.15 |

- Two-way classical least squares mixture model $\mathbf{X} = \mathbf{CS}'$,
 - E.g.
 - \mathbf{X} the measured spectra, \mathbf{C} the true concentrations, \mathbf{S} the pure spectra
- If pure spectra, \mathbf{S} , known, \mathbf{C} determined by a simple least squares regression step – $\mathbf{C} = \mathbf{X}(\mathbf{S}')^+$
- Thus, if all spectra known quantitation possible and no calibration samples are needed

- In three-way PARAFAC the model is $\mathbf{X}_k = \mathbf{A}\mathbf{D}_k\mathbf{B}'$,
 - \mathbf{X}_k the measured spectra at occasion k
 - \mathbf{A} the true concentrations
 - \mathbf{B} the pure spectra
 - \mathbf{C} (rows from \mathbf{D}_k) pure spectra/profiles etc.

- Pure spectra not necessary. They are found by decomposing with PARAFAC

- Thus, without prior knowledge quantitation possible and no calibration samples are needed except for fixing the scale

Fluorescence example



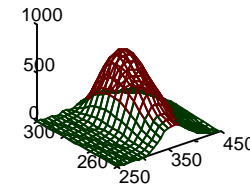
■ **Calibration set:** *One* sample with *one* analyte (2.67 μM Trp)

■ **Test set:** Two samples with three analytes each (Trp, Tyr, Phe)

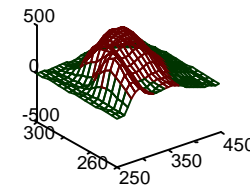
■ Three-component PARAFAC model fitted to the fluorescence data ($3 \times 201 \times 61$):

- **A** - estimated concentrations
- **B** - estimated emission spectra
- **C** - estimated excitation spectra

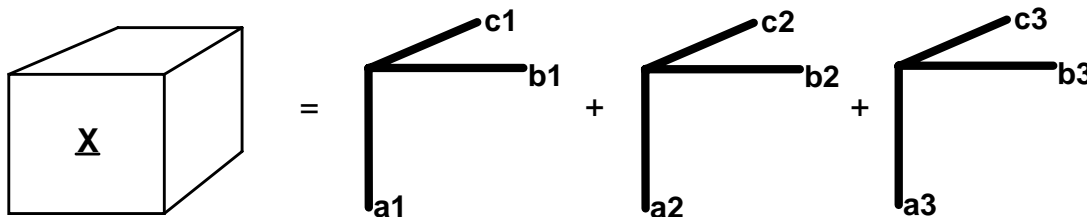
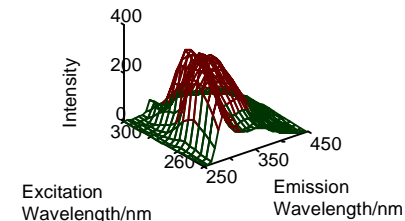
Standard (Trp)



Unknown 1 (Tyr, Trp, Phe)



Unknown 2 (Tyr, Trp, Phe)

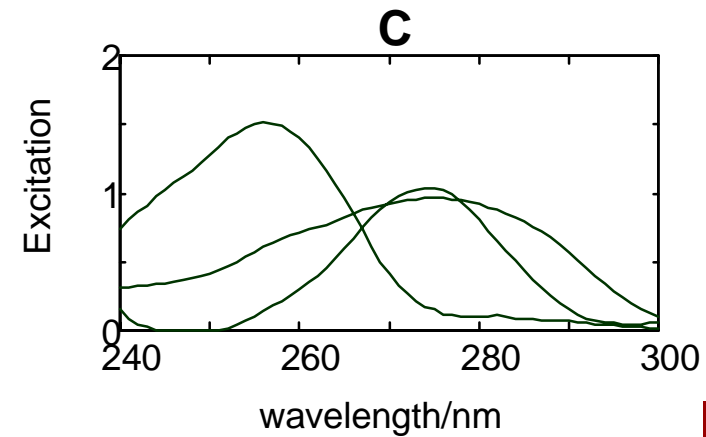
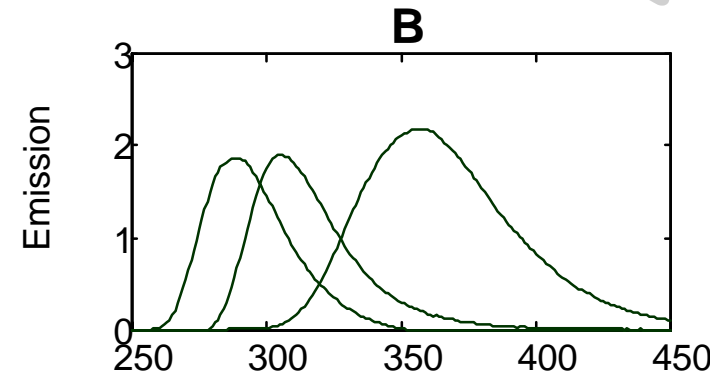


Fluorescence example



- Spectral loadings resemble pure spectra
- Sample scores - estimated concentration

$$\mathbf{A} = \begin{bmatrix} \text{Trp} & \text{Tyr} & \text{Phe} \\ 2.67 & 0 & 0 \\ 1.52 & 1.30 & 1.29 \\ .86 & 1.12 & 1.06 \end{bmatrix}, \quad \text{reference} = \begin{bmatrix} 2.67 \\ 1.58 \\ .88 \end{bmatrix}$$



■ How to make a prediction model

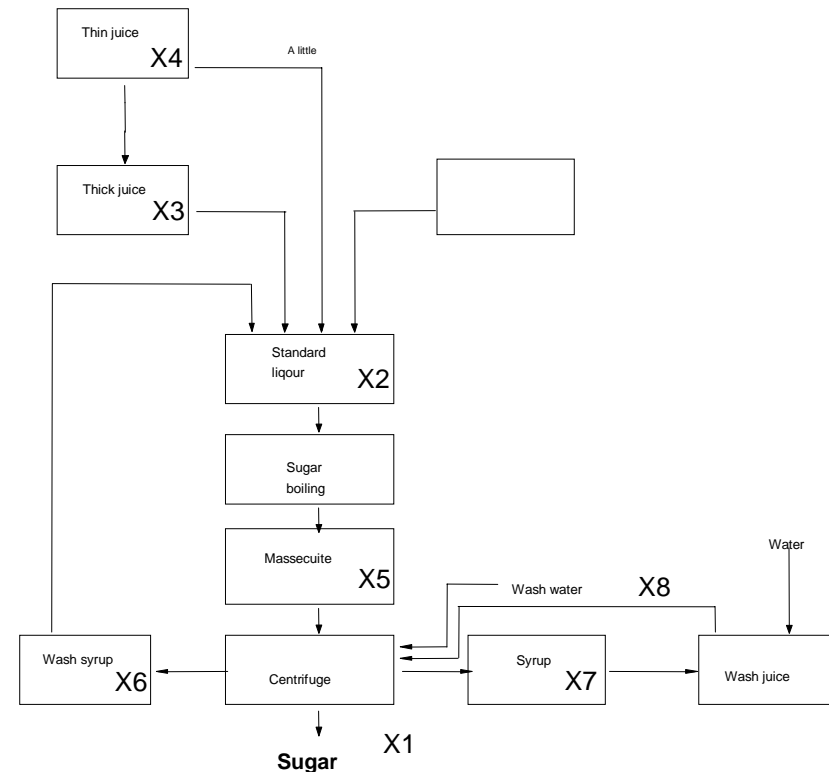
- Using sub-space regression, e.g. N-PLS
- Using 'second-order calibration' mostly with PARAFAC

■ Significant difference

- Predictions based on ordinary *regression* models only work if all interferences are varying independently in the calibration data ⇒
 - Many samples needed
 - Only similar samples can be predicted
- PARAFAC models work with only *one* calibration sample and with *unknown uncalibrated* interferences

■ Sugar made from beets

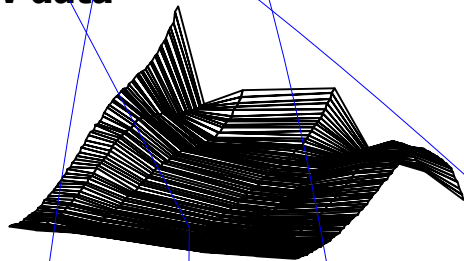
- End product sampled every 8'th hour during the three months of operation
- Measured spectrofluorometrically
- 260 samples



Sugar data model

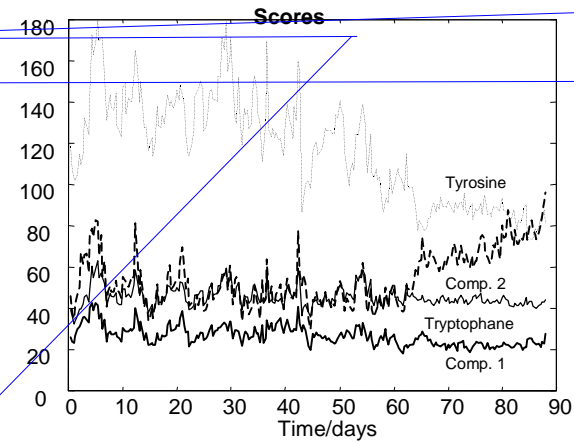
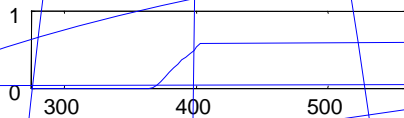


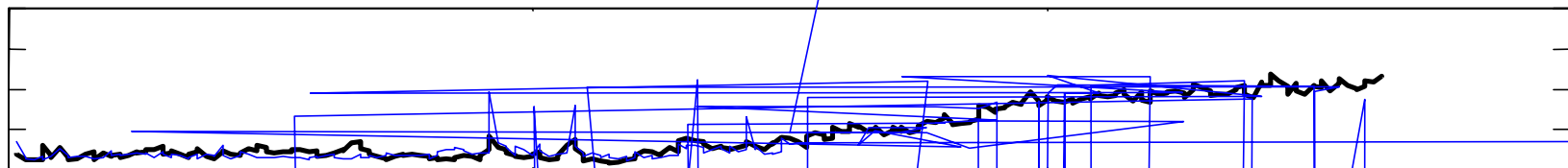
Raw data



Raw data: 268 samples from a sugar campaign - fluorescence landscapes

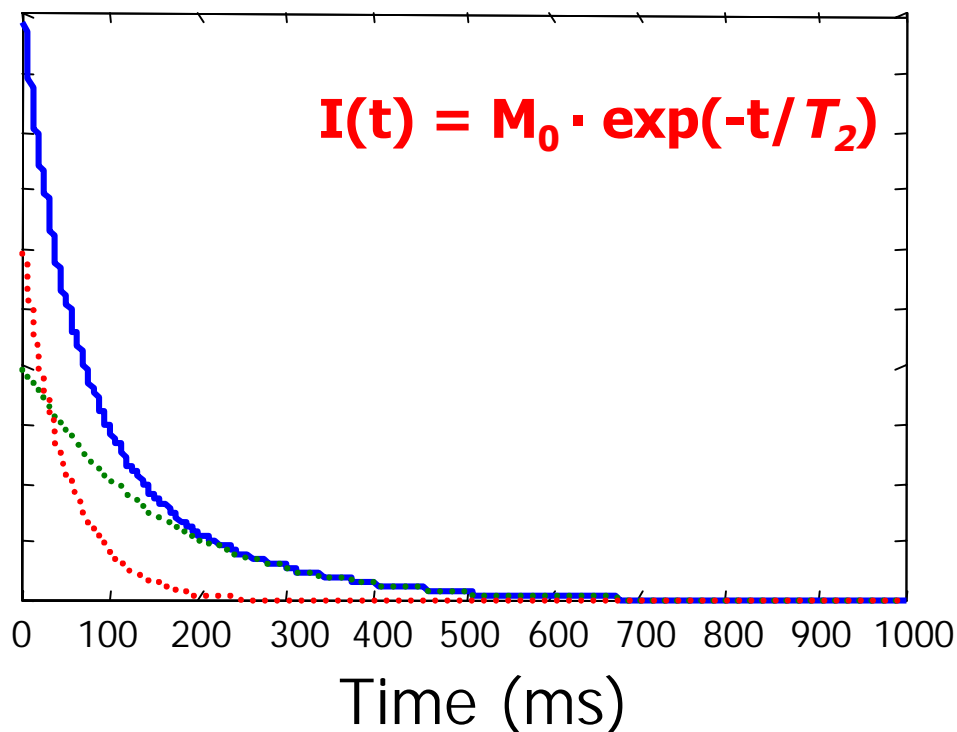
Concentrations





Two-way data but with very special structure in loadings

$$\mathbf{X} = \mathbf{AB}' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 27 & 9 & 3 & 1 \\ 16 & 8 & 4 & 2 \end{bmatrix}$$



| Sample 1 | Sample 2 |
|-----------|-----------|
| 27 | 16 |
| 9 | 8 |
| 3 | 4 |
| 1 | 2 |

$$\mathbf{X}_1 = \mathbf{AD}_1\mathbf{B}' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 9 & 3 & 1 \\ 8 & 4 & 2 \end{bmatrix}$$

SLAB 1 (\mathbf{X}_1)

| Sample 1 | Sample 2 |
|----------|----------|
| 27 | 16 |
| 9 | 8 |
| 3 | 4 |

$$\mathbf{X}_2 = \mathbf{AD}_2\mathbf{B}' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 9 & 3 & 1 \\ 8 & 4 & 2 \end{bmatrix}$$

SLAB 2 (\mathbf{X}_2)

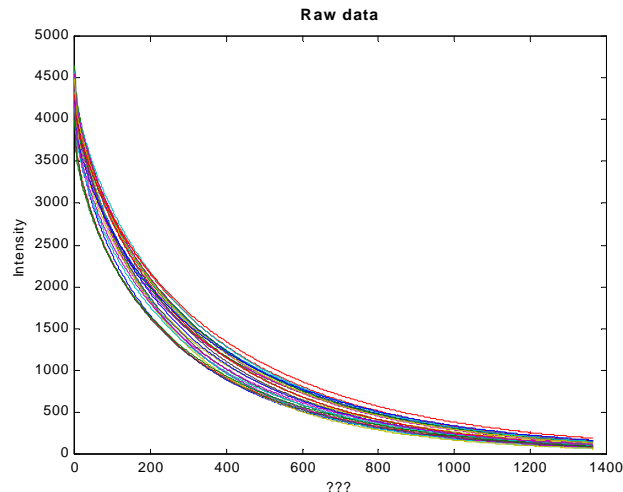
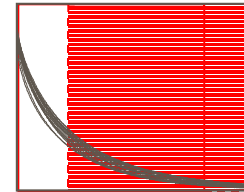
| Sample 1 | Sample 2 |
|----------|----------|
| 9 | 8 |
| 3 | 4 |
| 1 | 2 |

Example: Sensory quality



- **Predicting cooked potato quality from NMR on raw data**

- Potatoes cooked and served hot. Ten assessors evaluate texture profile, e.g. mealiness
- *Raw (!)* potato measured by NMR (CPMG pulse sequence)



A. K. Thybo et. al., *Food Science and Technology*, 33 (2):103-111, 2000.

Example: Sensory quality

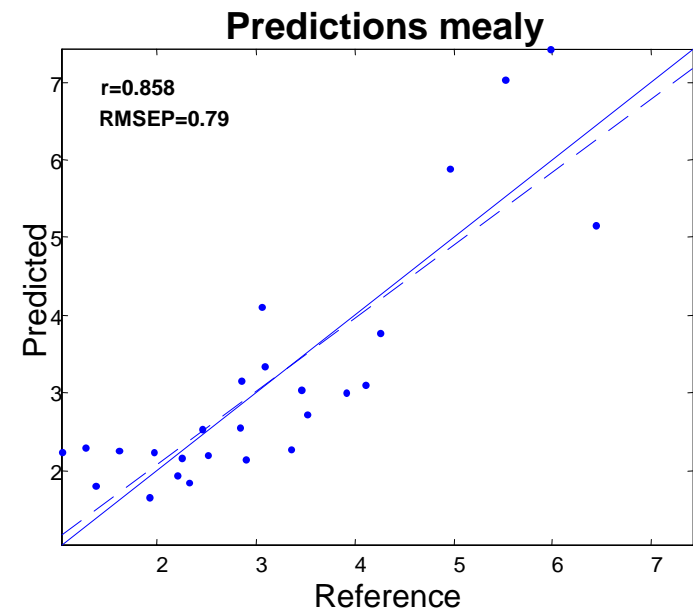
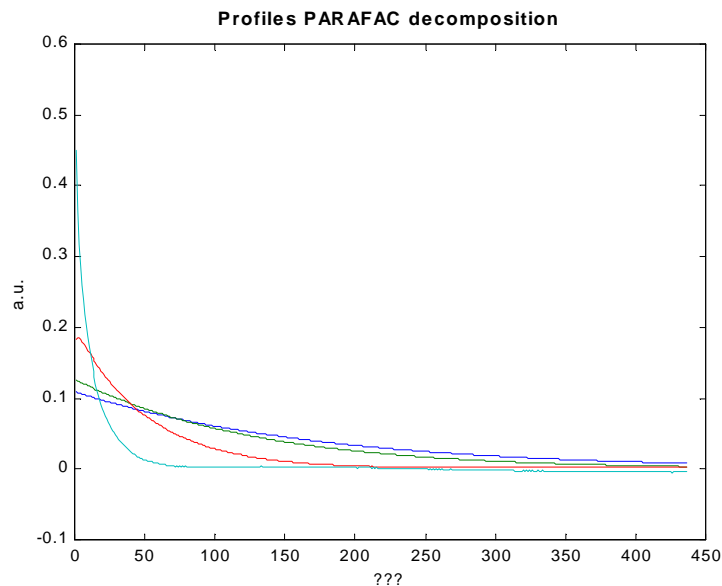


■ Decomposing NMR into meaningful latent variables

- PARAFAC on lagged data enables meaningful decomposition
- Four components adequate for describing NMR data

■ Predictions of mealiness from NMR

- Predicting sensory quality of *cooked* potatoes from amount of latent variables (i.e. directly from NMR of *raw* potatoes)



■ Using constraints in multi-way modeling

■ Example

■ Instead of 'PCA': $|| \mathbf{X} - \mathbf{AB}' ||$

■ fit the model: $|| \mathbf{X} - \mathbf{AB}' ||,$

subject to **A** and **B** are nonnegative

■ Constraints are essential in two-way curve resolution because the model is unidentified

■ In three-way curve resolution the model is often unique but constraints are still useful

Why constraints?



- **Obtain sensible parameters**

- Ex.: Require chromatographic profiles to have but one peak

- **Obtain unique solution**

- Ex.: Use selective channels in data to obtain uniqueness

- **Test hypotheses**

- Ex.: Investigate if tryptophane is present in sample

- **Avoiding degeneracy and numerical problems**

- Ex.: Enabling a PARAFAC model of data otherwise inappropriate for the mode

- **Speed up algorithms**

- Ex.: Use truncated bases to reexpress problem by a smaller problem

- **Enable quantitative analysis of qualitative data**

- Ex.: Incorporate sex and job type in a model for predicting income

Typical constraints



| | Spectroscopy | Chromatography | FIA | Kinetics |
|--------------------------------------|---------------------|-----------------------|------------|-----------------|
| Auto- & cross correlation | ✓ | ✓ | ✓ | ✓ |
| Uncertainty | ✓ | ✓ | ✓ | ✓ |
| Nonnegativity | ✓ | ✓ | ✓ | ✓ |
| Unimodality | (✓) | ✓ | ✓ | |
| Selectivity | | ✓ | ✓ | |
| Smoothness | ✓ | ✓ | ✓ | ✓ |
| Known spectra | ✓ | | | |
| ... | | | | |



■ **Model of a single samp**

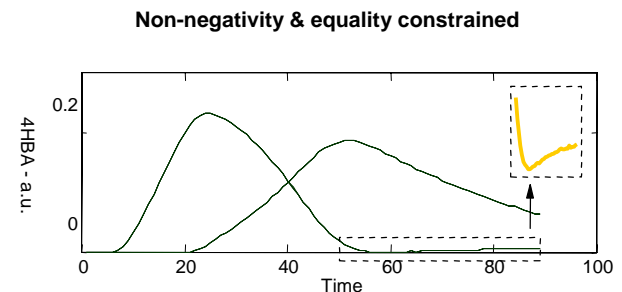
Effect of constraints



- Eq : Equality of summed profiles
- NNLS : Non-negativity of all parameters
- ULSR : Unimodality of FI Agrams/time profiles
- Fix : Fixing purely acidic/basic times to only reflect acidic/basic analytes

| SPECTRA | 2HBA acidic | 2HBA basic | 3HBA acidic | 3HBA basic | 4HBA acidic | 4HBA basic |
|------------------|----------------|---------------|----------------|---------------|----------------|---------------|
| Eq | 0.9893* | 0.9871* | 0.9689* | 0.7647* | 0.9106* | 0.9211* |
| NNLS | 0.9944 | 0.9117* | 0.9952 | 0.9241 | 0.9974 | 0.9977 |
| NNLS/Eq | 0.9946 | 0.9312* | 0.9953 | 0.9988 | 0.9965 | 0.9971 |
| NNLS/ULSR/Eq | 0.9946 | 0.9590* | 0.9953 | 0.9989 | 0.9966 | 0.9943 |
| NNLS/ULSR/Fix/Eq | 0.9946 | 0.9989 | 0.9954 | 0.9986 | 0.9961 | 0.9977 |

Correlations



■ Algorithms

- Tucker, N-PLS are unproblematic, even for large data sets
- PARAFAC problematic
 - Multiple local minima
 - 2-factor degeneracy (model-problem): no solution?
 - Solution sensitive to correct dimensionality (not sequential)
 - Slow convergence

■ Other problems currently considered

- Statistical measures (rank/DoF problems)
- Diagnostics for choosing the number of components
- Handling highly structured error covariances
- Handling huge amounts of missing data

■ **Uniqueness** (mainly PARAFAC)

- Pure spectra
- Second order advantage
 - No interferences, only one standard

■ **Better structural model**

- Robustness/noise reduction
- Simpler model
- Interpretation
- Better predictions

■ References

- Harshman, Lundy. *Comp. Stat. Data Anal.*, 1994, **18**, 39
- Leurgans, Ross. *Statist. Sci.*, 1992, **7**, 289
- Smilde. *Chemom. Intell. Lab. Syst.*, 1992, **5**, 143
- Bro. *J. Chemom.* 1996, **10**, 47
- Bro, *Chemom. Intell. Lab. Syst.*, 1997, **38**, 149
- Sidiropoulos & Bro. *PARAFAC Techniques for Signal separation*. In: *Signal Processing Advances in Communications*, edited by P. Stoica, G. B. Giannakis, Y. Hua, and L. Tong, Prentice-Hall, 2000.

■ Software & info

- <http://www.models.kvl.dk> (Free matlab code'n'course, database of papers)
- http://www.ece.umn.edu/users/nikos/public_html/3SPICE/3SPICEmain.html
TRIPLE SPICE – Multi-way analysis in signal processing – matlab, presentations etc.
- <http://www.eigenvector.com> (Matlab code)
- <http://www.fsw.leidenuniv.nl/~kroonenb> (Stand alone)