Tensor decompositions in Engineering

Pierre Comon

July 21, 2008

Canonical decomposition

Goal: Given a tensor $\mathbf{T}$ of order d, defined in space $V_1 \otimes \ldots \otimes V_d$ on $\mathbb{R}$ or $\mathbb{C}$, with fixed bases

Canonical decomposition

Goal: Given a tensor $\mathbf{T}$ of order d, defined in space $V_1 \otimes \ldots \otimes V_d$ on $\mathbb{R}$ or $\mathbb{C}$, with fixed bases

- Determine the largest P for which the decomposition

$$\mathbf{T} = \sum_{p=1}^{P} \mathbf{a}(p) \otimes \mathbf{b}(p) \otimes \mathbf{c}(p) \otimes \ldots \mathbf{d}(p)$$

is *unique* (depending on P, d, and dimensions)

Canonical decomposition

Goal: Given a tensor $\mathbf{T}$ of order d, defined in space $V_1 \otimes \ldots \otimes V_d$ on $\mathbb{R}$ or $\mathbb{C}$, with fixed bases

- Determine the largest P for which the decomposition

$$\mathbf{T} = \sum_{p=1}^{P} \mathbf{a}(p) \otimes \mathbf{b}(p) \otimes \mathbf{c}(p) \otimes \ldots \mathbf{d}(p)$$

  is *unique* (depending on P, d, and dimensions)
- Or at least when there are *finitely* many

Canonical decomposition

Goal: Given a tensor $\mathbf{T}$ of order d, defined in space $V_1 \otimes \ldots \otimes V_d$ on $\mathbb{R}$ or $\mathbb{C}$, with fixed bases

- Determine the largest P for which the decomposition

$$\mathbf{T} = \sum_{p=1}^{P} \mathbf{a}(p) \otimes \mathbf{b}(p) \otimes \mathbf{c}(p) \otimes \ldots \mathbf{d}(p)$$

  is *unique* (depending on P, d, and dimensions)
- Or at least when there are *finitely* many
- Provide an algorithm to *compute* the collections of vectors $\mathbf{a}(p)$, $\mathbf{b}(p)$, $\mathbf{c}(p)$, $\ldots \mathbf{d}(p)$

Application areas

1. Antenna Array Processing: Telecommunications (Cellular, Satellite, Military), Radar, Sonar, Biomedical (EchoGraphy, ElectroEncephaloGraphy, ElectroCardioGraphy)...

Application areas

1. Antenna Array Processing: Telecommunications (Cellular, Satellite, Military), Radar, Sonar, Biomedical (EchoGraphy, ElectroEncephaloGraphy, ElectroCardioGraphy)...
   - Techniques using High Order Statistics (e.g. *cumulants*)

Application areas

1. Antenna Array Processing: Telecommunications (Cellular, Satellite, Military), Radar, Sonar, Biomedical (EchoGraphy, ElectroEncephaloGraphy, ElectroCardioGraphy)...
   - Techniques using High Order Statistics (e.g. *cumulants*)
   - Deterministic Techniques exploiting receiver *diversities*

Application areas

1. Antenna Array Processing: Telecommunications (Cellular, Satellite, Military), Radar, Sonar, Biomedical (EchoGraphy, ElectroEncephaloGraphy, ElectroCardioGraphy)...
    - Techniques using High Order Statistics (e.g. *cumulants*)
    - Deterministic Techniques exploiting receiver *diversities*
2. Data & Factor Analysis (e.g. Psychometrics, Chemometrics, Food Sciences, Environment...)

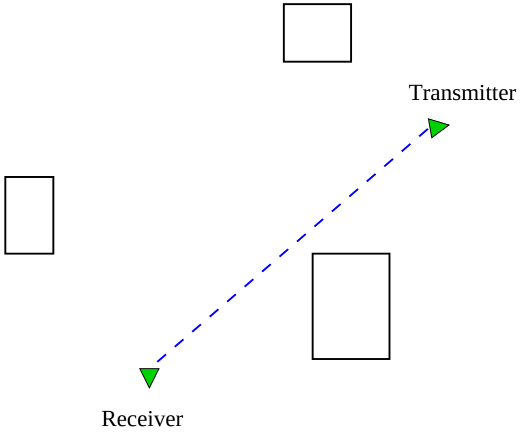
Application areas

1. Antenna Array Processing: Telecommunications (Cellular, Satellite, Military), Radar, Sonar, Biomedical (EchoGraphy, ElectroEncephaloGraphy, ElectroCardioGraphy)...
   - Techniques using High Order Statistics (e.g. *cumulants*)
   - Deterministic Techniques exploiting receiver *diversities*
2. Data & Factor Analysis (e.g. Psychometrics, Chemometrics, Food Sciences, Environment...)
3. Arithmetic Complexity Theory

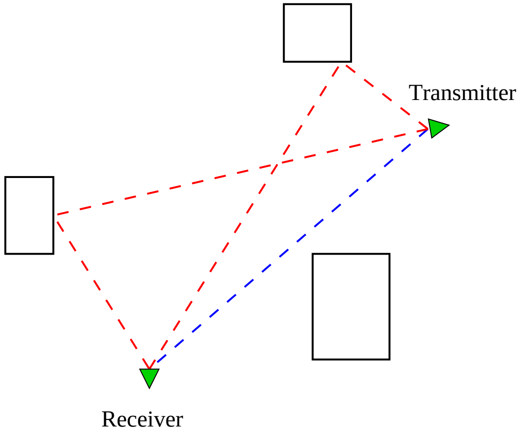
Application areas

1. Antenna Array Processing: Telecommunications (Cellular, Satellite, Military), Radar, Sonar, Biomedical (EchoGraphy, ElectroEncephaloGraphy, ElectroCardioGraphy)...
   - Techniques using High Order Statistics (e.g. *cumulants*)
   - Deterministic Techniques exploiting receiver *diversities*
2. Data & Factor Analysis (e.g. Psychometrics, Chemometrics, Food Sciences, Environment...)
3. Arithmetic Complexity Theory
4. Many other fields
   e.g. Medical imaging, Differential Geometry, Speech, Machine Learning, Control...

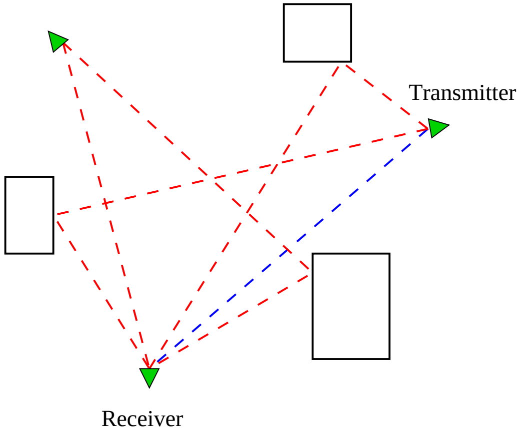
Example: Antenna Array Processing (1)



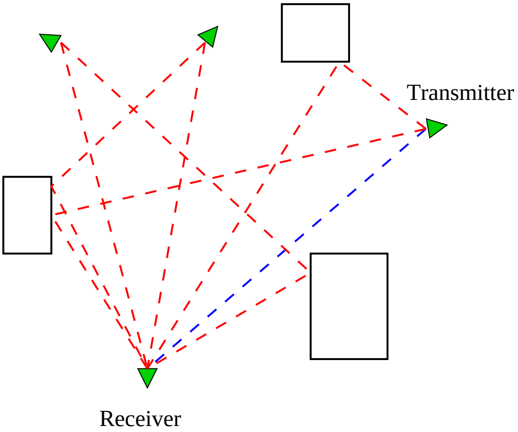
Example: Antenna Array Processing (1)



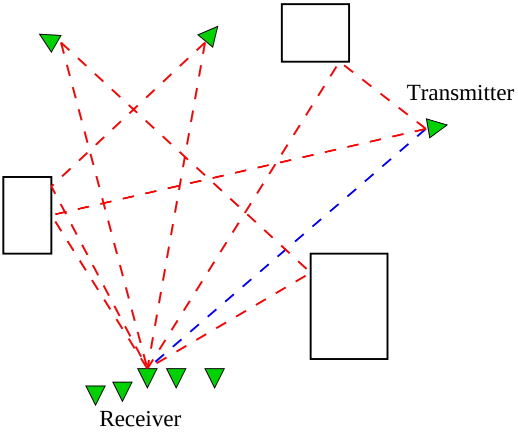
Example: Antenna Array Processing (1)



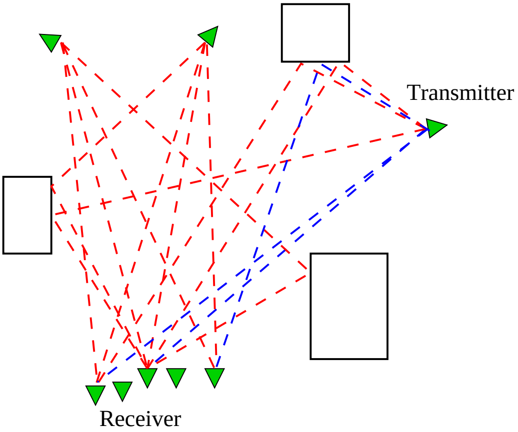
Example: Antenna Array Processing (1)



Example: Antenna Array Processing (1)



Example: Antenna Array Processing (1)



Example: Antenna Array Processing (2)

Modeling the signals received on an array of antennas generally leads to a *matrix decomposition*:

$$T_{ij} = \sum_q \sum_\ell a_{iq\ell} \sum_k h_{q\ell k} \; s_{kqj}$$

i: space	k: symbol time	**a**: receiver geometry
j: time	q: transmitter	**h**: global channel impulse response
	ℓ: path	**s**: transmitted signal

Example: Antenna Array Processing (2)

Modeling the signals received on an array of antennas generally leads to a *matrix decomposition*:

$$T_{ijp} = \sum_q \sum_\ell a_{iq\ell} \sum_k h_{q\ell kp} s_{kqj}$$

i: space	k: symbol time	$\mathbf{a}$: receiver geometry
j: time	q: transmitter	$\mathbf{h}$: global channel impulse response
	ℓ: path	$\mathbf{s}$: transmitted signal

But in the presence of additional *diversity*, a tensor can be constructed, thanks to new variable p

Example: Antenna Array Processing (3)

New variable p can represent:

- Oversampling (sample index),
- Spreading code (chip index),
- Frequency (multicarrier),
- Geometrical invariance (subarray index),
- Polarization
- Nonstationarity...

Example: Antenna Array Processing (3)

New variable p can represent:

- Oversampling (sample index),
- Spreading code (chip index),
- Frequency (multicarrier),
- Geometrical invariance (subarray index),
- Polarization
- Nonstationarity...

Warning: tensor should not have proportional matrix slices (degeneration)

Data not in tensor format

If no additional diversity available, problem cannot be solved, unless other properties on sources exist:

- Statistical independence
- Sparsity
- Discrete...

Data not in tensor format

If no additional diversity available, problem cannot be solved, unless other properties on sources exist:

- Statistical independence ⇒ *symmetric tensors*
- Sparsity
- Discrete...

Example: Speech

The Coktail Party problem

$$\text{In free space: } T_{ij} = \sum_q a_{iq} \sum_k h_q(k)\, s_q(j-k)$$

Remark: Data not in tensor format $\Rightarrow$ statistical tools

Example: Fluorescence Spectroscopy

An optical excitation produces several effects

Example: Fluorescence Spectroscopy

An optical excitation produces several effects

- Rayleigh diffusion
- Raman diffusion
- Fluorescence

Example: Fluorescence Spectroscopy

An optical excitation produces several effects

- Rayleigh diffusion
- Raman diffusion
- Fluorescence
  At low concentrations, Beer-Lambert law can be linearized

$$I(\lambda_e, \lambda_f, k) = I_o \sum_n \gamma_n(\lambda_f)\, \epsilon_n(\lambda_e)\, c_{k,n}$$

Example: Fluorescence Spectroscopy

An optical excitation produces several effects

- Rayleigh diffusion
- Raman diffusion
- Fluorescence
  At low concentrations, Beer-Lambert law can be linearized

$$I(\lambda_e, \lambda_f, k) = I_o \sum_n \gamma_n(\lambda_f)\, \epsilon_n(\lambda_e)\, c_{k,n}$$

- But there are also non linear effects:
  - Screen effect (if some components are too concentrated)
  - Reabsorption of fluorescent emissions by the medium

Example: Factor Analysis

Food Sciences:
one of the numerous application areas

judges $\times$ *products* $\times$ *sensory properties*

$$T_{ijk} = \sum_p A_{ip}\, B_{jp}\, C_{kp}$$

Particular constraints

- Toeplitz matrix (convolution)
- Symmetries (partial or total, real or Hermitean)
- Positivity
- Blocks $\rightarrow$ not $r-$secant of Segre variety...

The practical problem

- In practice, the tensor we are given is subject to errors (model, calculation) or noise, and is hence *generic*.

The practical problem

- In practice, the tensor we are given is subject to errors (model, calculation) or noise, and is hence *generic*.
- If we expect a rank different from generic, we have an *approximation problem* $\Rightarrow$ optimization

Direct vs Inverse

Two formulations possible if rank bounded by dimension:

1. *Direct:* look for **A**, **B**, … **D**:

$$\min_{\mathbf{A},\mathbf{B},..\mathbf{D}} ||\, T_{ijk..\ell} - \sum_{p=1}^{P} A_{ip}\, B_{jp}\, \dots\, D_{\ell p}||^2$$

i.e. decompose **T** into a sum of P rank-one terms

Direct vs Inverse

Two formulations possible if rank bounded by dimension:

1. *Direct:* look for $\mathbf{A}$, $\mathbf{B}$, $\dots \mathbf{D}$:

$$\min_{\mathbf{A},\mathbf{B},..\mathbf{D}} || T_{ijk..\ell} - \sum_{p=1}^{P} A_{ip}\, B_{jp}\, \dots\, D_{\ell p}||^2$$

   i.e. decompose $\mathbf{T}$ into a sum of P rank-one terms

2. *Inverse:* look for $\mathbf{A}'$, $\mathbf{B}'$, $\dots \mathbf{D}'$:

$$\min_{\mathbf{A}',\mathbf{B}',..\mathbf{D}'} \sum_{mnp..q \neq ppp..p} |\sum_{ijk..\ell} T_{ijk..\ell}\, A'_{im}\, B'_{jn}\, \dots\, D'_{\ell q}|^2$$

   i.e. try to “diagonalize” $\mathbf{T}$ by linear invertible change of coordinates in each space

Orthogonal decomposition

If **A**, **B**, **C**, **D** orthogonal, the two formulations are equivalent:

Orthogonal decomposition

If **A**, **B**, **C**, **D** orthogonal, the two formulations are equivalent:

1. *Direct:*

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\boldsymbol{\Delta}} ||T_{ijk\ell} - \sum_{p=1}^{P} A_{ip}\, B_{jp}\, C_{kp}\, D_{\ell p}\, \Delta_{pppp}||^2$$

Orthogonal decomposition

If **A**, **B**, **C**, **D** orthogonal, the two formulations are equivalent:

1. *Direct:*

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\mathbf{\Delta}} ||T_{ijk\ell} - \sum_{p=1}^{P} A_{ip}\, B_{jp}\, C_{kp}\, D_{\ell p}\, \Delta_{pppp}||^2$$

2. *Inverse:* $\min ||\sum_{ijk\ell} A_{pi}\, B_{qj}\, C_{rk}\, D_{s\ell}\; T_{ijk\ell} - \Delta_{pppp}\, \delta_{pqrs}||^2$ or

$$\max_{\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}} \sum_{p} |\sum_{ijk\ell} A_{pi}\, B_{pj}\, C_{pk}\, D_{p\ell}\; T_{ijk\ell}|^2$$

Orthogonal decomposition

If **A**, **B**, **C**, **D** orthogonal, the two formulations are equivalent:

1. *Direct:*

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\boldsymbol{\Delta}} ||T_{ijk\ell} - \sum_{p=1}^{P} A_{ip}\, B_{jp}\, C_{kp}\, D_{\ell p}\, \Delta_{pppp}||^2$$

2. *Inverse:* $\min ||\sum_{ijk\ell} A_{pi}\, B_{qj}\, C_{rk}\, D_{s\ell}\, T_{ijk\ell} - \Delta_{pppp}\, \delta_{pqrs}||^2$ or

$$\max_{\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}} \sum_{p} |\sum_{ijk\ell} A_{pi}\, B_{pj}\, C_{pk}\, D_{p\ell}\, T_{ijk\ell}|^2$$

Proof. The Frobenius norm is invariant under orthogonal change of coordinates. $\square$

General decomposition

- If rank exceeds dimensions (occurs generically), only the direct formulation is possible:

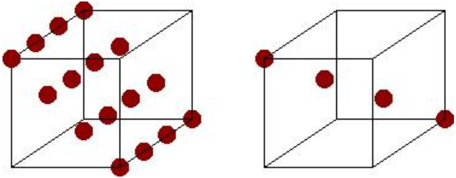
$$\min_{\mathbf{A},\mathbf{B},..\mathbf{D}} ||\, T_{ijk..\ell} - \sum_{p=1}^{P} A_{ip}\, B_{jp}\, \dots\, D_{\ell p}||^2$$

- Rank reduction is often necessary to restore uniqueness

Examples of objectives for orthogonal decomposition

$$C_{ijk\ell} \stackrel{\text{def}}{=} \sum_p Q_{ip}\, Q_{jq}\, Q_{kr}\, Q_{\ell s}\, T_{pqrs}$$

- $\Upsilon_{CoM}(\mathbf{Q}) \stackrel{\text{def}}{=} \sum_i C_{iiii}^2$ (maximize diagonal entries)
- $\Upsilon_{STO}(\mathbf{Q}) \stackrel{\text{def}}{=} \sum_{ij} C_{iiij}^2$ (jointly diagonalize 3rd order slices)
- $\Upsilon_{JAD}(\mathbf{Q}) \stackrel{\text{def}}{=} \sum_{ijk} C_{iijk}^2$ (jointly diagonalize matrix slices)



Matrix slices diagonalization $\neq$ Tensor diagonalization

Problem P2

1. At each step, a plane rotation is computed and yields the *global maximum* of the objective Υ restricted to one variable

$$\max_{\mathbf{Q}} \sum_{p=1}^{n} \Big| \sum_{ijk\ell} Q_{ip}\, Q_{jp}\, Q_{kp}\, Q_{\ell p}\, \mathcal{C}_{ijk\ell} \Big|^2$$

Problem P2

1. At each step, a plane rotation is computed and yields the *global maximum* of the objective Υ restricted to one variable

$$\max_{\mathbf{Q}} \sum_{p=1}^{n} \Big| \sum_{ijk\ell} Q_{ip}\, Q_{jp}\, Q_{kp}\, Q_{\ell p}\, \mathcal{C}_{ijk\ell} \Big|^2$$

2. There is no proof that the sequence of successive plane rotations yields the global maximum, in the general case (i.e. for symmetric tensors of dimension n and general form)

Problem P2

1. At each step, a plane rotation is computed and yields the *global maximum* of the objective Υ restricted to one variable

$$\max_{\mathbf{Q}} \sum_{p=1}^{n} \Big| \sum_{ijk\ell} Q_{ip}\, Q_{jp}\, Q_{kp}\, Q_{\ell p}\, \mathcal{C}_{ijk\ell} \Big|^2$$

2. There is no proof that the sequence of successive plane rotations yields the global maximum, in the general case (i.e. for symmetric tensors of dimension n and general form)
3. Yet, no counter-example has been found since 1991

Computation: bibliographical comments

- Orthogonal diagonalization of symmetric tensors:

Computation: bibliographical comments

- Orthogonal diagonalization of symmetric tensors:
  - JAD in 2 modes: [DeLe78] ($\mathbb{R}$), [CardS93] ($\mathbb{C}$), with matrix exponential [TanaF07]

Computation: bibliographical comments

- Orthogonal diagonalization of symmetric tensors:
  - JAD in 2 modes: [DeLe78] ($\mathbb{R}$), [CardS93] ($\mathbb{C}$), with matrix exponential [TanaF07]
  - JAD 2 modes with positive definite matrices: [Flur86]

Computation: bibliographical comments

- Orthogonal diagonalization of symmetric tensors:
  - JAD in 2 modes: [DeLe78] ($\mathbb{R}$), [CardS93] ($\mathbb{C}$), with matrix exponential [TanaF07]
  - JAD 2 modes with positive definite matrices: [Flur86]
  - JAD in 3 modes: [DelaDV01]

Computation: bibliographical comments

- Orthogonal diagonalization of symmetric tensors:
  - JAD in 2 modes: [DeLe78] ($\mathbb{R}$), [CardS93] ($\mathbb{C}$), with matrix exponential [TanaF07]
  - JAD 2 modes with positive definite matrices: [Flur86]
  - JAD in 3 modes: [DelaDV01]
  - direct diago without slicing (pairs): [Como92] [Como94]

Computation: bibliographical comments

- Orthogonal diagonalization of symmetric tensors:
  - JAD in 2 modes: [DeLe78] ($\mathbb{R}$), [CardS93] ($\mathbb{C}$), with matrix exponential [TanaF07]
  - JAD 2 modes with positive definite matrices: [Flur86]
  - JAD in 3 modes: [DelaDV01]
  - direct diago without slicing (pairs): [Como92] [Como94]
- Orthogonal diagonalization of non symmetric tensors:

Computation: bibliographical comments

- Orthogonal diagonalization of symmetric tensors:
  - JAD in 2 modes: [DeLe78] ($\mathbb{R}$), [CardS93] ($\mathbb{C}$), with matrix exponential [TanaF07]
  - JAD 2 modes with positive definite matrices: [Flur86]
  - JAD in 3 modes: [DelaDV01]
  - direct diago without slicing (pairs): [Como92] [Como94]
- Orthogonal diagonalization of non symmetric tensors:
  - ALS type [Kroo83] [Kier92]

Computation: bibliographical comments

- Orthogonal diagonalization of symmetric tensors:
  - JAD in 2 modes: [DeLe78] ($\mathbb{R}$), [CardS93] ($\mathbb{C}$), with matrix exponential [TanaF07]
  - JAD 2 modes with positive definite matrices: [Flur86]
  - JAD in 3 modes: [DelaDV01]
  - direct diago without slicing (pairs): [Como92] [Como94]
- Orthogonal diagonalization of non symmetric tensors:
  - ALS type [Kroo83] [Kier92]
  - ALS on pairs: [MartV08] [SoreC08]

Computation: bibliographical comments

- Orthogonal diagonalization of symmetric tensors:
  - JAD in 2 modes: [DeLe78] ($\mathbb{R}$), [CardS93] ($\mathbb{C}$), with matrix exponential [TanaF07]
  - JAD 2 modes with positive definite matrices: [Flur86]
  - JAD in 3 modes: [DelaDV01]
  - direct diago without slicing (pairs): [Como92] [Como94]
- Orthogonal diagonalization of non symmetric tensors:
  - ALS type [Kroo83] [Kier92]
  - ALS on pairs: [MartV08] [SoreC08]
  - JAD in 2 modes ($\mathbb{R}$): [Pesq01]

Computation: bibliographical comments

- Orthogonal diagonalization of symmetric tensors:
  - JAD in 2 modes: [DeLe78] ($\mathbb{R}$), [CardS93] ($\mathbb{C}$), with matrix exponential [TanaF07]
  - JAD 2 modes with positive definite matrices: [Flur86]
  - JAD in 3 modes: [DelaDV01]
  - direct diago without slicing (pairs): [Como92] [Como94]
- Orthogonal diagonalization of non symmetric tensors:
  - ALS type [Kroo83] [Kier92]
  - ALS on pairs: [MartV08] [SoreC08]
  - JAD in 2 modes ($\mathbb{R}$): [Pesq01]
  - Matrix exponential [SoreICD08]

Problem P3

Now allow general invertible transforms

- Ill-poseness of optimization over set of invertible matrices

Possible solutions:

Problem P3

Now allow general invertible transforms

- Ill-poseness of optimization over set of invertible matrices

Possible solutions:

- Impose a constraint like $\det \mathbf{A} \geq \eta > 0$?

Problem P3

Now allow general invertible transforms

- Ill-poseness of optimization over set of invertible matrices

Possible solutions:

- Impose a constraint like $\det \mathbf{A} \geq \eta > 0$?
- restrict $\mathbf{A}$ to diagonally dominant, i.e. $|A_{ii}| \geq \eta + \sum_{j \neq i} |A_{ij}|$, and parameterize it as a product of matrices of the form $(\mathbf{I} + \mathbf{W})$?

Problem P3

Now allow general invertible transforms

- Ill-poseness of optimization over set of invertible matrices

Possible solutions:

- Impose a constraint like $\det \mathbf{A} \geq \eta > 0$?
- restrict $\mathbf{A}$ to diagonally dominant, i.e. $|A_{ii}| \geq \eta + \sum_{j \neq i} |A_{ij}|$, and parameterize it as a product of matrices of the form $(\mathbf{I} + \mathbf{W})$?
- Decompose as $\mathbf{A} = \mathbf{Q}\,\mathbf{R}$ where $\mathbf{R}$ triangular with $R_{ii} \geq \eta > 0$?

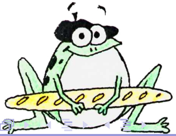
Problem P3

Now allow general invertible transforms

- Ill-poseness of optimization over set of invertible matrices

Possible solutions:

- Impose a constraint like $\det \mathbf{A} \geq \eta > 0$?
- restrict $\mathbf{A}$ to diagonally dominant, i.e. $|A_{ii}| \geq \eta + \sum_{j \neq i} |A_{ij}|$, and parameterize it as a product of matrices of the form $(\mathbf{I} + \mathbf{W})$?
- Decompose as $\mathbf{A} = \mathbf{Q}\,\mathbf{R}$ where $\mathbf{R}$ triangular with $R_{ii} \geq \eta > 0$?
- Just add a penalty term of the form $\log\det \mathbf{A}$ to the objective ?



Computation: bibliographical comments

- Invertible diagonalization of a collection of symmetric matrices:

Computation: bibliographical comments

- Invertible diagonalization of a collection of symmetric matrices:
  - maximizes iteratively a lower bound to the decrease on a probabilistic objective [Pham01]

Computation: bibliographical comments

- Invertible diagonalization of a collection of symmetric matrices:
    - maximizes iteratively a lower bound to the decrease on a probabilistic objective [Pham01]
    - alternately minimize $\sum_q ||\mathbf{T}(q) - \mathbf{A}\,\mathbf{\Lambda}(q)\,\mathbf{A}^\mathsf{T}||^2$ wrt $\mathbf{A}$ and $\mathbf{\Lambda}(q)$ [Yere02]. See also: [Li07] [VollO06]

Computation: bibliographical comments

- Invertible diagonalization of a collection of symmetric matrices:
  - maximizes iteratively a lower bound to the decrease on a probabilistic objective [Pham01]
  - alternately minimize $\sum_q ||\mathbf{T}(q) - \mathbf{A}\,\mathbf{\Lambda}(q)\,\mathbf{A}^\mathsf{T}||^2$ wrt $\mathbf{A}$ and $\mathbf{\Lambda}(q)$ [Yere02]. See also: [Li07] [VollO06]
  - minimizes iteratively $\sum_q ||\mathbf{T}(q) - \mathbf{A}\,\mathbf{\Lambda}(q)\,\mathbf{A}^\mathsf{T}||^2$ under the assumption that $\mathbf{A}$ is diagonally dominant [Zieh04] .

Computation: bibliographical comments

- Invertible diagonalization of a collection of symmetric matrices:
  - maximizes iteratively a lower bound to the decrease on a probabilistic objective [Pham01]
  - alternately minimize $\sum_q ||\mathbf{T}(q) - \mathbf{A}\,\boldsymbol{\Lambda}(q)\,\mathbf{A}^\mathsf{T}||^2$ wrt $\mathbf{A}$ and $\boldsymbol{\Lambda}(q)$ [Yere02]. See also: [Li07] [VollO06]
  - minimizes iteratively $\sum_q ||\mathbf{T}(q) - \mathbf{A}\,\boldsymbol{\Lambda}(q)\,\mathbf{A}^\mathsf{T}||^2$ under the assumption that $\mathbf{A}$ is diagonally dominant [Zieh04] .
- Invertible diagonalization of a collection of non symmetric matrices:

Computation: bibliographical comments

- Invertible diagonalization of a collection of symmetric matrices:
  - maximizes iteratively a lower bound to the decrease on a probabilistic objective [Pham01]
  - alternately minimize $\sum_q ||\mathbf{T}(q) - \mathbf{A}\,\mathbf{\Lambda}(q)\,\mathbf{A}^\mathsf{T}||^2$ wrt $\mathbf{A}$ and $\mathbf{\Lambda}(q)$ [Yere02]. See also: [Li07] [VollO06]
  - minimizes iteratively $\sum_q ||\mathbf{T}(q) - \mathbf{A}\,\mathbf{\Lambda}(q)\,\mathbf{A}^\mathsf{T}||^2$ under the assumption that $\mathbf{A}$ is diagonally dominant [Zieh04] .
- Invertible diagonalization of a collection of non symmetric matrices:
  - factor $\mathbf{A}$ into orthogonal and triangular parts, and perform a joint Schur decomposition [DelaDV04].

Computation: bibliographical comments

- Invertible diagonalization of a collection of symmetric matrices:
  - maximizes iteratively a lower bound to the decrease on a probabilistic objective [Pham01]
  - alternately minimize $\sum_q ||\mathbf{T}(q) - \mathbf{A}\,\mathbf{\Lambda}(q)\,\mathbf{A}^{\mathsf{T}}||^2$ wrt $\mathbf{A}$ and $\mathbf{\Lambda}(q)$ [Yere02]. See also: [Li07] [VollO06]
  - minimizes iteratively $\sum_q ||\mathbf{T}(q) - \mathbf{A}\,\mathbf{\Lambda}(q)\,\mathbf{A}^{\mathsf{T}}||^2$ under the assumption that $\mathbf{A}$ is diagonally dominant [Zieh04] .
- Invertible diagonalization of a collection of non symmetric matrices:
  - factor $\mathbf{A}$ into orthogonal and triangular parts, and perform a joint Schur decomposition [DelaDV04].
  - others: algorithms applicable to underdetermined case work here [AlbeFCC05]

Now rank larger than dimensions

Again we don't know well:

- Uniqueness
- Computation

What we know rather well

- Uniqueness in symmetric case (AH theorem)
- Computation in dimension 2 (Sylvester, matrix pencils...)

Why symmetric tensors are important

- In examples of Antenna array processing (Telecommunications, Speech, Sonar...), source signals are statistically independent
- This yields equations

$$\mathrm{E}\{f(s_i)\,g(s_j)\} = \mathrm{E}\{f(s_i)\}\,\mathrm{E}\{g(s_j)\}$$

for all $i \neq j$ and any functions f and g

Characteristic functions

First c.f.

Characteristic functions

First c.f.

- Real Scalar: $\Phi_x(t) \stackrel{\text{def}}{=} \mathrm{E}\{e^{\jmath\, tx}\} = \int_u e^{\jmath\, tu}\, dF_x(u)$
- Real Multivariate: $\Phi_{\mathbf{x}}(\mathbf{t}) \stackrel{\text{def}}{=} \mathrm{E}\{e^{\jmath\, \mathbf{t}^\mathsf{T}\mathbf{x}}\} = \int_{\mathbf{u}} e^{\jmath\, \mathbf{t}^\mathsf{T}\mathbf{u}}\, dF_{\mathbf{x}}(\mathbf{u})$

Characteristic functions

First c.f.

- Real Scalar: $\Phi_x(t) \stackrel{\text{def}}{=} \mathrm{E}\{e^{\jmath\, tx}\} = \int_u e^{\jmath\, tu}\, dF_x(u)$
- Real Multivariate: $\Phi_{\mathbf{x}}(\mathbf{t}) \stackrel{\text{def}}{=} \mathrm{E}\{e^{\jmath\, \mathbf{t}^\mathsf{T}\mathbf{x}}\} = \int_{\mathbf{u}} e^{\jmath\, \mathbf{t}^\mathsf{T}\mathbf{u}}\, dF_{\mathbf{x}}(\mathbf{u})$

Second c.f.

Characteristic functions

First c.f.

- Real Scalar: $\Phi_x(t) \stackrel{\text{def}}{=} \mathrm{E}\{e^{\jmath\, tx}\} = \int_u e^{\jmath\, tu}\, dF_x(u)$
- Real Multivariate: $\Phi_{\mathbf{x}}(\mathbf{t}) \stackrel{\text{def}}{=} \mathrm{E}\{e^{\jmath\, \mathbf{t}^\mathsf{T}\mathbf{x}}\} = \int_{\mathbf{u}} e^{\jmath\, \mathbf{t}^\mathsf{T}\mathbf{u}}\, dF_{\mathbf{x}}(\mathbf{u})$

Second c.f.

- $\Psi(\mathbf{t}) \stackrel{\text{def}}{=} \log \Phi(\mathbf{t})$
- Properties:
  - Always exists in the neighborhood of 0
  - Uniquely defined as long as $\Phi(\mathbf{t}) \neq 0$

Characteristic functions (cont'd)

- Properties of the 2nd Characteristic function (cont'd):
  - if a c.f. $\Psi_x(t)$ is a polynomial, then its degree is at most 2 and x is Gaussian *(Marcinkiewicz'1938)* [Luka70]
  - if (x, y) statistically independent, then

$$\Psi_{x,y}(u, v) = \Psi_x(u) + \Psi_y(v) \tag{1}$$

Characteristic functions (cont'd)

- Properties of the 2nd Characteristic function (cont'd):
  - if a c.f. $\Psi_x(t)$ is a polynomial, then its degree is at most 2 and x is Gaussian *(Marcinkiewicz'1938)* [Luka70]
  - if (x, y) statistically independent, then

$$\Psi_{x,y}(u, v) = \Psi_x(u) + \Psi_y(v) \tag{1}$$

Proof.

$$\begin{aligned} \Psi_{x,y}(u, v) &= \log[\mathrm{E}\{\exp(ux + vy)\}] \\ &= \log[\mathrm{E}\{\exp(ux)\}\,\mathrm{E}\{\exp(vy)\}]. \end{aligned}$$

□

General problem: Blind identification

Linear statistical model

$$\mathbf{x} = \mathbf{H}\,\mathbf{s} + \mathbf{b} \tag{2}$$

with
$$\begin{cases} \mathbf{x}: & K \times 1 \text{ random} \\ \mathbf{s}: & P \times 1 \text{ random with } \textit{stat. independent} \text{ entries} \\ \mathbf{H}: & K \times P \text{ deterministic} \\ \mathbf{b}: & \text{errors (may be removed for } P \text{ large enough)} \end{cases}$$

Problem posed in terms of Characteristic Functions

- If s_p independent and $\mathbf{x} = \mathbf{H}\,\mathbf{s}$, we have the *core equation*:

$$\Psi_x(\mathbf{u}) = \sum_p \Psi_{s_p}\left(\sum_q u_q\, H_{qp}\right) \tag{3}$$

Proof. □

Problem posed in terms of Characteristic Functions

- If s_p independent and $\mathbf{x} = \mathbf{H}\,\mathbf{s}$, we have the *core equation*:

$$\Psi_x(\mathbf{u}) = \sum_p \Psi_{s_p}\left(\sum_q u_q\, H_{qp}\right) \tag{3}$$

Proof.

- Plug $\mathbf{x} = \mathbf{H}\,\mathbf{s}$, in definition of Ψ_x and get

$$\Phi_x(\mathbf{u}) \stackrel{\text{def}}{=} \mathrm{E}\{\exp(\mathbf{u}^\mathsf{T}\mathbf{H}\,\mathbf{s})\} = \mathrm{E}\{\exp(\sum_{p,q} u_q\, H_{qp}\, s_p)\}$$

$\square$

Problem posed in terms of Characteristic Functions

- If s_p independent and $\mathbf{x} = \mathbf{H}\,\mathbf{s}$, we have the *core equation*:

$$\Psi_x(\mathbf{u}) = \sum_p \Psi_{s_p}\left(\sum_q u_q\, H_{qp}\right) \tag{3}$$

Proof.

- Plug $\mathbf{x} = \mathbf{H}\,\mathbf{s}$, in definition of Ψ_x and get

$$\Phi_x(\mathbf{u}) \stackrel{\text{def}}{=} \mathrm{E}\{\exp(\mathbf{u}^{\mathsf{T}}\mathbf{H}\,\mathbf{s})\} = \mathrm{E}\{\exp(\sum_{p,q} u_q\, H_{qp}\, s_p)\}$$

- Since s_p independent, $\Phi_x(\mathbf{u}) = \prod_p \mathrm{E}\{\exp(\sum_q u_q\, H_{qp}\, s_p)\}$

$\square$

Problem posed in terms of Characteristic Functions

- If s_p independent and $\mathbf{x} = \mathbf{H}\,\mathbf{s}$, we have the *core equation*:

$$\Psi_x(\mathbf{u}) = \sum_p \Psi_{s_p}\left(\sum_q u_q\, H_{qp}\right) \tag{3}$$

Proof.

- Plug $\mathbf{x} = \mathbf{H}\,\mathbf{s}$, in definition of Ψ_x and get

$$\Phi_x(\mathbf{u}) \stackrel{\text{def}}{=} \mathrm{E}\{\exp(\mathbf{u}^{\mathsf{T}}\mathbf{H}\,\mathbf{s})\} = \mathrm{E}\{\exp(\sum_{p,q} u_q\, H_{qp}\, s_p)\}$$

- Since s_p independent, $\Phi_x(\mathbf{u}) = \prod_p \mathrm{E}\{\exp(\sum_q u_q\, H_{qp}\, s_p)\}$
- Taking the log concludes.

$\square$

Problem posed in terms of Characteristic Functions

- If s_p independent and $\mathbf{x} = \mathbf{H}\,\mathbf{s}$, we have the *core equation*:

$$\Psi_x(\mathbf{u}) = \sum_p \Psi_{s_p}\left(\sum_q u_q\, H_{qp}\right) \tag{3}$$

Proof.

- Plug $\mathbf{x} = \mathbf{H}\,\mathbf{s}$, in definition of Ψ_x and get

$$\Phi_x(\mathbf{u}) \stackrel{\text{def}}{=} \mathrm{E}\{\exp(\mathbf{u}^{\mathsf{T}}\mathbf{H}\,\mathbf{s})\} = \mathrm{E}\{\exp(\sum_{p,q} u_q\, H_{qp}\, s_p)\}$$

- Since s_p independent, $\Phi_x(\mathbf{u}) = \prod_p \mathrm{E}\{\exp(\sum_q u_q\, H_{qp}\, s_p)\}$
- Taking the log concludes.

$\square$

Problem: Decompose a mutlivariate function into a sum of univariate ones

Equations derived from the C.F.

- Assumption: functions ψ_p, $1 \leq p \leq P$ admit finite derivatives up to order r in a neighborhood of the origin, containing $\mathcal{G}$.

Equations derived from the C.F.

- Assumption: functions ψ_p, $1 \le p \le P$ admit finite derivatives up to order r in a neighborhood of the origin, containing $\mathcal{G}$.
- Then, Taking $r = 3$ as a working example:

$$\frac{\partial^3 \Psi_x}{\partial u_i \partial u_j \partial u_k}(\mathbf{u}) = \sum_{p=1}^{P} H_{ip}\, H_{jp}\, H_{kp} \;\; \psi_p^{(3)}\Big(\sum_{q=1}^{K} u_q\, H_{qp}\Big)$$

Equations derived from the C.F.

- Assumption: functions ψ_p, $1 \leq p \leq P$ admit finite derivatives up to order r in a neighborhood of the origin, containing $\mathcal{G}$.
- Then, Taking $r = 3$ as a working example:

$$\frac{\partial^3 \Psi_x}{\partial u_i \partial u_j \partial u_k}(\mathbf{u}) = \sum_{p=1}^{P} H_{ip}\, H_{jp}\, H_{kp}\;\; \psi_p^{(3)}(\sum_{q=1}^{K} u_q\, H_{qp})$$

- If $L > 1$ point in grid $\mathcal{G}$, then yields another mode in tensor

Putting the problem in tensor form

- Takes at $L > 1$ points on a grid:

$$T_{ijk\ell} = \sum_{p} H_{ip}\, H_{jp}\, H_{kp}\, B_{\ell p}$$

or $\qquad \mathbf{T} = \sum_p \mathbf{h}(p) \otimes \mathbf{h}(p) \otimes \mathbf{h}(p) \otimes \mathbf{b}(p)$
where tensor $\mathbf{T}$ is $K \times K \times K \times L$, and partially symmetric

Putting the problem in tensor form

- Takes at $L > 1$ points on a grid:

$$T_{ijk\ell} = \sum_p H_{ip}\, H_{jp}\, H_{kp}\, B_{\ell p}$$

  or $\quad \mathbf{T} = \sum_p \mathbf{h}(p) \otimes \mathbf{h}(p) \otimes \mathbf{h}(p) \otimes \mathbf{b}(p)$
  where tensor $\mathbf{T}$ is $K \times K \times K \times L$, and partially symmetric

- Take only equations at the origin: *Cumulant tensor*:

$$\mathbf{T} = \sum_p \mathbf{h}(p) \otimes \mathbf{h}(p) \otimes \mathbf{h}(p)$$

Putting the problem in tensor form

- Takes at $L > 1$ points on a grid:

$$T_{ijk\ell} = \sum_p H_{ip}\, H_{jp}\, H_{kp}\, B_{\ell p}$$

  or $\quad \mathbf{T} = \sum_p \mathbf{h}(p) \otimes \mathbf{h}(p) \otimes \mathbf{h}(p) \otimes \mathbf{b}(p)$
  where tensor $\mathbf{T}$ is $K \times K \times K \times L$, and partially symmetric

- Take only equations at the origin: *Cumulant tensor*:

$$\mathbf{T} = \sum_p \mathbf{h}(p) \otimes \mathbf{h}(p) \otimes \mathbf{h}(p)$$

- One can get equations at *arbitrary orders*.

Joint use of different derivative orders

Example

- Derivatives of order 3:

$$T^{(3)}_{ijk\ell} = \sum_p H_{ip}\, H_{jp}\, H_{kp}\, B_{\ell p}$$

Joint use of different derivative orders

Example

- Derivatives of order 3:

$$T^{(3)}_{ijk\ell} = \sum_p H_{ip}\, H_{jp}\, H_{kp}\, B_{\ell p}$$

- Derivatives of order 4:

$$T^{(4)}_{ijkm\ell} = \sum_p H_{ip}\, H_{jp}\, H_{kp}\, H_{mp}\, C_{\ell p}$$

Joint use of different derivative orders

Example

- Derivatives of order 3:

$$T^{(3)}_{ijk\ell} = \sum_p H_{ip}\, H_{jp}\, H_{kp}\, B_{\ell p}$$

- Derivatives of order 4:

$$T^{(4)}_{ijkm\ell} = \sum_p H_{ip}\, H_{jp}\, H_{kp}\, H_{mp}\, C_{\ell p}$$

- Derivatives of orders 3 and 4:

$$T_{ijk\ell}[m] = \sum_p H_{ip}\, H_{jp}\, H_{kp}\, D_{\ell p}[m]$$

with $D_{\ell p}[m] = H_{mp}\, C_{\ell p}$ and $D_{\ell p}[0] = B_{\ell p}$.

Problem P18

- Let two tensors $\mathbf{T}_1 \in \mathcal{A} = \mathcal{S}^3 V_1 \otimes V_2$ and $\mathbf{T}_2 \in \mathcal{B} = \mathcal{S}^2 V_1 \otimes V_3$ be defined by decompositions
$$\mathbf{T}_1 = \sum_{p=1}^{r} \mathbf{a}(p) \otimes \mathbf{a}(p) \otimes \mathbf{a}(p) \otimes \mathbf{b}(p) \text{ and } \mathbf{T}_2 = \sum_{p=1}^{r} \mathbf{a}(p) \otimes \mathbf{a}(p) \otimes \mathbf{c}(p)$$
where r is the generic rank in $\mathcal{B}$.
- We are given $\tilde{\mathbf{T}}_1 = \mathbf{T}_1 + \mathbf{E}_1$ and $\tilde{\mathbf{T}}_2 = \mathbf{T}_2 + \mathbf{E}_2$ where $\mathbf{E}_i$ are small.
- How can we compute the $\mathbf{a}(p)$'s, $1 \leq p \leq r$ from $\tilde{\mathbf{T}}_1$ and $\tilde{\mathbf{T}}_2$?

Alexander-Hirschowitz theorem

Theorem (1995) For $d > 2$, the generic rank of a dth order symmetric tensor of dimension K is *always* equal to the lower bound

$$\bar{R}_s = \left\lceil \frac{\binom{K+d-1}{d}}{K} \right\rceil \tag{4}$$

except for the following cases:
$(d, K) \in \{(3,5), (4,3), (4,4), (4,5)\}$, for which it should be increased by 1 (i.e. only a *finite number* of exceptions, also called *defective* cases)

Values of the Generic Rank (1)

Symmetric tensors of order d and dimension K

d \ K	2	3	4	5	6	7	8
3	**2**	4	**5**	**8**	10	**12**	**15**
4	3	**6**	**10**	**15**	**21**	**30**	42

$$\bar{R}_s \geq \frac{1}{K}\binom{K+d-1}{d}$$

Bold: exceptions to the ceil rule: $\bar{R}_s = \lceil \frac{1}{K}\binom{K+d-1}{d} \rceil$, sometimes called *defective* cases.
Green: lower bound $\frac{1}{K}\binom{K+d-1}{d}$ is integer and nondefective, hence finite number of solutions with proba 1

Uniqueness

- Uniqueness, or at least finite number of solutions
- *Terracini's lemma* allows to compute the dimension of any secant variety via the tangent space
- and in particular Segre, Veronese, or any other enjoying special symmetry properties.

Values of the Generic Rank (2)

Warning: for *unsymmetric* tensors of order d and dimension K, the generic rank is different

d \ K	2	3	4	5	6	7
3	**2**	**5**	7	10	14	19
4	4	**9**	20	37	62	97

$$\bar{R} \geq \frac{K^d}{Kd - d + 1}$$

Bold: exceptions to the ceil rule: $\bar{R} = \lceil \frac{K^d}{Kd-d+1} \rceil$.

Green: lower bound $\frac{K^d}{Kd-d+1}$ is integer and nondefective

Numerical computation of the Generic Rank

Mapping (for unsymmetric tensors):

$$\{\mathbf{u}(\ell), \mathbf{v}(\ell), \dots, \mathbf{w}(\ell),\ 1 \leq \ell \leq r\} \quad \xrightarrow{\varphi} \quad \sum_{\ell=1}^{r} \mathbf{u}(\ell) \otimes \mathbf{v}(\ell) \otimes \dots \otimes \mathbf{w}(\ell)$$

$$\{\mathbb{C}^{n_1} \otimes \dots \otimes \mathbb{C}^{n_d}\}^r \quad \xrightarrow{\varphi} \quad \mathcal{A}$$

- The *smallest r* for wich $rank(\mathrm{Jacobian}(\varphi)) = \prod_i n_i$ is the generic rank, $\bar{R}$.
- Example of use of *Terracini's lemma*

First example of computation of Generic Rank

$$\{\mathbf{a}(\ell), \mathbf{b}(\ell), \mathbf{c}(\ell)\} \xrightarrow{\varphi} \mathbf{T} = \sum_{\ell=1}^{r} \mathbf{a}(\ell) \otimes \mathbf{b}(\ell) \otimes \mathbf{c}(\ell)$$

$\mathbf{T}$ has coordinate vector: $\sum_{\ell=1}^{r} \mathbf{a}(\ell) \otimes \mathbf{b}(\ell) \otimes \mathbf{c}(\ell)$. Hence the Jacobian of φ is the $r(n_1 + n_2 + n_3) \times n_1 n_2 n_3$ matrix:

$$\mathbf{J} = \begin{bmatrix} \mathbf{I}_{n_1} & \otimes & \mathbf{b}^{\mathsf{T}}(1) & \otimes & \mathbf{c}^{\mathsf{T}}(1) \\ \vdots & \otimes & \vdots & \otimes & \vdots \\ \mathbf{I}_{n_1} & \otimes & \mathbf{b}^{\mathsf{T}}(r) & \otimes & \mathbf{c}^{\mathsf{T}}(r) \\ \mathbf{a}(1)^{\mathsf{T}} & \otimes & \mathbf{I}_{n_2} & \otimes & \mathbf{c}^{\mathsf{T}}(1) \\ \vdots & \otimes & \vdots & \otimes & \vdots \\ \mathbf{a}(r)^{\mathsf{T}} & \otimes & \mathbf{I}_{n_2} & \otimes & \mathbf{c}^{\mathsf{T}}(r) \\ \mathbf{a}(1)^{\mathsf{T}} & \otimes & \mathbf{b}(1)^{\mathsf{T}} & \otimes & \mathbf{I}_{n_3} \\ \vdots & \otimes & \vdots & \otimes & \vdots \\ \mathbf{a}(r)^{\mathsf{T}} & \otimes & \mathbf{b}(r)^{\mathsf{T}} & \otimes & \mathbf{I}_{n_3} \end{bmatrix} \quad \text{and} \quad \left\{ \begin{array}{l} \text{rank}\{\mathbf{J}\} = dim(\mathrm{Im}(\varphi)) \\ \bar{R} = Min\{r : \ \mathrm{Im}\{\varphi\} = \mathcal{A}\} \end{array} \right.$$

Problem P5

- Similar theorem as AH for unsymmetric tensors?
- that is, decomposition of homogeneous polynomials of degree d but partial degree 1 into sum of products of linear forms
- Uniqueness?

Second Example: third order real tensors with symmetric slices

Typical ranks for $N_1 \times N_2 \times N_2$ arrays, with $N_2 \times N_2$ real symmetric slices.

N_1 N_2	2	3	4	5
2	2,3	3,4	4,5	5,6
3	3	4	**6**	**7**
4	3	4,5	**6**	**8**
5	3	5,6	**7**	**9**
6	3	6	**7**	**9**
7	3	6	**7**	**10**
8	3	6	**8**	**10**
9	3	6	9,10	**11**
10	3	6	10	**11**

Bold: smallest typical ranks computed numerically.

Plain: known typical ranks; in $\mathbb{C}$, the smallest value is generic.

Genericity in $\mathbb{R}$ vs. $\mathbb{C}$

Define $\mathcal{Z}_r = \{\text{tensors of rank } r\}$

- A rank r is *typical* if $\mathcal{Z}_r$ is Zariski-dense

Genericity in $\mathbb{R}$ vs. $\mathbb{C}$

Define $\mathcal{Z}_r = \{\text{tensors of rank } r\}$

- A rank r is *typical* if $\mathcal{Z}_r$ is Zariski-dense
- In the complex field, there is only one typical rank, called the *generic rank*.

Genericity in $\mathbb{R}$ vs. $\mathbb{C}$

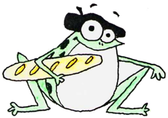
Define $\mathcal{Z}_r = \{\text{tensors of rank } r\}$

- A rank r is *typical* if $\mathcal{Z}_r$ is Zariski-dense
- In the complex field, there is only one typical rank, called the *generic rank*.
- In the real field, there can be *several* typical ranks (smallest equals generic rank in $\mathbb{C}$)

Problem P10

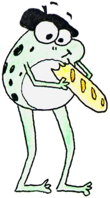
Case of the Real field

- Similar result as AH theorem in the real field?
- i.e. values of typical ranks for any order and dimensions



Problem P13

- Uniqueness results for tensors enjoying partial symmetries
- e.g., symmetric in the 3 first modes and not not in others
- e.g. tensors with symmetries in some modes and Hermitean symmetries in others
- Generic/typical ranks?
- Existence/well-poseness?
- Computation?



Problem P7

Computation of a decomposition via flattening matrices when unique (e.g. general tensors with subgeneric ranks [ChiaC06])

Let V_i be spaces of dimension n_i, and for some chosen ℓ, let

$$p \stackrel{\text{def}}{=} \sum_{i=1}^{\ell} n_i, \; q \stackrel{\text{def}}{=} \sum_{i=\ell+1}^{d} n_i, \;\; \mathcal{B} \stackrel{\text{def}}{=} V_1 \otimes \dots V_\ell \text{ and } \mathcal{C} \stackrel{\text{def}}{=} V_{\ell+1} \otimes \dots V_d$$

Let $\mathbf{T} \in \mathcal{B} \otimes \mathcal{C}$ be a tensor of rank at most $min(p, q)$.

- Associate $\mathbf{T}$ with a linear operator φ from $\mathcal{B}^* \to \mathcal{C}$, defined by a matrix $\mathbf{M}$ of size $p \times q$.
- Compute a basis $\{\mathbf{a}(k)\}_{1 \le k \le r}$ of $Im\{\varphi\}$
- Find *all* linear combinations $\mathbf{b} = \sum_k \lambda_k \mathbf{a}(k)$ such that $\mathbf{b}$ represents a tensor of rank 1 in $\mathcal{C}$.

How can we solve this quadratic system of equations in λ_i ?

Problem P4

Lower rank approximation

- One can reduce the rank by truncating the basis $\{\mathbf{a}(k)\}$ of $Im\{\varphi\}$
- This may restore uniqueness
- Sometimes hint on expected rank from practical problem
- But ill posed problem because of lack of closeness

FOOBI algorithms

Given a $K^2 \times P$ matrix $\mathbf{H}^{\odot 2}$, find a real orthogonal matrix $\mathbf{Q}$ such that the P columns of $\mathbf{H}^{\odot 2}\,\mathbf{Q}$ are of the form $\mathbf{h}[p] \otimes \mathbf{h}[p]^*$

FOOBI algorithms

Given a $K^2 \times P$ matrix $\mathbf{H}^{\odot 2}$, find a real orthogonal matrix $\mathbf{Q}$ such that the P columns of $\mathbf{H}^{\odot 2}\,\mathbf{Q}$ are of the form $\mathbf{h}[p] \otimes \mathbf{h}[p]^*$

- **FOOBI:** use the K^4 determinantal equations characterizing rank-1 matrices $\mathbf{h}[p]\,\mathbf{h}[p]^{\mathsf{H}}$ of the form:
  $\phi(\mathbf{X},\mathbf{Y})_{ijk\ell} = x_{ij}y_{\ell k} - x_{ik}y_{\ell j} + y_{ij}x_{\ell k} - y_{ik}x_{\ell j}$

FOOBI algorithms

Given a $K^2 \times P$ matrix $\mathbf{H}^{\odot 2}$, find a real orthogonal matrix $\mathbf{Q}$ such that the P columns of $\mathbf{H}^{\odot 2}\,\mathbf{Q}$ are of the form $\mathbf{h}[p] \otimes \mathbf{h}[p]^*$

- **FOOBI:** use the K^4 determinantal equations characterizing rank-1 matrices $\mathbf{h}[p]\,\mathbf{h}[p]^{\mathsf{H}}$ of the form: $\phi(\mathbf{X}, \mathbf{Y})_{ijk\ell} = x_{ij} y_{\ell k} - x_{ik} y_{\ell j} + y_{ij} x_{\ell k} - y_{ik} x_{\ell j}$
- **FOOBI2:** use the K^2 equations of the form: $\Phi(\mathbf{X}, \mathbf{Y}) = \mathbf{X}\,\mathbf{Y} + \mathbf{Y}\,\mathbf{X} - \mathrm{trace}\{\mathbf{X}\}\mathbf{Y} - \mathrm{trace}\{\mathbf{Y}\}\mathbf{X}$ where matrices $\mathbf{X}$ and $\mathbf{Y}$ are $K \times K$ Hermitean.

Computation: bibliographical comments

Decomposition of symmetric tensors with rank larger than dimension:

- FOOBI: Quite interesting recent algorithms based on matrix slices and rank-1 detecting criteria [DelaCC07]

  Lieven DeLathauwer

Decomposition of general tensors with rank larger than dimensions:

- Basically all iterative

Computation via iterative algorithms

Many practitioners execute more or less brute force minimizations of $||\mathbf{T} - \sum_{p=1}^{r} \mathbf{u}_p \otimes \mathbf{v}_p \otimes \dots \otimes \mathbf{w}_p||^2$

- Gradient with fixed or variable (ad-hoc) stepsize
- Alternate Least Squares (ALS)
- Levenberg-Marquardt
- Newton
- Conjugate gradient...

Remarks

- Hessian is generally huge, but sparse
- Problem of local minima:
  - *ELS variants* for all of the above
  - *initial point* could be provided by flattening matrices

Conclusion

Lack of special purpose algorithmic tools. Suboptimal because:

- Some “open” problems listed
- either minimize 2 successive criteria instead of a single one
- or treat a tensor as a collection of matrices
- or ignore some symmetries
- or are iterative without global convergence proof
- or need rank reduction
- or all together

Conclusion

Lack of special purpose algorithmic tools. Suboptimal because:

- Some "open" problems listed
- either minimize 2 successive criteria instead of a single one
- or treat a tensor as a collection of matrices
- or ignore some symmetries
- or are iterative without global convergence proof
- or need rank reduction
- or all together

One starts to know the scale of our ignorance...

Ignorance is the necessary condition for human being happiness.
Anatole France (1844-1924)