Multivariate statistical analyses for MEG data

Bratislav Mišić

Montréal Neurological Institute

November 7th 2016
Towards multivariate analysis

Estimating functional connectivity


Niso et al. (2013) Neuroinformatics
Bastos & Schoffelen (2016) Front Syst Neurosci
Colcough et al. (2016) NeuroImage
Constructing functional networks

atlas-guided beamforming → 'virtual sensor' time-series → functional bands → Hilbert-transform

functional brain networks

Dunkley et al. (2015) *NeuroImage: Clinical*
Why multivariate statistics?

1. how to operationalize network property?
2. how to deal with more variables than observations?
3. how to relate multiple data sets to one another?
Principal component analysis (PCA)

Hotelling (1933) *J Educ Psychol*
Principal component analysis (PCA)

Hotelling (1933) *J Educ Psychol*
Principal component analysis (PCA)

Hotelling (1933) *J Educ Psychol*
Principal component analysis (PCA)

Hotelling (1933) *J Educ Psychol*
Principal component analysis (PCA)

Hotelling (1933) *J Educ Psychol*
Maximizing variance

find new variable \( z = X \mathbf{u} \)

choose \( \mathbf{u} \) to maximize \( \text{var}(z) \)

under the constraint \( \mathbf{u}'\mathbf{u} = 1 \)

\[
\text{var}(z) = \frac{1}{n-1} \mathbf{u}'X'X\mathbf{u} = \mathbf{u}'\mathbf{R}\mathbf{u}
\]

since \( \mathbf{R} = \frac{1}{n-1} X'X \)
Maximizing variance

\[ L = u'Ru - \lambda(u'u - 1) \]

\[ \frac{\partial L}{\partial u} = 2Ru - 2u\lambda = 0 \]

\[ Ru = u\lambda \]

\[ (R - \lambda I)u = 0 \]

eigenvalue \( \lambda \) (variance) & eigenvector \( u \) (weights)

\[ var(z) = u'Ru = u'u\lambda = \lambda \]
Singular value decomposition

Spectral decomposition:
\[ \text{EIG}(X'X) = U \Lambda U' \]
\[ \text{EIG}(XX') = V \Lambda V' \]

Singular value decomposition:
\[ \text{SVD}(X) = USV' \]

Eckart & Young (1936) *Psychometrika*
Singular value decomposition

Spectral decomposition:
\[ \text{EIG}(X'X) = U \Lambda U' \]
\[ \text{EIG}(XX') = V \Lambda V' \]

Singular value decomposition:
\[ \text{SVD}(X) = USV' \]

Eckart & Young (1936) *Psychometrika*
A family of techniques

PCA: $\text{SVD}(X)$

PLS: $\text{SVD}(X'Y)$

CCA: $\text{SVD}((X'X')^{-1/2}(X'Y)(Y'Y)^{-1/2})$

Worsley et al. (1997) *NeuroImage*

Tijl et al. (2005) *Handbook of Geometric Computing - Springer*

McIntosh & Mišić (2013) *Annu Rev Psychol*
Related techniques: ICA and FA

image: Gaël Varoquaux
Example: resting-state MEG networks

Brookes et al. (2011) Proc Natl Acad Sci USA
Partial least squares (PLS)

Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*
Partial least squares (PLS)

\[ n \text{ subjects} \]
\[ \text{k connections} \]
\[ X \]

\[ \text{p contrasts / behaviours} \]
\[ X'Y \]
\[ \text{U} \]
\[ \text{S} \]
\[ \text{V'} \]

\[ k = \]

Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*
Partial least squares (PLS)

Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*
Partial least squares (PLS)

Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*
How many components to retain?

\[
\text{\% covariance} = \frac{s_i^2}{\sum_j s_j^2}
\]
Statistical significance: permutation tests

Edgington (1965, 1969) *J Psychol*

McIntosh et al. (1996, 2004) *NeuroImage*
Statistical significance: permutation tests

Edgington (1965, 1969) *J Psychol*
McIntosh et al. (1996, 2004) *NeuroImage*
How many components to retain?
How many components to retain?

![Graph showing covariance accounted for (%) against latent variable. The x-axis represents the latent variable, and the y-axis represents the covariance accounted for (%). The graph includes a line at a p-value of 0.05, indicating a significant threshold.]

- Covariance accounted for (%)
- Latent variable
- P-value

- 0 20 40 60 80 100 120 140 160
- 0 2 4 6 8 10
- 0 0.2 0.4 0.6 0.8 1
- 0 0.05 0.1
How many components to retain?

The diagram shows a plot with the x-axis labeled as "latent variable" and the y-axis labeled as "covariance accounted for (%)". The p-value is also shown on the same y-axis range. The dots on the graph are selected for further analysis when the p-value is below 0.05.
Which connections to focus on?

- **n subjects**
- **k connections**
- **p contrasts / behaviours**

\[
\begin{align*}
  \text{n subjects} & : X \\
  \text{k connections} & : X'Y \\
  \text{p contrasts / behaviours} & : U \\
  \end{align*}
\]
Reliability: bootstrapping

Efron & Tibshirani (1986) *Stat Sci*
McIntosh et al. (1996, 2004) *NeuroImage*
Reliability: bootstrapping

\[ X \times Y = X'Y \]

\[ X_{boot} \times Y_{boot} = X_{boot}'Y_{boot} = X_{boot}'Y_{boot} = X_{boot}'Y_{boot} \]

\[ bsr = \frac{weight}{SE(weight)} \]

Efron & Tibshirani (1986) *Stat Sci*

McIntosh et al. (1996, 2004) *NeuroImage*
Example: connectivity differentiates PTSD from mTBI

$p < 10^{-10}$, crossblock covariance = 31%

connections

bootstrap ratio (reliability of connection)

connections

Mišić et al. (2016) J Neurosci
Individual participants

- $n$ subjects
- $k$ connections

$F_i = XU_i$

Connection strength

Statistical weight
Example: connectivity and symptom severity in PTSD

\[ p < 10^{-10}, \text{crossblock covariance} = 31\% \]

![Bar chart showing contrasts between groups: military controls, PTSD, civilian controls, and mTBI. The y-axis represents optimal contrast, with green bars for resting 1 and blue bars for resting 2.]

![Scatter plot showing a positive correlation between brain scores (functional connectivity) and PTSD Check List (PCL-M) scores. The Pearson correlation coefficient \( r = 0.63 \), and the significance level \( p < 0.01 \).]
Adding other dimensions

Krishnan et al. (2011) *NeuroImage*
Adding other dimensions

Krishnan et al. (2011) *NeuroImage*
Example: connectivity and mood in major depression

LV 1, p = 0.006, crossblock covariance = 53.5%

Berman et al. (2014) *NeuroImage*
Cross-validation: split-half resampling

1. Randomly permute data.
2. Decompose data \( D \) with SVD: \( D = USV' \).
3. Split into halves and project:
   - For the first half: \( U_1 = D_1 VS^{-1} \) and \( V_1 = D_1 US^{-1} \).
   - For the second half: \( U_2 = D_2 VS^{-1} \) and \( V_2 = D_2 US^{-1} \).
4. Correlate \( U_1 \) & \( U_2 \) and \( V_1 \) & \( V_2 \).
5. Mean correlation across splits.
6. Compare against null distribution.

Strother et al. (2002) *NeuroImage*
Canonical correlation analysis (CCA)

\[ SVD\left( (X'X')^{-1/2}(X'Y)(Y'Y)^{-1/2} \right) \]

Hotelling (1936) *Biometrika*
Canonical correlation analysis (CCA)

\[ \text{SVD}\left((X'X')^{-1/2}(X'Y)(Y'Y)^{-1/2}\right) \]

Hotelling (1936) *Biometrika*
Canonical correlation analysis (CCA)

\[ \text{SVD}((X'X')^{-1/2}(X'Y)(Y'Y)^{-1/2}) \]

n subjects
k connections
p contrasts / behaviours

\( X \)
\( Y \)

Hotelling (1936) *Biometrika*
Canonical correlation analysis (CCA)

$$\text{SVD}\left(\left(\mathbf{X}'\mathbf{X}'\right)^{-1/2}\left(\mathbf{X}'\mathbf{Y}\right)\left(\mathbf{Y}'\mathbf{Y}\right)^{-1/2}\right)$$

$n$ subjects

$k$ connections

$p$ contrasts / behaviours

Hotelling (1936) *Biometrika*
Example: relating connectivity and behaviour

Smith et al. (2015) *Nat Neurosci*
Linear discriminant analysis (LDA)

\[ \text{SVD}(W^{-1}B) \]

image: https://mlalgorithim.wordpress.com/
Friston et al. (1995) *NeuroImage*
Extensions

- sparse/regularized solutions, e.g.
  PLS-CA: Beaton et al. (2015) Psychol Meth

- extensions to 3+ data sets, e.g.
  PARAFAC: Bro (1997) Chemometr Intell Lab
  Multiway PLS: Wold et al. (1987) J Chemometrics

- nonlinear dependencies

- Bayesian implementations, e.g.
  IBFA: Virtanen et al. (2011) ICML-II

- prediction
Limitations and considerations

- overfitting
- linear
- unique partitioning of variance/covariance
- inference on individual variables
Multivariate models embody network property.
All techniques entail unique assumptions.
Many linear multivariate techniques are related.
Multivariate techniques are versatile.
Resources

- Matlab toolbox for dimensionality reduction
  https://lvdmaaten.github.io/drtoolbox/
- PLS toolbox
  https://www.rotman-baycrest.on.ca/index.php?section=84
- sparse CCA
  http://statweb.stanford.edu/~tibs/Correlate/
- coming soon to Brainstorm
Matching randomized components

\[ X'Y = USV' \iff X'_{\text{boot}} Y = U_{\text{boot}} S_{\text{boot}} V'_{\text{boot}} \]