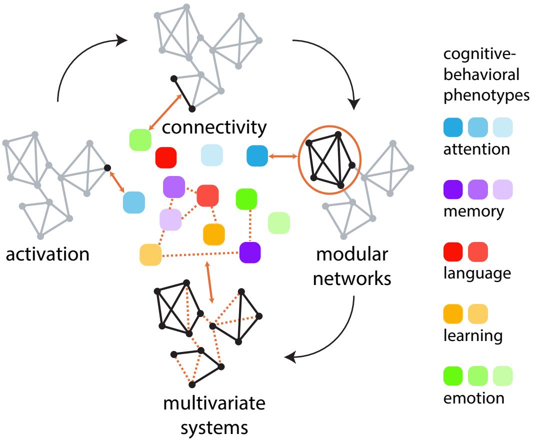
Multivariate Linear Models

Bratislav Mišić

NEUR 603

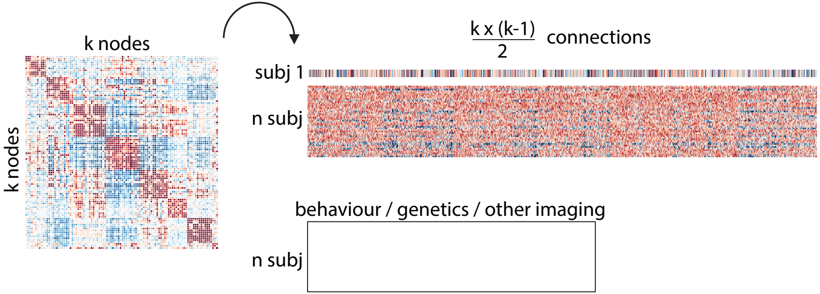
March 21^{st} 2018

Towards multivariate analysis



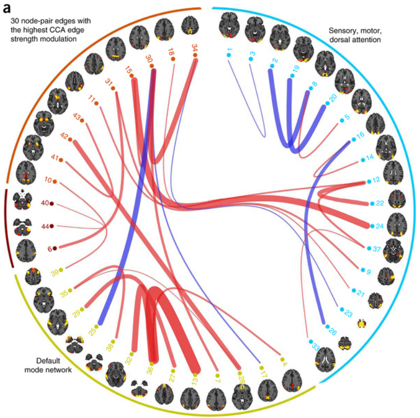
Mišić & Sporns (2016) *Curr Opin Neurobiol*

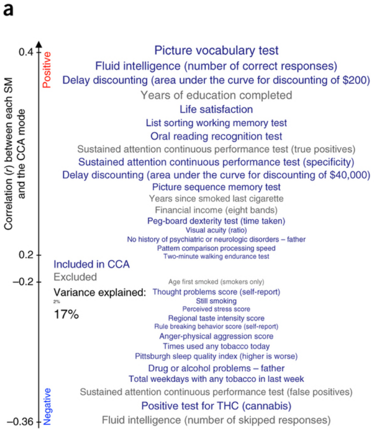
Why multivariate statistics?



1. how to operationalize network property?
2. how to deal with with more variables than observations?
3. how to relate multiple data sets to one another?

Example: relating connectivity and behaviour



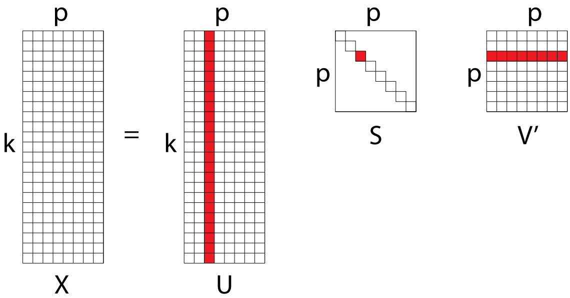


Smith et al. (2015) *Nat Neurosci*

Singular value decomposition

Spectral decomposition:
$\mathrm{EIG}(\mathbf{X}'\mathbf{X}) = \mathbf{U\Lambda U}'$
$\mathrm{EIG}(\mathbf{XX}') = \mathbf{V\Lambda V}'$

Singular value decomposition:
$\mathrm{SVD}(\mathbf{X}) = \mathbf{USV}'$



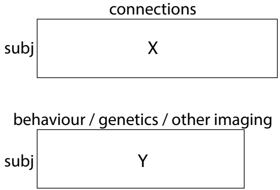
Eckart & Young (1936) *Psychometrika*

A family of techniques

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

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

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



Worsley et al. (1997) *NeuroImage*

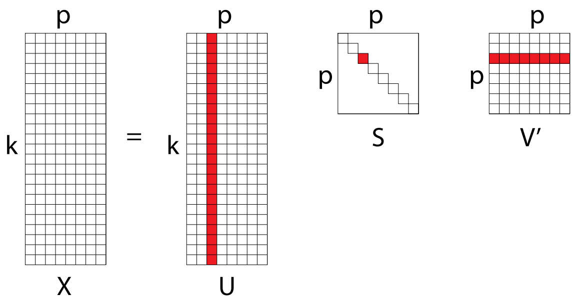
De Bie et al. (2005) *Handbook of Geometric Computing - Springer*

McIntosh & Mišić (2013) *Annu Rev Psychol*

Singular value decomposition

Spectral decomposition:
$\mathrm{EIG}(\mathbf{X}'\mathbf{X}) = \mathbf{U\Lambda U}'$
$\mathrm{EIG}(\mathbf{XX}') = \mathbf{V\Lambda V}'$

Singular value decomposition:
$\mathrm{SVD}(\mathbf{X}) = \mathbf{USV}'$



Eckart & Young (1936) *Psychometrika*

EIG vs SVD

SVD: $\mathbf{X} = \mathbf{USV}'$

$$
\begin{aligned}
\text{EIG: } \mathbf{X}'\mathbf{X} &= (\mathbf{VS}'\mathbf{U}')(\mathbf{USV}') \\
&= \mathbf{VS}'(\mathbf{U}'\mathbf{U})\mathbf{SV}' \\
&= \mathbf{V}(\mathbf{S}'\mathbf{S})\mathbf{V}'
\end{aligned}
$$

$$
\begin{aligned}
\text{EIG: } \mathbf{XX}' &= (\mathbf{USV}')(\mathbf{VS}'\mathbf{U}') \\
&= \mathbf{US}(\mathbf{V}'\mathbf{V})\mathbf{S}'\mathbf{U}' \\
&= \mathbf{U}(\mathbf{SS}')\mathbf{U}'
\end{aligned}
$$

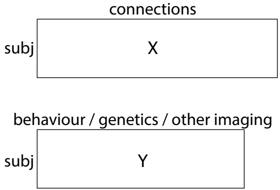
1. eigenvector of $(\mathbf{XX}')$ = left singular vector $(\mathbf{U})$
2. eigenvector of $(\mathbf{X}'\mathbf{X})$ = right singular vector $(\mathbf{V})$
3. eigenvalue = squared singular value

A family of techniques

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

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

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

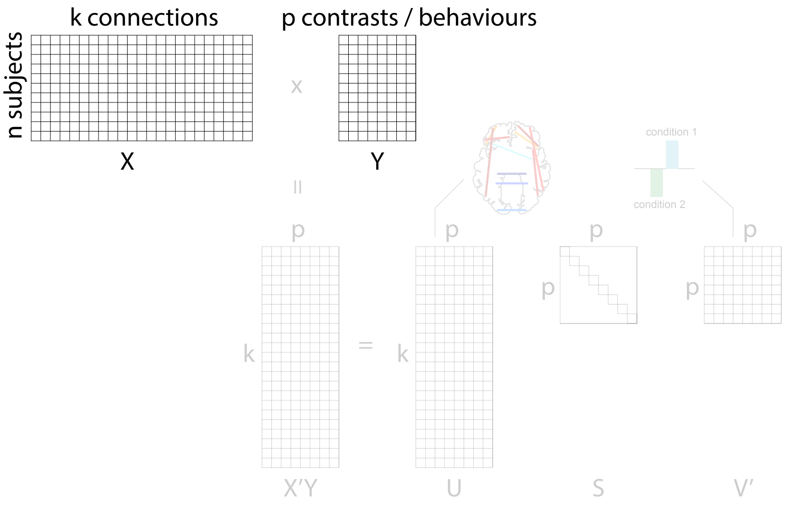


Worsley et al. (1997) *NeuroImage*

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

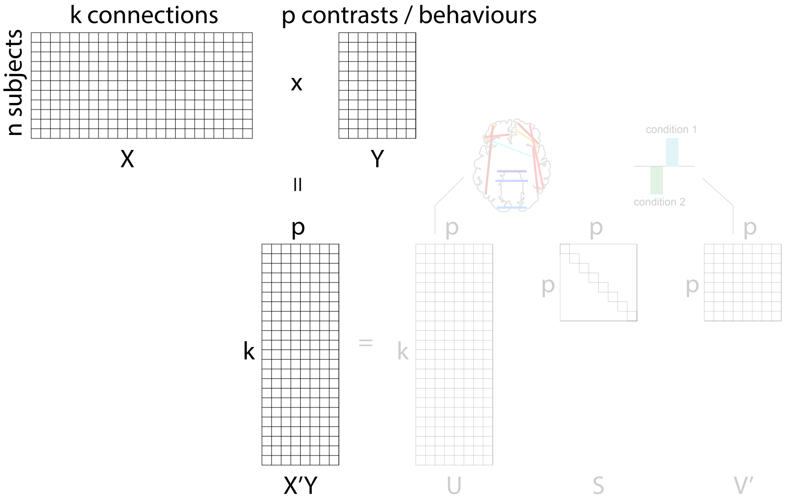
McIntosh & Mišić (2013) *Annu Rev Psychol*

Partial least squares (PLS)



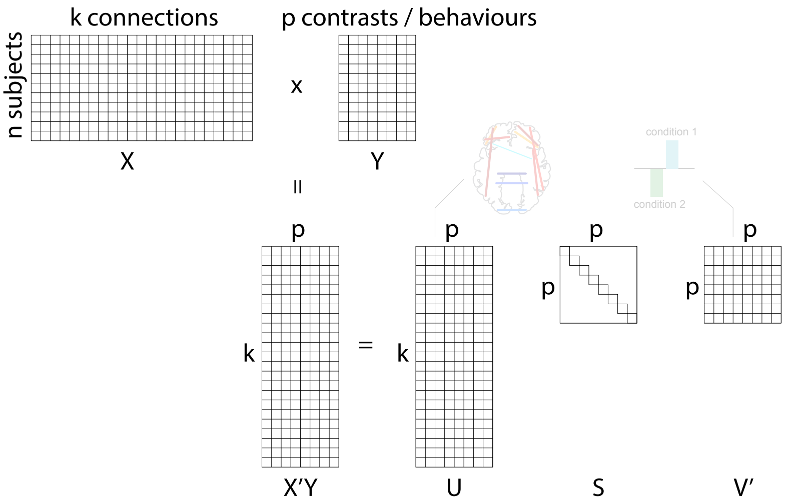
Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*

Partial least squares (PLS)



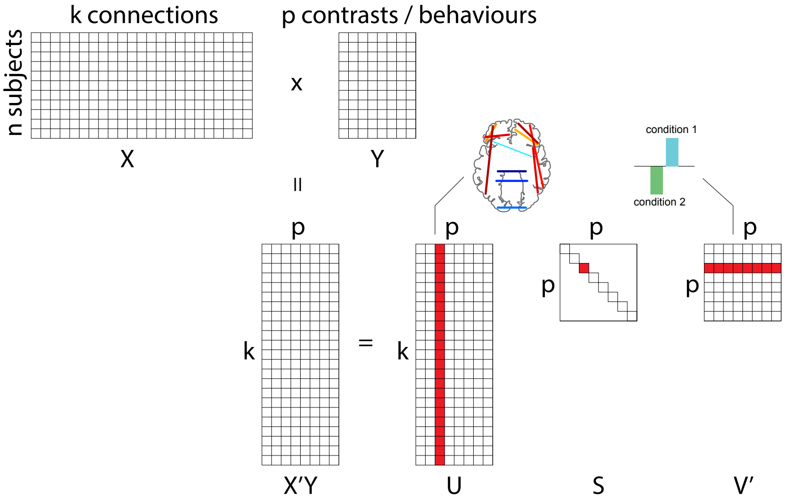
Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*

Partial least squares (PLS)



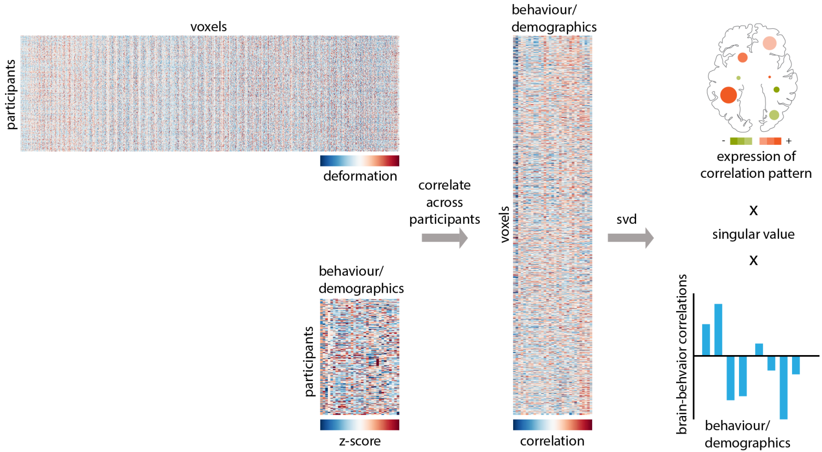
Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*

Partial least squares (PLS)



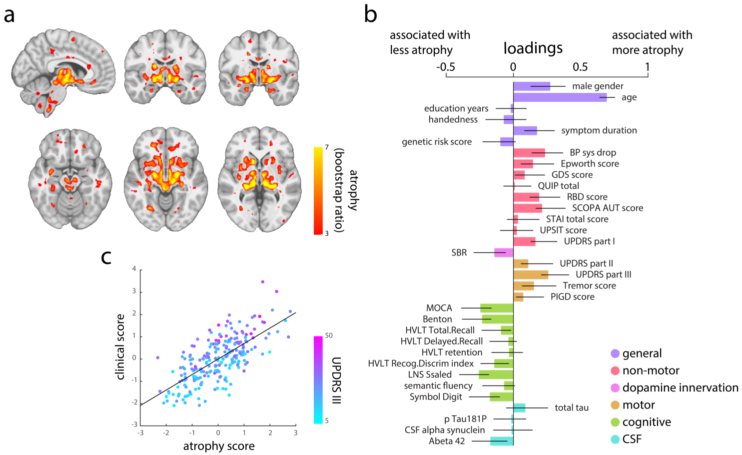
Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*

Example: clinical-anatomical signature of Parkinson's



Zeighami et al. (2018) *NeuroImage*

Example: clinical-anatomical signature of Parkinson's



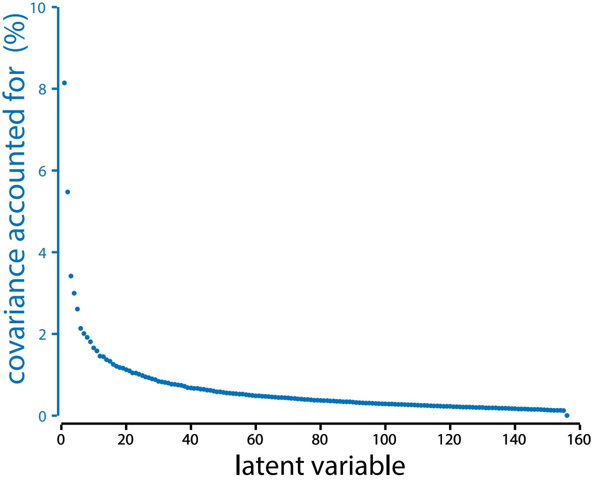
Zeighami et al. (2018) *NeuroImage*

Follow-up questions

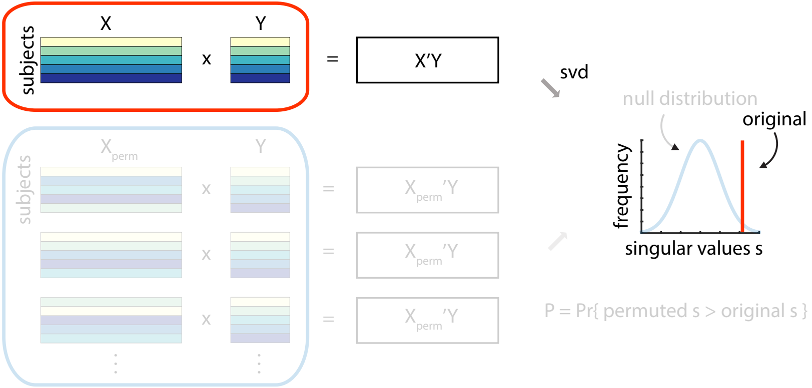
- which components are statistically meaningful?
- which variables are most important?
- how well to individual subjects express the multivariate pattern?

How many components to retain?

$$\%\ \text{covariance} = s_i^2 / \sum_j s_j^2$$



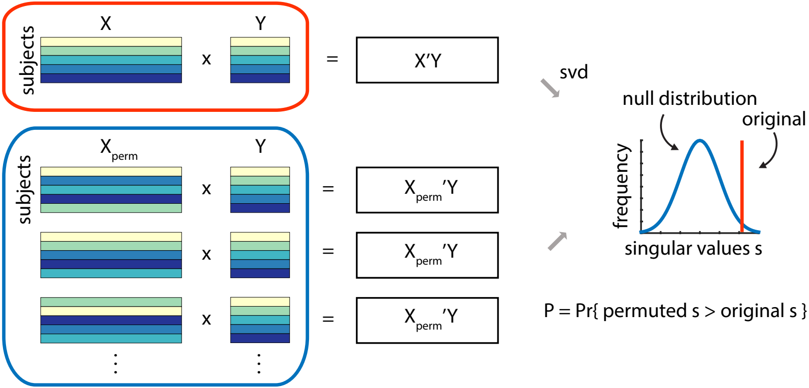
Statistical significance: permutation tests



Edgington (1965, 1969) *J Psychol*

McIntosh et al. (1996, 2004) *NeuroImage*

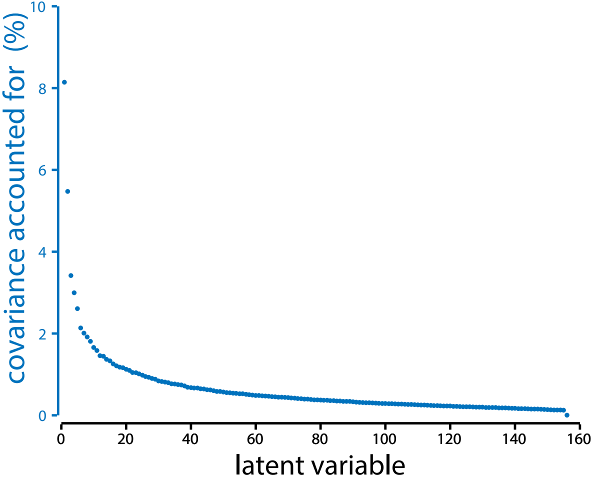
Statistical significance: permutation tests



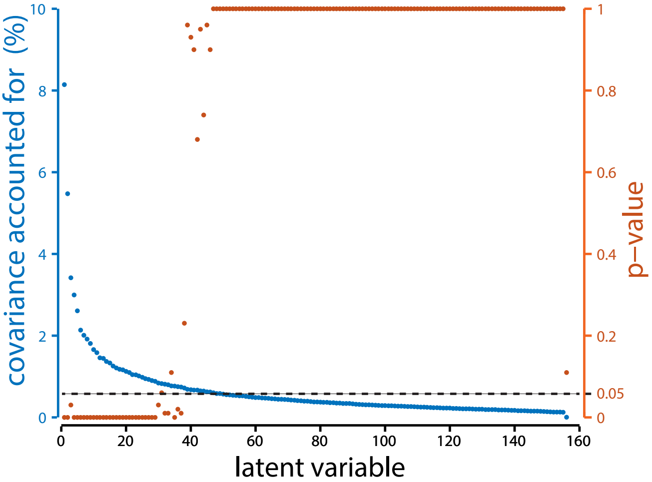
Edgington (1965, 1969) *J Psychol*

McIntosh et al. (1996, 2004) *NeuroImage*

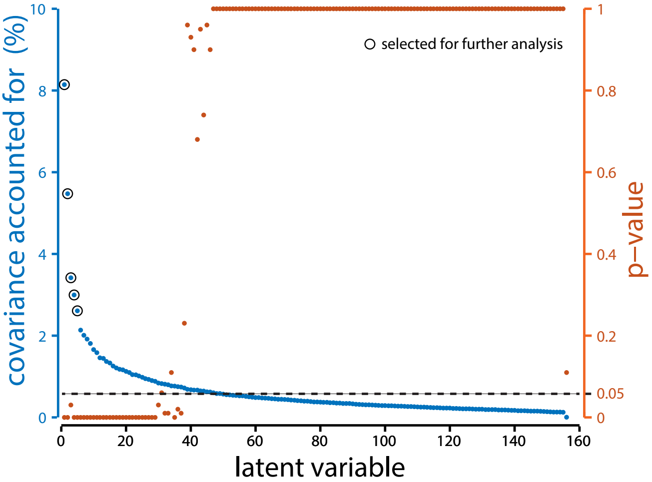
How many components to retain?



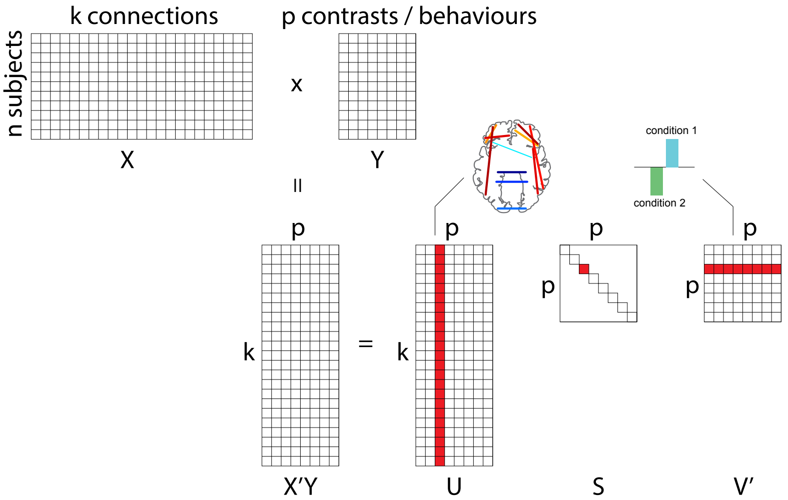
How many components to retain?



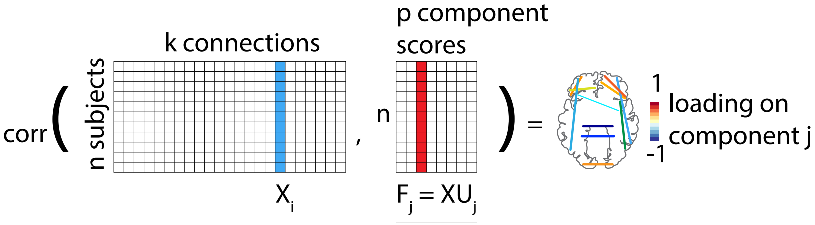
How many components to retain?



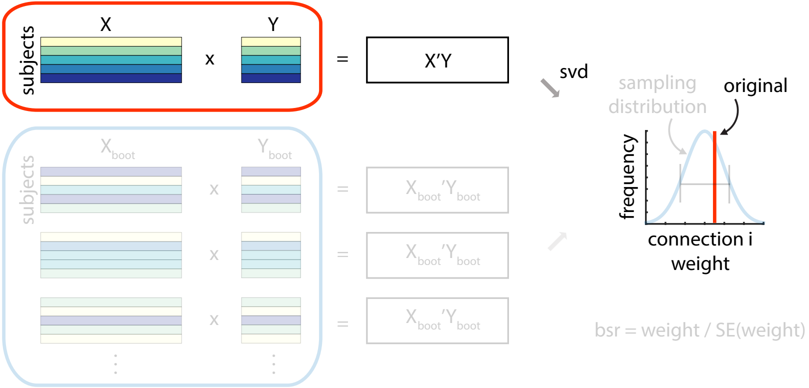
Which variables are important?



Shared variance: loadings



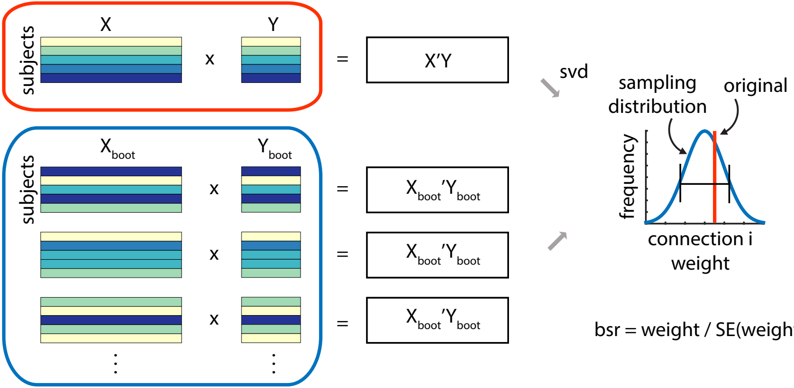
Reliability: bootstrapping



Efron & Tibshirani (1986) *Stat Sci*

McIntosh et al. (1996, 2004) *NeuroImage*

Reliability: bootstrapping

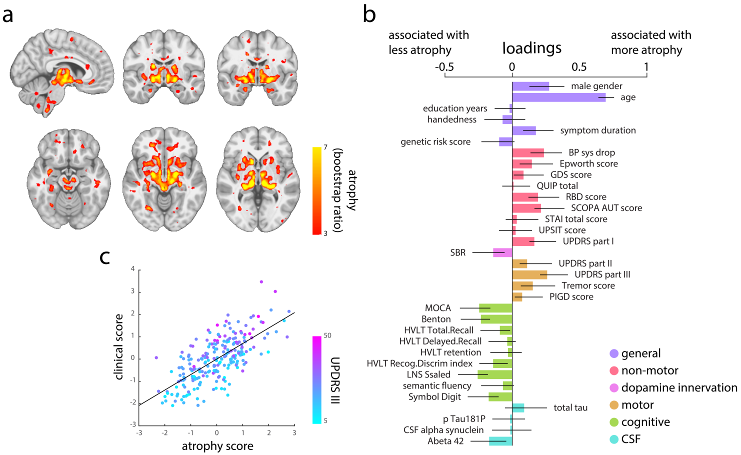


bsr = weight / SE(weight)

Efron & Tibshirani (1986) *Stat Sci*

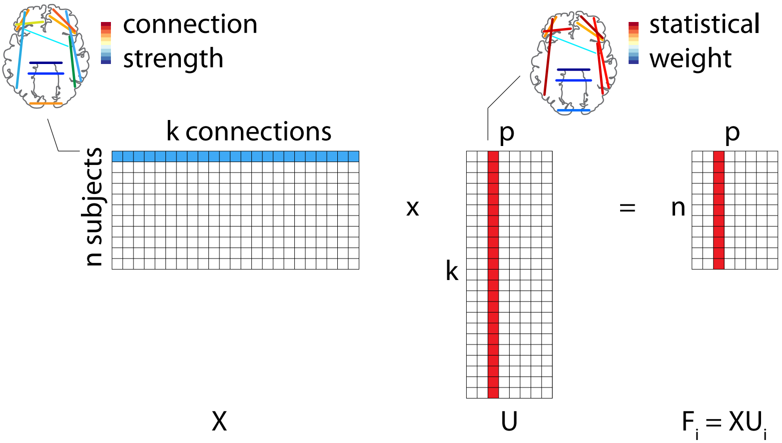
McIntosh et al. (1996, 2004) *NeuroImage*

Example: clinical-anatomical signature of Parkinson's

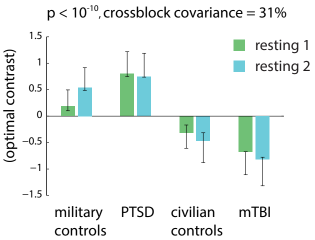


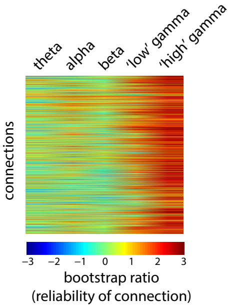
Zeighami et al. (2018) *NeuroImage*

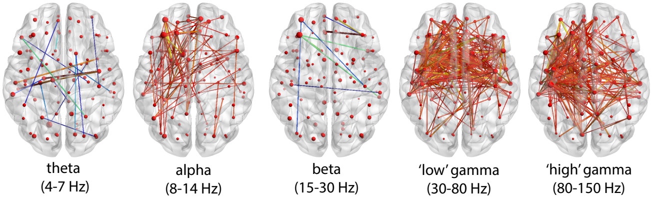
Individual participants



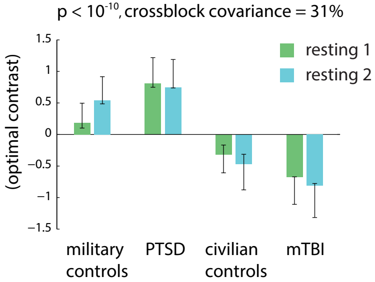
Example: connectivity differentiates PTSD from mTBI

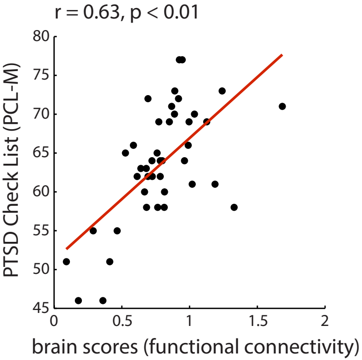


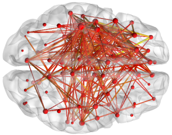




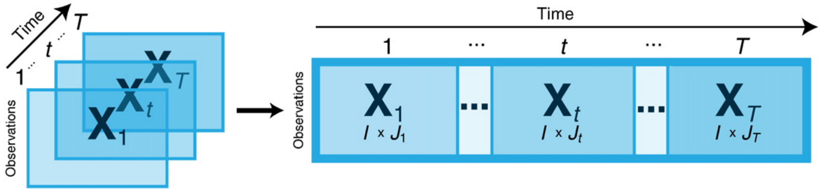
Example: connectivity and symptom severity in PTSD





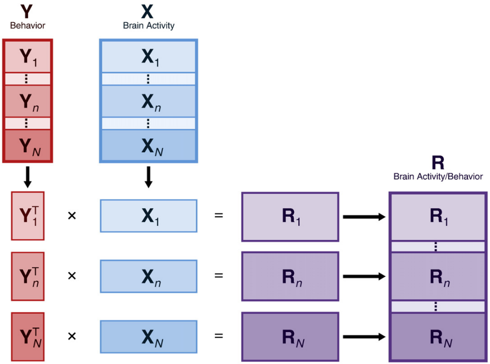


Adding other dimensions



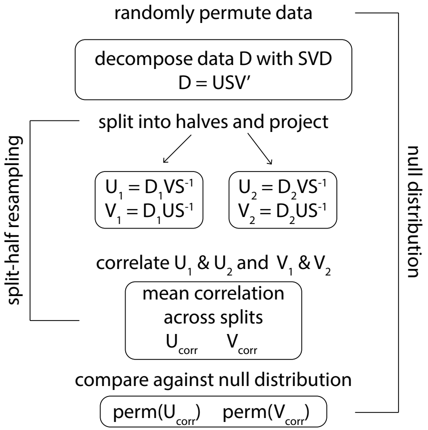
Krishnan et al. (2011) *NeuroImage*

Adding other dimensions



Krishnan et al. (2011) *NeuroImage*

Cross-validation: split-half resampling

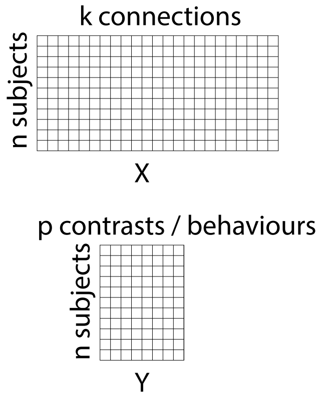


Strother et al. (2002) *NeuroImage*

Kovacevic et al. (2013) *New perspectives in Partial Least Squares* (Ed: Abdi)

Canonical correlation analysis (CCA)

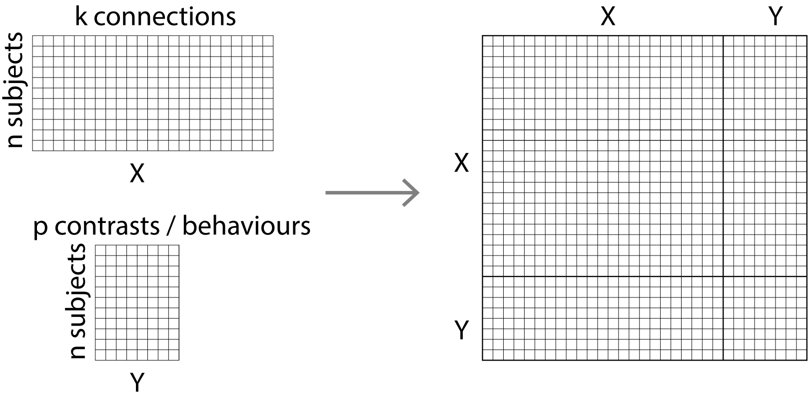
$$\mathrm{SVD}((\mathbf{X}'\mathbf{X}')^{-1/2}(\mathbf{X}'\mathbf{Y})(\mathbf{Y}'\mathbf{Y})^{-1/2})$$



Hotelling (1936) *Biometrika*

Canonical correlation analysis (CCA)

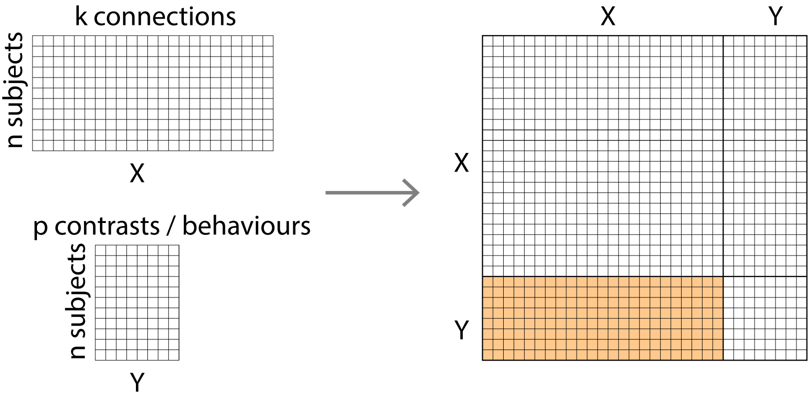
$$\mathrm{SVD}((\mathbf{X}'\mathbf{X}')^{-1/2}(\mathbf{X}'\mathbf{Y})(\mathbf{Y}'\mathbf{Y})^{-1/2})$$



Hotelling (1936) *Biometrika*

Canonical correlation analysis (CCA)

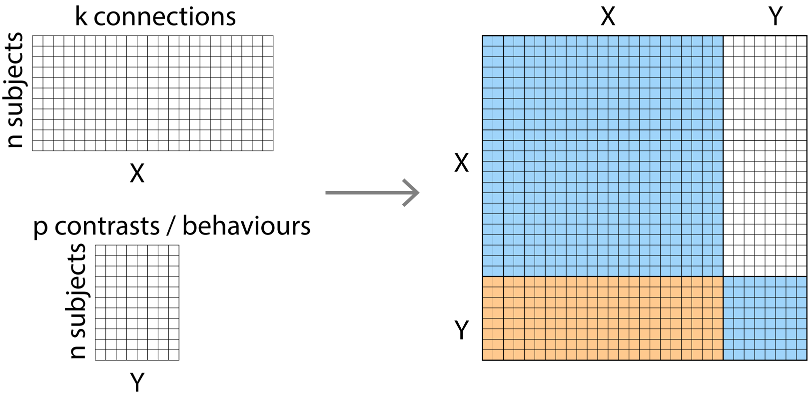
$$\mathrm{SVD}((\mathbf{X}'\mathbf{X}')^{-1/2}(\mathbf{X}'\mathbf{Y})(\mathbf{Y}'\mathbf{Y})^{-1/2})$$



Hotelling (1936) *Biometrika*

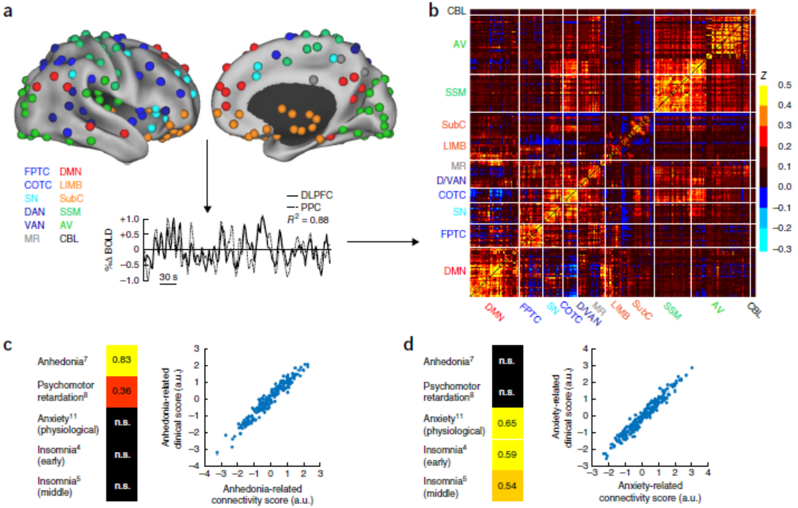
Canonical correlation analysis (CCA)

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



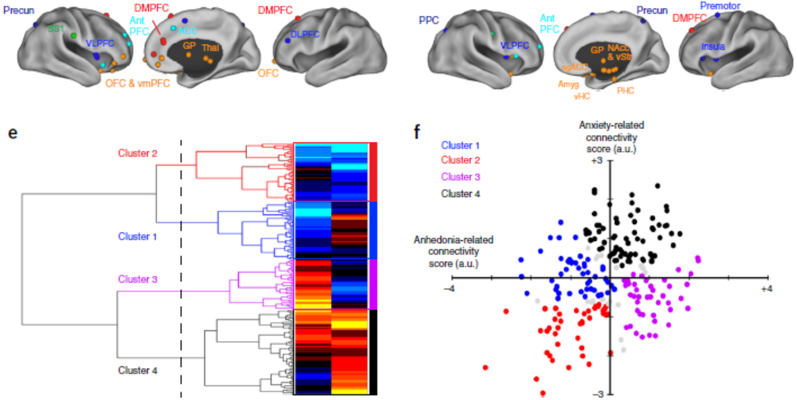
Hotelling (1936) *Biometrika*

Behaviour- and connectivity-defined subtypes of depression



Drysdale et al. (2016) *Nat Med*

Behaviour- and connectivity-defined subtypes of depression



Drysdale et al. (2016) *Nat Med*

Linear discriminant analysis (LDA)

- LDA: maximize between- vs. within-group variance
- SVD($\mathbf{W}^{-1}\mathbf{B}$)

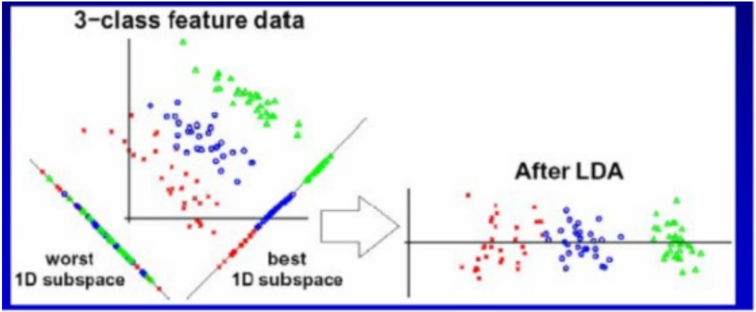


image: https://mlalgorithm.wordpress.com/

Friston et al. (1995) *NeuroImage*

Extensions

- sparse/regularized solutions, e.g.
  sCCA: Witten & Tibshirani (2009) *Stat Appl Genet Mol Biol*
  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

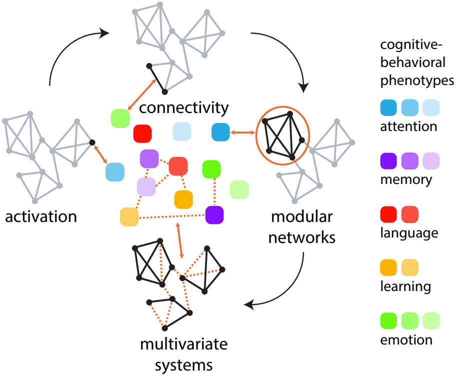
- overfitting
- linear
- unique partitioning of variance/covariance
- inference on individual variables

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/

Summary

- multivariate models embody network property
- all techniques entail unique assumptions
- many linear multivariate techniques are related
- multivariate techniques are versatile



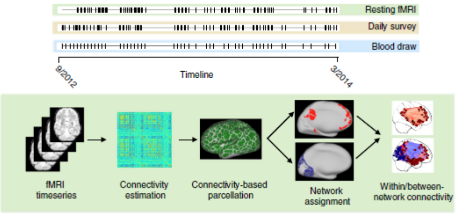
Demo: *MyConnectome*

ARTICLE

Received 9 Sep 2015 | Accepted 10 Oct 2015 | Published 9 Dec 2015 DOI: 10.1038/ncomms9885 OPEN

Long-term neural and physiological phenotyping of a single human

Russell A. Poldrack[1,2,3,4], Timothy O. Laumann[5], Oluwasanmi Koyejo[4], Brenda Gregory[3], Ashleigh Hover[3], Mei-Yen Chen[1], Krzysztof J. Gorgolewski[4], Jeffrey Luci[2,3], Sung Jun Joo[1], Ryan L. Boyd[1], Scott Hunicke-Smith[6], Zack Booth Simpson[7], Thomas Caven[8], Vanessa Sochat[9], James M. Shine[4], Evan Gordon[5], Abraham Z. Snyder[5], Babatunde Adeyemo[5], Steven E. Petersen[5], David C. Glahn[10,11], D. Reese Mckay[10,11], Joanne E. Curran[12], Harald H.H. Göring[13], Melanie A. Carless[13], John Blangero[12], Robert Dougherty[14], Alexander Leemans[15], Daniel A. Handwerker[16], Laurie Frick[3], Edward M. Marcotte[7,17] & Jeanette A. Mumford[1]



Poldrack et al. (2015) *Nat Commun*