Almost everything you need to know about PLS

Part 2: How to do PLS in Matlab and R

Jenny Rieck & Derek Beaton

November 21, 2017

The BIG outline

- Part 1: Background & Examples

The BIG outline

- Part 1: Background & Examples
  - It was on Oct. 24 (YOU MISSED IT).

The BIG outline

- Part 1: Background & Examples
  - It was on Oct. 24 (YOU MISSED IT).
  - It's ok, we'll do a refresher.

The BIG outline

- Part 1: Background & Examples
  - It was on Oct. 24 (YOU MISSED IT).
  - It's ok, we'll do a refresher.
- Part 2: PLS in Matlab & R

The BIG outline

- Part 1: Background & Examples
  - It was on Oct. 24 (YOU MISSED IT).
  - It's ok, we'll do a refresher.
- Part 2: PLS in Matlab & R
  - RIGHT NOW.

The BIG outline

- Part 1: Background & Examples
  - It was on Oct. 24 (YOU MISSED IT).
  - It's ok, we'll do a refresher.
- Part 2: PLS in Matlab & R
  - RIGHT NOW.
  - Put knowledge into practice

Part 2 outline

- Refresher

Part 2 outline

- Refresher
  - Principal component analysis (PCA)

Part 2 outline

- Refresher
  - Principal component analysis (PCA)
  - Partial least squares (PLS)

Part 2 outline

- Refresher
  - Principal component analysis (PCA)
  - Partial least squares (PLS)
    - Refresher

Part 2 outline

- Refresher
    - Principal component analysis (PCA)
    - Partial least squares (PLS)
        - Refresher
        - Example

Part 2 outline

- Refresher
  - Principal component analysis (PCA)
  - Partial least squares (PLS)
    - Refresher
    - Example
- Baby example

Part 2 outline

- Refresher
    - Principal component analysis (PCA)
    - Partial least squares (PLS)
        - Refresher
        - Example
- Baby example
    - Single subject block design (fMRI)

Part 2 outline

- Refresher
  - Principal component analysis (PCA)
  - Partial least squares (PLS)
    - Refresher
    - Example
- Baby example
  - Single subject block design (fMRI)
  - First with PLSGui (Matlab)

Part 2 outline

- Refresher
    - Principal component analysis (PCA)
    - Partial least squares (PLS)
        - Refresher
        - Example
- Baby example
    - Single subject block design (fMRI)
    - First with PLSGui (Matlab)
    - Then in R

Part 2 outline

- Refresher
  - Principal component analysis (PCA)
  - Partial least squares (PLS)
    - Refresher
    - Example
- Baby example
  - Single subject block design (fMRI)
  - First with PLSGui (Matlab)
  - Then in R
- Real examples

Part 2 outline

- Refresher
  - Principal component analysis (PCA)
  - Partial least squares (PLS)
    - Refresher
    - Example
- Baby example
  - Single subject block design (fMRI)
  - First with PLSGui (Matlab)
  - Then in R
- Real examples
  - Face matching in aging

Part 2 outline

- Refresher
  - Principal component analysis (PCA)
  - Partial least squares (PLS)
    - Refresher
    - Example
- Baby example
  - Single subject block design (fMRI)
  - First with PLSGui (Matlab)
  - Then in R
- Real examples
  - Face matching in aging
  - Incidental encoding, memory peformance in aging

Part 2 outline

- Refresher
  - Principal component analysis (PCA)
  - Partial least squares (PLS)
    - Refresher
    - Example
- Baby example
  - Single subject block design (fMRI)
  - First with PLSGui (Matlab)
  - Then in R
- Real examples
  - Face matching in aging
  - Incidental encoding, memory peformance in aging
  - Reward processing, genetics, & stress in CUD

Part 2 outline

- Refresher
  - Principal component analysis (PCA)
  - Partial least squares (PLS)
    - Refresher
    - Example
- Baby example
  - Single subject block design (fMRI)
  - First with PLSGui (Matlab)
  - Then in R
- Real examples
  - Face matching in aging
  - Incidental encoding, memory peformance in aging
  - Reward processing, genetics, & stress in CUD
- And beyond!

Principal Components Analysis

What is PCA?

- Visualize high dimensional data

What is PCA?

- Visualize high dimensional data
- Orthogonal transformation

What is PCA?

- Visualize high dimensional data
- Orthogonal transformation
- Dimensionality reduction

What is PCA?

- Visualize high dimensional data
- Orthogonal transformation
- Dimensionality reduction
- Modern overviews

What is PCA?

- Visualize high dimensional data
- Orthogonal transformation
- Dimensionality reduction
- Modern overviews
    - S. Wold et al., (1987)

What is PCA?

- Visualize high dimensional data
- Orthogonal transformation
- Dimensionality reduction
- Modern overviews
    - S. Wold et al., (1987)
    - Jollife (2002)

What is PCA?

- Visualize high dimensional data
- Orthogonal transformation
- Dimensionality reduction
- Modern overviews
  - S. Wold et al., (1987)
  - Jollife (2002)
  - Abdi & Williams (2010)

What is PCA?

- Find "components"

What is PCA?

- Find "components"
    - Components are new variables that are combinations of old variables

What is PCA?

- Find "components"
  - Components are new variables that are combinations of old variables
- Components explain maximum possible variance

What is PCA?

- Find “components”
  - Components are new variables that are combinations of old variables
- Components explain maximum possible variance
  - Conditional to orthogonality

Visual example

Visual example

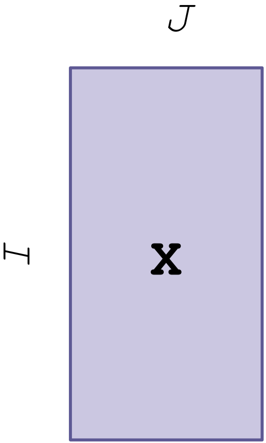
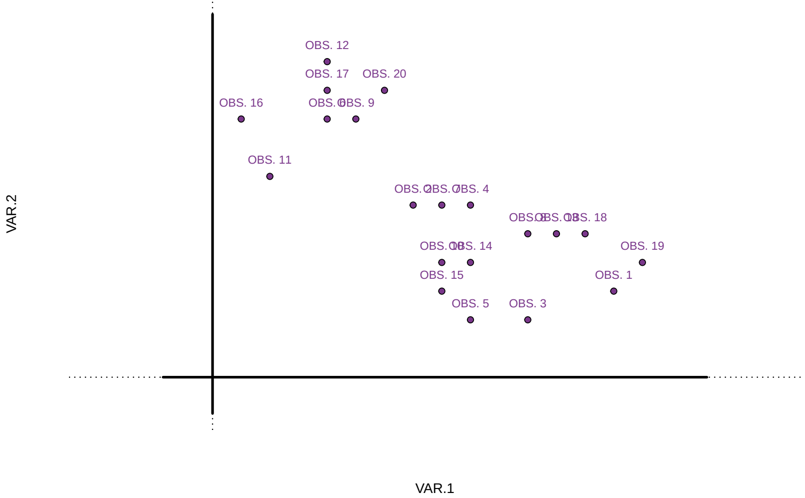
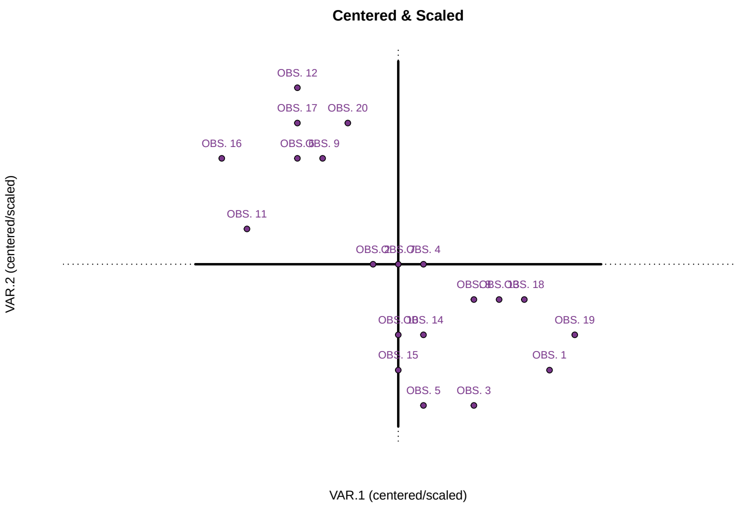


Figure 1: The kind of data we usually expect for PCA

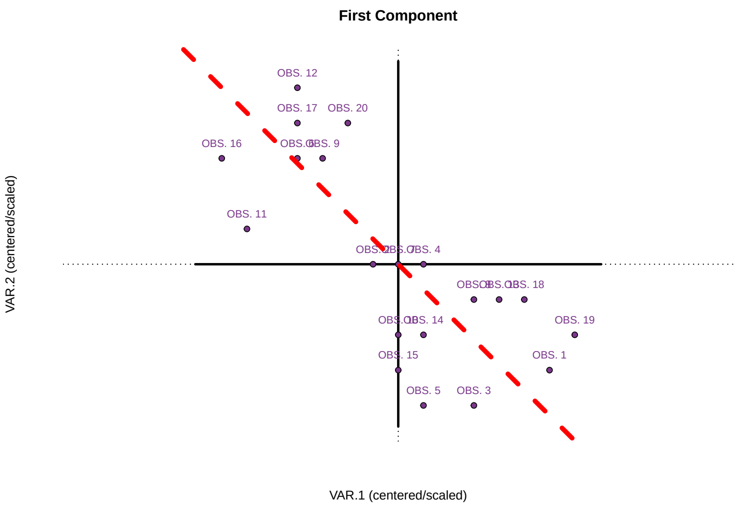
Visual example



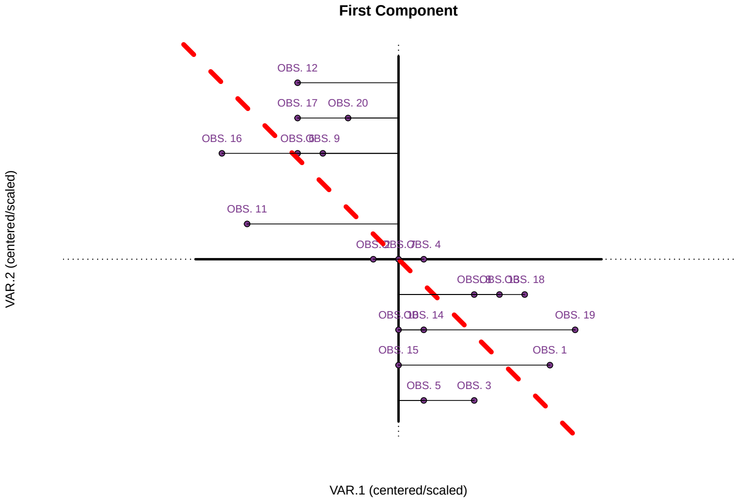
Visual example



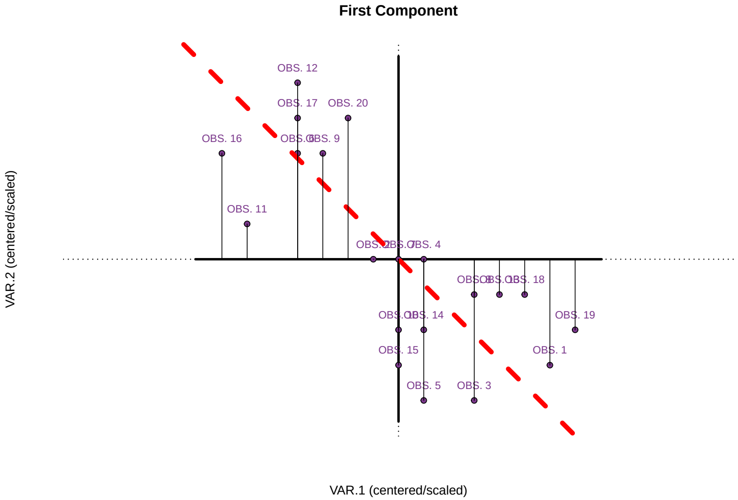
Visual example



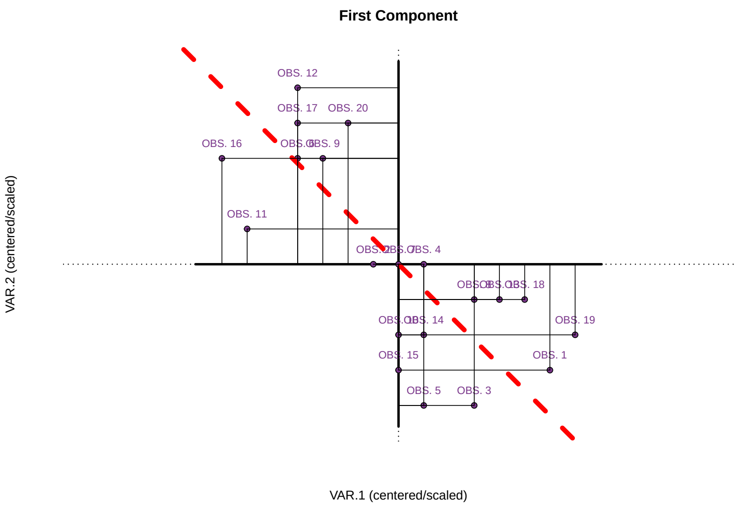
Visual example



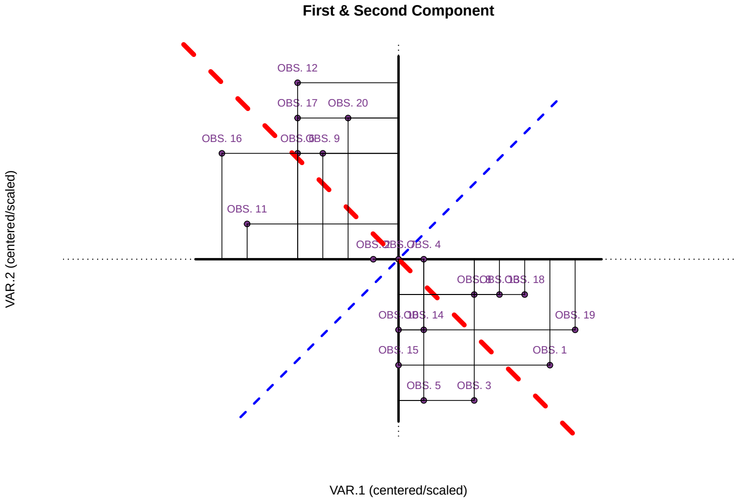
Visual example



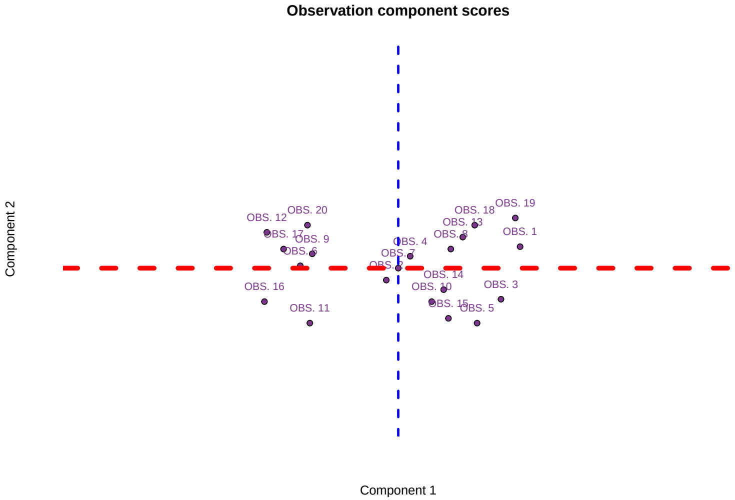
Visual example



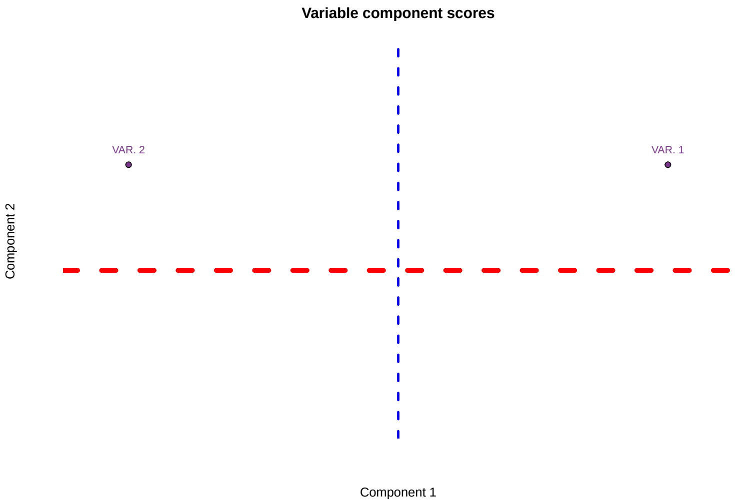
Visual example



Visual example



Visual example



Background

What is PCA?

- The basis of modern techniques

What is PCA?

- The basis of modern techniques
  - Factor analyses

What is PCA?

- The basis of modern techniques
  - Factor analyses
  - Independent components analysis

What is PCA?

- The basis of modern techniques
    - Factor analyses
    - Independent components analysis
    - Partial least squares

What is PCA?

- The basis of modern techniques
    - Factor analyses
    - Independent components analysis
    - Partial least squares
    - Discriminant analyses

What is PCA?

- The basis of modern techniques
  - Factor analyses
  - Independent components analysis
  - Partial least squares
  - Discriminant analyses
  - Multi-table (e.g., MFA, GCCA)

What is PCA?

- A special case of the singular value decomposition (SVD)

What is PCA?

- A special case of the singular value decomposition (SVD)
- Which means (almost) everything else is, too

Singular value decomposition

- The SVD is one of the most ubiquitous and important tools

Singular value decomposition

- The SVD is one of the most ubiquitous and important tools
- If you want more:

Singular value decomposition

- The SVD is one of the most ubiquitous and important tools
- If you want more:
  - S. Wold et al., (1987)

Singular value decomposition

- The SVD is one of the most ubiquitous and important tools
- If you want more:
    - S. Wold et al., (1987)
    - Jollife (2002)

Singular value decomposition

- The SVD is one of the most ubiquitous and important tools
- If you want more:
    - S. Wold et al., (1987)
    - Jollife (2002)
    - Abdi & Williams (2010)

Singular value decomposition

- The SVD is one of the most ubiquitous and important tools
- If you want more:
  - S. Wold et al., (1987)
  - Jollife (2002)
  - Abdi & Williams (2010)
  - And many others...

Formalization

SVD

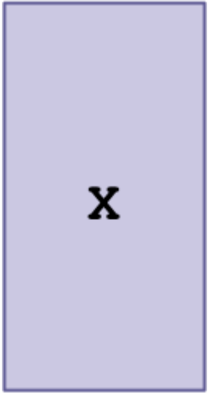


Figure 2: The shape of the data

SVD

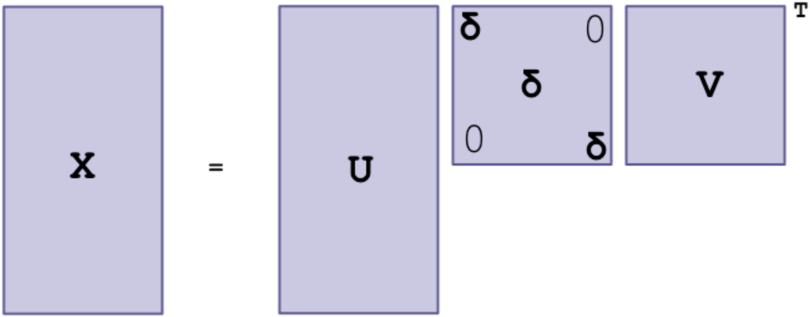


Figure 3: SVD breaks down the data

SVD

Notation

SVD

Notation

- x - a scalar

SVD

Notation

- x - a scalar
- $\mathbf{a}$ - a vector

SVD

Notation

- x - a scalar
- $\mathbf{a}$ - a vector
- $\mathbf{A}$ - a matrix

SVD

Notation

- x - a scalar
- $\mathbf{a}$ - a vector
- $\mathbf{A}$ - a matrix
- $\mathbf{A}^T$ - transpose

SVD

Notation

- x - a scalar
- $\mathbf{a}$ - a vector
- $\mathbf{A}$ - a matrix
- $\mathbf{A}^T$ - transpose
- $\mathbf{AB}$ - multiplication

SVD

Given a matrix $\mathbf{X}$ we generally assume that

SVD

Given a matrix $\mathbf{X}$ we generally assume that

- column-wise centered

SVD

Given a matrix **X** we generally assume that

- column-wise centered
- column-wise scaled (e.g., z-scores or sums of squares $= 1$)

SVD

The SVD of $\mathbf{X}$ of size $I \times J$:

SVD

The SVD of $\mathbf{X}$ of size $I \times J$:

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^{\mathbf{T}} \quad (1)$$

SVD

The SVD of $\mathbf{X}$ of size $I \times J$:

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^{\mathbf{T}} \tag{1}$$

$\mathbf{U}$ and $\mathbf{V}$ are orthonormal such that

SVD

The SVD of $\mathbf{X}$ of size $I \times J$:

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^{\mathbf{T}} \tag{1}$$

$\mathbf{U}$ and $\mathbf{V}$ are orthonormal such that

$$\mathbf{U}^{\mathbf{T}}\mathbf{U} = \mathbf{I} = \mathbf{V}^{\mathbf{T}}\mathbf{V} \tag{2}$$

SVD

$\mathbf{X} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^{\mathsf{T}}$ such that $\mathbf{U}^{\mathsf{T}}\mathbf{U} = \mathbf{I} = \mathbf{V}^{\mathsf{T}}\mathbf{V}$

SVD

$\mathbf{X} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^{\mathbf{T}}$ such that $\mathbf{U}^{\mathbf{T}}\mathbf{U} = \mathbf{I} = \mathbf{V}^{\mathbf{T}}\mathbf{V}$

- where rank is L (i.e., number of extractable components)

SVD

$\mathbf{X} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^\mathsf{T}$ such that $\mathbf{U}^\mathsf{T}\mathbf{U} = \mathbf{I} = \mathbf{V}^\mathsf{T}\mathbf{V}$

- where rank is L (i.e., number of extractable components)
- $\mathbf{U}$ is $I \times L$ (left singular vectors; rows of $\mathbf{X}$)

SVD

$\mathbf{X} = \mathbf{U\Delta V^T}$ such that $\mathbf{U^TU} = \mathbf{I} = \mathbf{V^TV}$

- where rank is L (i.e., number of extractable components)
- $\mathbf{U}$ is $I \times L$ (left singular vectors; rows of $\mathbf{X}$)
- $\mathbf{V}$ is $J \times L$ (right singular vectors; columns of $\mathbf{X}$)

SVD

$$\mathbf{X} = \mathbf{U\Delta V^T} \text{ such that } \mathbf{U^TU} = \mathbf{I} = \mathbf{V^TV}$$

SVD

$\mathbf{X} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^{\mathsf{T}}$ such that $\mathbf{U}^{\mathsf{T}}\mathbf{U} = \mathbf{I} = \mathbf{V}^{\mathsf{T}}\mathbf{V}$

- $\boldsymbol{\Delta}$ is $L \times L$ diagonal matrix

SVD

$\mathbf{X} = \mathbf{U\Delta V^T}$ such that $\mathbf{U^TU} = \mathbf{I} = \mathbf{V^TV}$

- $\mathbf{\Delta}$ is $L \times L$ diagonal matrix
- $\mathrm{diag}\{\mathbf{\Delta}\} = \boldsymbol{\delta}$ are singular values (decreasing)

SVD

$\mathbf{X} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^{\mathbf{T}}$ such that $\mathbf{U}^{\mathbf{T}}\mathbf{U} = \mathbf{I} = \mathbf{V}^{\mathbf{T}}\mathbf{V}$

- $\boldsymbol{\Delta}$ is $L \times L$ diagonal matrix
- $\mathrm{diag}\{\boldsymbol{\Delta}\} = \boldsymbol{\delta}$ are singular values (decreasing)
- $\boldsymbol{\lambda} = \boldsymbol{\delta}^{2}$ are the eigenvalues (variance)

SVD

$$\mathbf{X} = \mathbf{U}\mathbf{\Delta}\mathbf{V}^{\mathbf{T}} \text{ such that } \mathbf{U}^{\mathbf{T}}\mathbf{U} = \mathbf{I} = \mathbf{V}^{\mathbf{T}}\mathbf{V}$$

SVD

$\mathbf{X} = \mathbf{U}\mathbf{\Delta}\mathbf{V}^\mathbf{T}$ such that $\mathbf{U}^\mathbf{T}\mathbf{U} = \mathbf{I} = \mathbf{V}^\mathbf{T}\mathbf{V}$

- $\mathbf{F}_I = \mathbf{U}\mathbf{\Delta}$ (row component scores)

SVD

$\mathbf{X} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^\mathbf{T}$ such that $\mathbf{U}^\mathbf{T}\mathbf{U} = \mathbf{I} = \mathbf{V}^\mathbf{T}\mathbf{V}$

- $\mathbf{F}_I = \mathbf{U}\boldsymbol{\Delta}$ (row component scores)
- $\mathbf{F}_J = \mathbf{V}\boldsymbol{\Delta}$ (column component scores)

SVD

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^{\mathsf{T}} \text{ such that } \mathbf{U}^{\mathsf{T}}\mathbf{U} = \mathbf{I} = \mathbf{V}^{\mathsf{T}}\mathbf{V}$$

SVD

$\mathbf{X} = \mathbf{U \Delta V^T}$ such that $\mathbf{U^T U} = \mathbf{I} = \mathbf{V^T V}$

- $\mathbf{F}_I = \mathbf{U\Delta} = \mathbf{XV}$

SVD

$\mathbf{X} = \mathbf{U\Delta V^T}$ such that $\mathbf{U^TU} = \mathbf{I} = \mathbf{V^TV}$

- $\mathbf{F}_I = \mathbf{U\Delta} = \mathbf{XV}$
- $\mathbf{F}_J = \mathbf{V\Delta} = \mathbf{X}^T\mathbf{U}$

Partial Least Squares

Refresher

PLS means a lot of things

- Projection onto latent structures

PLS means a lot of things

- Projection onto latent structures
  - Probably the most accurate name

PLS means a lot of things

- Projection onto latent structures
  - Probably the most accurate name
  - But also probably too broad a definition

PLS means a lot of things

Partial least squares sounds like ordinary least squares

- When we have two matrices: **X** and **Y**

PLS means a lot of things

Partial least squares sounds like ordinary least squares

- When we have two matrices: $\mathbf{X}$ and $\mathbf{Y}$
- OLS: $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$

PLS means a lot of things

Partial least squares sounds like ordinary least squares

- When we have two matrices: $\mathbf{X}$ and $\mathbf{Y}$
- OLS: $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$
- PLS: $\mathbf{X}^T\mathbf{Y}$

PLS means a lot of things

- Partial least squares path modelling (PLS-PM)

PLS means a lot of things

- Partial least squares path modelling (PLS-PM)
- Partial least squares regression (PLSR)

PLS means a lot of things

- Partial least squares path modelling (PLS-PM)
- Partial least squares regression (PLSR)
- Partial least squares correlation (PLSC)

PLS means a lot of things

- Partial least squares path modelling (PLS-PM)
- Partial least squares regression (PLSR)
- Partial least squares correlation (PLSC)
  - This is the one we'll talk about today

PLSC has many. . .

Names

- Inter-battery (factor) analysis (Tucker, 1958)

PLSC has many...

Names

- Inter-battery (factor) analysis (Tucker, 1958)
- Covariance between two fields (Bretherton, Smith, & Wallace, 1992)

PLSC has many...

Names

- Inter-battery (factor) analysis (Tucker, 1958)
- Covariance between two fields (Bretherton, Smith, & Wallace, 1992)
- PLS-SVD (Tenenhaus, 2005)

PLSC has many. . .

Names

- Inter-battery (factor) analysis (Tucker, 1958)
- Covariance between two fields (Bretherton, Smith, & Wallace, 1992)
- PLS-SVD (Tenenhaus, 2005)
- Co-inertia analysis (Dray, 2014)

PLSC has many. . .

Friends

- Reduced Rank Regression (a.k.a. redundancy analysis)

PLSC has many. . .

Friends

- Reduced Rank Regression (a.k.a. redundancy analysis)
- Canonical Correlation Analysis

PLSC has many...

Friends

- Reduced Rank Regression (a.k.a. redundancy analysis)
- Canonical Correlation Analysis
- (Fisher's) Linear Discriminant Analysis

PLSC has many. . .

Friends

- Reduced Rank Regression (a.k.a. redundancy analysis)
- Canonical Correlation Analysis
- (Fisher's) Linear Discriminant Analysis
- PLS-correspondence analysis

PLSC

History

- McIntosh, Bookstein, Haxby, & Grady (1996)

PLSC

History

- McIntosh, Bookstein, Haxby, & Grady (1996)
- Bookstein (1992)

PLSC

History

- McIntosh, Bookstein, Haxby, & Grady (1996)
- Bookstein (1992)
- Tucker (1958)

PLSC

Modern overviews

- McIntosh & Lobaugh (2004)

PLSC

Modern overviews

- McIntosh & Lobaugh (2004)
- Krishnan et al., (2011)

PLSC

It's effectively just PCA applied to the cross product of two matrices measured on the same observations:

PLSC

It's effectively just PCA applied to the cross product of two matrices measured on the same observations:

- $\mathbf{X}$ which is $I \times J$

PLSC

It's effectively just PCA applied to the cross product of two matrices measured on the same observations:

- $\mathbf{X}$ which is $I \times J$
- $\mathbf{Y}$ which is $I \times K$

PLSC via the SVD

Compute the relationship between $\mathbf{X}$ and $\mathbf{Y}$

$$\mathbf{R} = \mathbf{X}^{\mathsf{T}}\mathbf{Y} \tag{3}$$

PLSC via the SVD

Compute the relationship between **X** and **Y**

$$\mathbf{R} = \mathbf{X}^{\mathbf{T}}\mathbf{Y} \tag{3}$$

Compute the SVD of **R**

$$\mathbf{R} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^{\mathbf{T}} \tag{4}$$

PLSC via the SVD

Compute the relationship between $\mathbf{X}$ and $\mathbf{Y}$

$$\mathbf{R} = \mathbf{X}^{\mathsf{T}}\mathbf{Y} \tag{3}$$

Compute the SVD of $\mathbf{R}$

$$\mathbf{R} = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^{\mathsf{T}} \tag{4}$$

Compute the latent variables

$$\mathbf{L}_{\mathbf{X}} = \mathbf{X}\mathbf{U} \text{ and } \mathbf{L}_{\mathbf{Y}} = \mathbf{Y}\mathbf{V} \tag{5}$$

PLSC via the SVD

Almost everything is the same:

- $\boldsymbol{\Delta}$ are singular values

PLSC via the SVD

Almost everything is the same:

- $\mathbf{\Delta}$ are singular values
- $\mathbf{F}_J = \mathbf{U}\mathbf{\Delta}$ (component scores for variables of $\mathbf{X}$)

PLSC via the SVD

Almost everything is the same:

- $\mathbf{\Delta}$ are singular values
- $\mathbf{F}_J = \mathbf{U\Delta}$ (component scores for variables of $\mathbf{X}$)
- $\mathbf{F}_K = \mathbf{V\Delta}$ (component scores for variables of $\mathbf{Y}$)

PLSC via the SVD

Almost everything is the same:

- $\mathbf{\Delta}$ are singular values
- $\mathbf{F}_J = \mathbf{U\Delta}$ (component scores for variables of $\mathbf{X}$)
- $\mathbf{F}_K = \mathbf{V\Delta}$ (component scores for variables of $\mathbf{Y}$)
- So our nomeclature will align with PCA

PLSC via the SVD

The new-ness

- $\mathbf{L_X} = \mathbf{XU}$ express the individuals w.r.t. $\mathbf{X}$

PLSC via the SVD

The new-ness

- $\mathbf{L_X} = \mathbf{XU}$ express the individuals w.r.t. $\mathbf{X}$
- $\mathbf{L_Y} = \mathbf{YV}$ express the individuals w.r.t. $\mathbf{Y}$

PLSC via the SVD

The new-ness

- $\mathbf{L_X} = \mathbf{XU}$ express the individuals w.r.t. $\mathbf{X}$
- $\mathbf{L_Y} = \mathbf{YV}$ express the individuals w.r.t. $\mathbf{Y}$
- Not in PCA

PLSC via the SVD

The new-ness

- $\mathbf{L_X} = \mathbf{XU}$ express the individuals w.r.t. $\mathbf{X}$
- $\mathbf{L_Y} = \mathbf{YV}$ express the individuals w.r.t. $\mathbf{Y}$
- Not in PCA
  - We'll call these "latent variable scores"

PLSC via the SVD

Maximizes the latent variables

PLSC via the SVD

Maximizes the latent variables

$$\mathbf{L}_{\mathbf{X}}^{T}\mathbf{L}_{\mathbf{Y}} = \boldsymbol{\Delta} \tag{6}$$

PLSC via the SVD

Maximizes the latent variables

$$\mathbf{L}_{\mathbf{X}}^{T}\mathbf{L}_{\mathbf{Y}} = \boldsymbol{\Delta} \tag{6}$$

When expanded

$(\mathbf{XU})^{\mathbf{T}}(\mathbf{YV}) = \boldsymbol{\Delta}$

PLSC via the SVD

Maximizes the latent variables

$$\mathbf{L}_{\mathbf{X}}^{T}\mathbf{L}_{\mathbf{Y}} = \boldsymbol{\Delta} \tag{6}$$

When expanded

$(\mathbf{XU})^{\mathbf{T}}(\mathbf{YV}) = \boldsymbol{\Delta}$

$\mathbf{U}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}}\mathbf{YV} = \boldsymbol{\Delta}$

PLSC via the SVD

Maximizes the latent variables

$$\mathbf{L}_{\mathbf{X}}^{T}\mathbf{L}_{\mathbf{Y}} = \boldsymbol{\Delta} \tag{6}$$

When expanded

$(\mathbf{XU})^{\mathbf{T}}(\mathbf{YV}) = \boldsymbol{\Delta}$

$\mathbf{U}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}}\mathbf{YV} = \boldsymbol{\Delta}$

$\mathbf{U}^{\mathbf{T}}\mathbf{RV} = \boldsymbol{\Delta}$

PLSC via the SVD

Maximizes the latent variables

$$\mathbf{L}_{\mathbf{X}}^{T}\mathbf{L}_{\mathbf{Y}} = \mathbf{\Delta} \tag{6}$$

When expanded

$(\mathbf{XU})^{\mathbf{T}}(\mathbf{YV}) = \mathbf{\Delta}$

$\mathbf{U}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}}\mathbf{YV} = \mathbf{\Delta}$

$\mathbf{U}^{\mathbf{T}}\mathbf{RV} = \mathbf{\Delta}$

$\mathbf{U}^{\mathbf{T}}\mathbf{U\Delta V}^{\mathbf{T}}\mathbf{V} = \mathbf{\Delta}$

PLSC via the SVD

Maximizes the latent variables

$$\mathbf{L}_{\mathbf{X}}^{T}\mathbf{L}_{\mathbf{Y}} = \mathbf{\Delta} \tag{6}$$

When expanded

$(\mathbf{XU})^{\mathbf{T}}(\mathbf{YV}) = \mathbf{\Delta}$

$\mathbf{U}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}}\mathbf{YV} = \mathbf{\Delta}$

$\mathbf{U}^{\mathbf{T}}\mathbf{RV} = \mathbf{\Delta}$

$\mathbf{U}^{\mathbf{T}}\mathbf{U\Delta V}^{\mathbf{T}}\mathbf{V} = \mathbf{\Delta}$

because

$\mathbf{U}^{\mathbf{T}}\mathbf{U} = \mathbf{I} = \mathbf{V}^{\mathbf{T}}\mathbf{V}$

PLSC via the SVD

It's effectively just PCA with some new-ness:

- $\mathbf{\Delta}$ are singular values

PLSC via the SVD

It's effectively just PCA with some new-ness:

- $\mathbf{\Delta}$ are singular values
- $\mathbf{U}$ akin to loadings (for variables of $\mathbf{X}$)

PLSC via the SVD

It's effectively just PCA with some new-ness:

- $\mathbf{\Delta}$ are singular values
- $\mathbf{U}$ akin to loadings (for variables of $\mathbf{X}$)
- $\mathbf{V}$ akin to loadings (for variables of $\mathbf{Y}$)

PLSC via the SVD

It's effectively just PCA with some new-ness:

- $\mathbf{\Delta}$ are singular values
- $\mathbf{U}$ akin to loadings (for variables of $\mathbf{X}$)
- $\mathbf{V}$ akin to loadings (for variables of $\mathbf{Y}$)
- $\mathbf{F}_J = \mathbf{U\Delta}$ (component scores for variables of $\mathbf{X}$)

PLSC via the SVD

It's effectively just PCA with some new-ness:

- $\mathbf{\Delta}$ are singular values
- $\mathbf{U}$ akin to loadings (for variables of $\mathbf{X}$)
- $\mathbf{V}$ akin to loadings (for variables of $\mathbf{Y}$)
- $\mathbf{F}_J = \mathbf{U\Delta}$ (component scores for variables of $\mathbf{X}$)
- $\mathbf{F}_K = \mathbf{V\Delta}$ (component scores for variables of $\mathbf{Y}$)

PLSC via the SVD

It's effectively just PCA with some new-ness:

- $\mathbf{\Delta}$ are singular values
- $\mathbf{U}$ akin to loadings (for variables of $\mathbf{X}$)
- $\mathbf{V}$ akin to loadings (for variables of $\mathbf{Y}$)
- $\mathbf{F}_J = \mathbf{U\Delta}$ (component scores for variables of $\mathbf{X}$)
- $\mathbf{F}_K = \mathbf{V\Delta}$ (component scores for variables of $\mathbf{Y}$)
- $\mathbf{L}_\mathbf{X} = \mathbf{XU}$ express the individuals w.r.t. $\mathbf{X}$

PLSC via the SVD

It's effectively just PCA with some new-ness:

- $\mathbf{\Delta}$ are singular values
- $\mathbf{U}$ akin to loadings (for variables of $\mathbf{X}$)
- $\mathbf{V}$ akin to loadings (for variables of $\mathbf{Y}$)
- $\mathbf{F}_J = \mathbf{U\Delta}$ (component scores for variables of $\mathbf{X}$)
- $\mathbf{F}_K = \mathbf{V\Delta}$ (component scores for variables of $\mathbf{Y}$)
- $\mathbf{L_X} = \mathbf{XU}$ express the individuals w.r.t. $\mathbf{X}$
- $\mathbf{L_Y} = \mathbf{YV}$ express the individuals w.r.t. $\mathbf{Y}$

A Glossary

For PCA nomenclature and PLSGui:

- $\mathbf{\Delta} = \mathbf{S}$ are singular values

A Glossary

For PCA nomenclature and PLSGui:

- $\mathbf{\Delta} = \mathbf{S}$ are singular values
- $\mathbf{U}$ saliences, brain latent variables (when you use brain data)

A Glossary

For PCA nomenclature and PLSGui:

- $\mathbf{\Delta} = \mathbf{S}$ are singular values
- $\mathbf{U}$ saliences, brain latent variables (when you use brain data)
- $\mathbf{V}$ saliences, design/behavior/etc... latent variables

A Glossary

For PCA nomenclature and PLSGui:

- $\mathbf{\Delta} = \mathbf{S}$ are singular values
- $\mathbf{U}$ saliences, brain latent variables (when you use brain data)
- $\mathbf{V}$ saliences, design/behavior/etc... latent variables
- $\mathbf{L_X} = \mathbf{XU}$ is brain scores (when you use brain data)

A Glossary

For PCA nomenclature and PLSGui:

- $\mathbf{\Delta} = \mathbf{S}$ are singular values
- $\mathbf{U}$ saliences, brain latent variables (when you use brain data)
- $\mathbf{V}$ saliences, design/behavior/etc... latent variables
- $\mathbf{L_X} = \mathbf{XU}$ is brain scores (when you use brain data)
- $\mathbf{L_Y} = \mathbf{YV}$ is design/behavior/etc... scores

Example

PLSC dataset

Via ADNI ($N = 569$)

- 3 groups of participants

PLSC dataset

Via ADNI ($N = 569$)

- 3 groups of participants
    - $N = 178$ healthy control

PLSC dataset

Via ADNI ($N = 569$)

- 3 groups of participants
  - $N = 178$ healthy control
  - $N = 275$ late MCI

PLSC dataset

Via ADNI ($N = 569$)

- 3 groups of participants
  - $N = 178$ healthy control
  - $N = 275$ late MCI
  - $N = 116$ AD

PLSC dataset

Via ADNI ($N = 569$)

- 3 groups of participants
  - $N = 178$ healthy control
  - $N = 275$ late MCI
  - $N = 116$ AD
- 8 neuropsych measures

PLSC dataset

Via ADNI ($N = 569$)

- 3 groups of participants
  - $N = 178$ healthy control
  - $N = 275$ late MCI
  - $N = 116$ AD
- 8 neuropsych measures
- 68 cortical thickness estimates (via Freesurfer)

Data matrices

$$\mathbf{X} = \begin{array}{cc} & \begin{array}{cccc} \text{BNT} & \text{Clock} & \ldots & \text{RAVLT} \end{array} \\ \begin{array}{c} \text{Subj}_1 \\ \text{Subj}_2 \\ \text{Subj}_3 \\ \vdots \\ \text{Subj}_{N\text{-}1} \\ \text{Subj}_N \end{array} & \begin{bmatrix} 10 & 5 & \ldots & 30 \\ 7 & 4 & \ldots & 26 \\ 3 & 0 & \ldots & 23 \\ \vdots & \vdots & & \vdots \\ 2 & 1 & \ldots & 18 \\ 8 & 4 & \ldots & 27 \end{bmatrix} \end{array} \qquad \mathbf{Y} = \begin{array}{cc} & \begin{array}{cccc} \text{R.IFG} & \text{L.IFG} & \ldots & \text{L.Fusi} \end{array} \\ \begin{array}{c} \text{Subj}_1 \\ \text{Subj}_2 \\ \text{Subj}_3 \\ \vdots \\ \text{Subj}_{N\text{-}1} \\ \text{Subj}_N \end{array} & \begin{bmatrix} 3.24 & 6.27 & \ldots & 2.32 \\ 5.89 & 0.26 & \ldots & 4.51 \\ 2.84 & 2.51 & \ldots & 1.17 \\ \vdots & \vdots & & \vdots \\ 1.96 & 8.9 & \ldots & 3.46 \\ 4.42 & 7.81 & \ldots & 1.96 \end{bmatrix} \end{array}$$

Figure 4: **X** and **Y** matrices in standard PLS

Standard PLSC scree

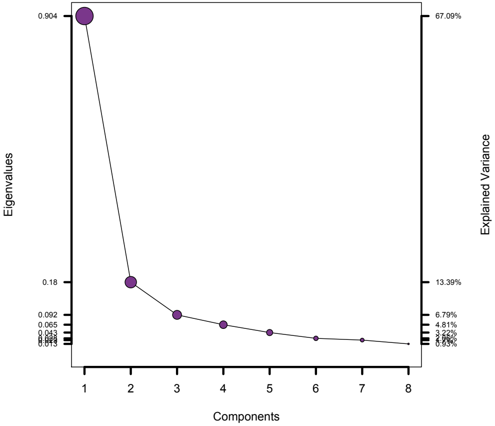


Figure 5:

How many components to interpret?

- A mix of art & science

How many components to interpret?

- A mix of art & science
- Use tests, effects sizes, and heuristics

How many components to interpret?

- A mix of art & science
- Use tests, effects sizes, and heuristics
- Inference tests (for later)

How many components to interpret?

- A mix of art & science
- Use tests, effects sizes, and heuristics
- Inference tests (for later)
    - Jackson (1993)

How many components to interpret?

- A mix of art & science
- Use tests, effects sizes, and heuristics
- Inference tests (for later)
  - Jackson (1993)
  - Peres-Neto et al., (2005)

How many components to interpret?

- A mix of art & science
- Use tests, effects sizes, and heuristics
- Inference tests (for later)
  - Jackson (1993)
  - Peres-Neto et al., (2005)
  - Dray (2008)

How many components to interpret?

- A mix of art & science
- Use tests, effects sizes, and heuristics
- Inference tests (for later)
    - Jackson (1993)
    - Peres-Neto et al., (2005)
    - Dray (2008)
    - Josse and Husson (2011)

How many components to interpret?

- A mix of art & science
- Use tests, effects sizes, and heuristics
- Inference tests (for later)
  - Jackson (1993)
  - Peres-Neto et al., (2005)
  - Dray (2008)
  - Josse and Husson (2011)
- We'll talk about the first 2

Two components

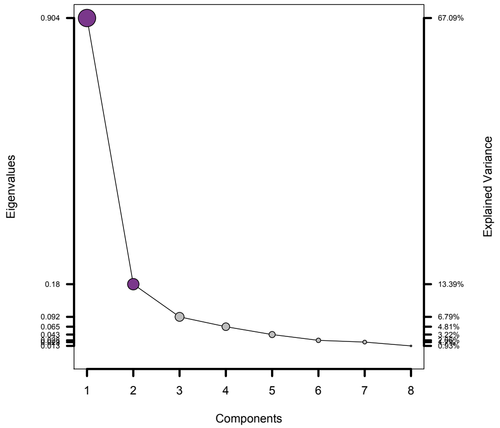


Figure 6:

Latent variables

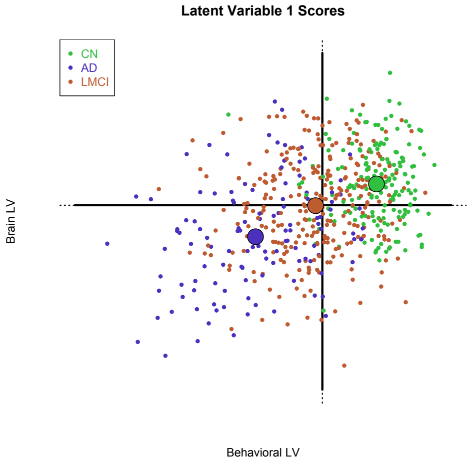


Figure 7:

Latent variables

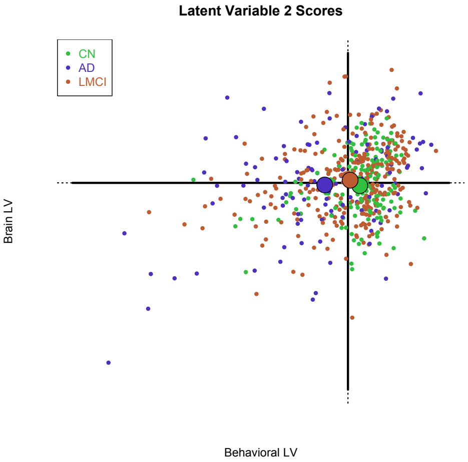


Figure 8:

Neuropsych component scores

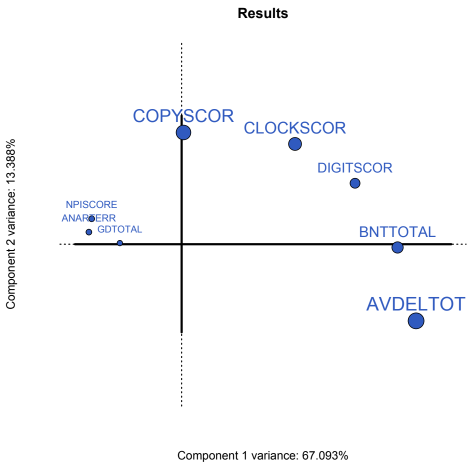


Figure 9:

Structural thickness component scores

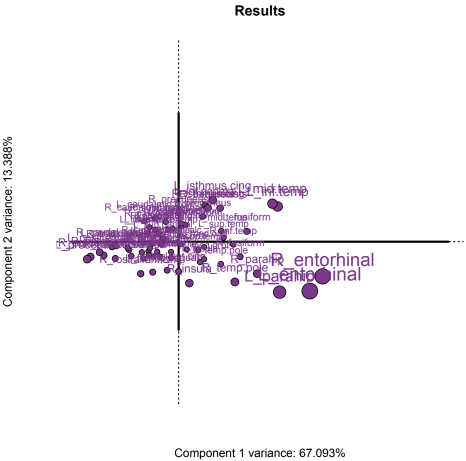


Figure 10:

LV1: Altogether now

- Let's focus on LV1

LV1: Altogether now

- Let's focus on LV1
- How can we put a story to the pictures?

LV1: Altogether now

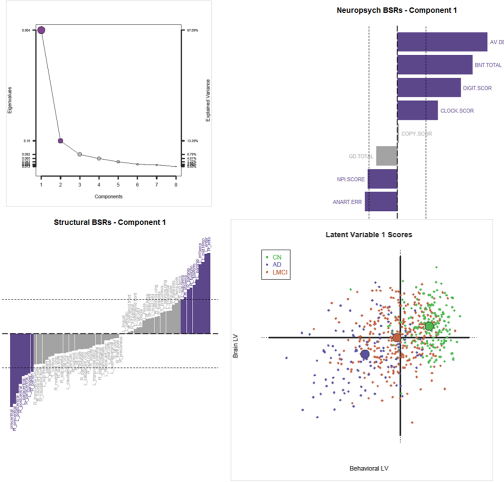


Figure 11: