Representational Component Models

Jörn Diedrichsen

Motor Control Group
Institute of Cognitive Neuroscience
University College London
Weighting of different features

\[ f_1, f_2, \ldots, f_k \]

Simple Model

\[ w_1 f_1, w_2 f_2, \ldots, w_k f_k \]

Weighted component model

Examples:
- Different aspects of objects: color, shape
- Different aspects of movements: force, sequence, timing
- Different layers of a computation vision model
Overview

• Features and representational components
• Factorial models (MANOVA)
• Linear representational models
• Nonlinear representational models
• Summary
Pattern for each condition is caused by different features, each associated with a feature pattern.
Pattern for each condition is caused by different features, each associated with a feature pattern.

Covariance matrix

\[ G = UU^T \]

\[ = \sum_{h=1}^{H} m_h w_h \sum_{h=1}^{H} w_h^T m_h^T \]

\[ = \sum_{h=1}^{H} m_h w_h w_h^T m_h^T \]

\[ = \sum_{h=1}^{H} w_h w_h^T m_h m_h^T \]

Assuming independence of feature patterns:

\[ w_i w_j^T = 0 \]

Distance matrix

\[ G = \sum_{h=1}^{H} \omega_h G_h \]

\[ D = \sum_{h=1}^{H} \omega_h D_h \]

The component weight is the variance or power of the feature pattern.
Representational components

- Often we want to assign a single weight for groups of features.
- Sometimes we think some features are not independent.
- Sometimes we don’t have individual features at all.

-> Representation components
Most often we do not weight single features, but groups of features: representation components.

\[
G = UU^T \\
= \sum_{h=1}^{H} M_h W_h \sum_{h=1}^{H} W_h^T M_h^T \\
= \sum_{h=1}^{H} M_h W_h W_h^T M_h^T \\
= \sum_{h=1}^{H} \omega_h M_h V_h M_h^T \\
= \sum_{h=1}^{H} \omega_h G_h
\]

Assuming independence of feature patterns across components:

\[
W_i W_j^T = 0
\]

Assuming covariance of feature patterns:

\[
W_h W_h^T = V_h \omega_h
\]

\[\omega_h\] The component weight is the variance or power of the feature pattern.
### Features and representational components

<table>
<thead>
<tr>
<th>Feature vectors</th>
<th>Feature correlation</th>
<th>Covariance component</th>
<th>Distance component</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_h$</td>
<td>$V_h$</td>
<td>$G_h = M_h V_h M_h^T$</td>
<td>$D_{ij} = G_{ii} + G_{jj} - 2G_{ij}$</td>
</tr>
</tbody>
</table>

- **Feature vectors**
  - $M_h$

- **Feature correlation**
  - $V_h$

- **Covariance component**
  - $G_h = M_h V_h M_h^T$

- **Distance component**
  - $D_{ij} = G_{ii} + G_{jj} - 2G_{ij}$

- **Not unique**
  - 

- **Unique**
  - 

**Diagrams**:

- Heatmaps and matrices representing $M_h$, $V_h$, $G_h$, and $D_{ij}$.
How do we estimate component weights?

\[ \mathbf{D} = \sum_{h=1}^{H} \omega_h \mathbf{D}_h \]

**Vectorise**

\[ \mathbf{D} \rightarrow \mathbf{d} \]

**Build component matrix**

\[ \mathbf{X} = \begin{bmatrix} \mathbf{d}_1 & \mathbf{d}_2 & \ldots \end{bmatrix} \]

**Ordinary least-squares**

\[ \hat{\omega} = \left( \mathbf{X}^T \mathbf{X} \right)^{-1} \mathbf{X}^T \hat{\mathbf{d}} \]
Features and representational components

- *Features* are variables encoded in neuronal elements
- Groups of features with similar encoding strength form a *representational component*
- Feature patterns of different components are assumed to be mutually independent
- Many feature sets can lead to the same representational component
- Models are uniquely specified via their component matrices (representational similarity trick)
- Component weights estimate variance (or power) of representations
Overview

• Features and representational components
• Factorial models (MANOVA)
• Linear representational models
• Nonlinear representational models
• Summary
Integrated vs. independent encoding

• Now, we assumed that different components are encoded independently in the brain
• This does not mean that they are encoded in different regions / voxel: only that their patterns are unrelated to each other

• **BUT:** Can we test this?
Integrated vs. independent encoding: factorial models

- Are two groups of features (variables) encoded independently or dependently?
- Vary the 2 factors in a fully crossed design
  - Condition (see / do) x Action (3 gestures)
  - Rhythm x Spatial sequence
  - Reach directions (3) x Grasps (3) ....
- Where is factor A encoded, where is B encoded?
- Are A and B encoded in an integrated or independent fashion?
- Is Factor A consistently encoded across levels of factor B ("cross-decoding")?
Factorial representational models (MANOVA)

Factor A

Factor B

Features

Representational components

Factor A

Factor B

Interaction

0 1

0 2

Component design matrix

Raw data

Regression coefficients

Pattern estimate

Dissimilarity matrix

First-level GLM

Searchlight Noise normalisation

Crossvalidation

\[ \tilde{\omega} = \left(X^T X\right)^{-1} X^T vec(\hat{D}) \]
Factorial representational models (MANOVA)

Factors

Factor A

Factor B

Features

Representational components

\((X^TX)^{-1}X^T\)

Cross “decoding” Pattern consistency
Allefeld et al. (2013)

\(\hat{\omega} = (X^TX)^{-1}X^T vec(\hat{D})\)
Factorial representational models (MANOVA)

Features

Factor A
- see
- do

Factor B
- grasp
- pinch
- grip

Interaction

Cross “decoding”
Pattern consistency
Allefeld et al. (2013)

Representational components

(\(X^TX\))^{-1} X^T

Component weights

\(\hat{\omega}\)

- \(\omega_A = 0.005\)
- \(\omega_B = 0\)
- \(\omega_I = 0.01\)
Factorial representational models (MANOVA)

- Factorial models can reveal mean encoding effect and interactions
- Component weight estimates are unbiased and can be directly tested in group analysis
- Main effects are assessed by pattern consistency across levels of the other variable (replaces cross-classification)
- Mathematically identical to approach suggested by Allefeld et al. (2013)
**Factorial MANOVA designs (example)**

<table>
<thead>
<tr>
<th>Temporal sequence</th>
<th>Spatial sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>S1</td>
</tr>
<tr>
<td></td>
<td>T1,O1</td>
</tr>
<tr>
<td>T2</td>
<td>S2</td>
</tr>
<tr>
<td></td>
<td>T2,O1</td>
</tr>
<tr>
<td>T3</td>
<td>S3</td>
</tr>
<tr>
<td></td>
<td>T3,O1</td>
</tr>
<tr>
<td></td>
<td>T1,O2</td>
</tr>
<tr>
<td></td>
<td>T2,O2</td>
</tr>
<tr>
<td></td>
<td>T3,O2</td>
</tr>
<tr>
<td></td>
<td>T1,O3</td>
</tr>
<tr>
<td></td>
<td>T2,O3</td>
</tr>
<tr>
<td></td>
<td>T3,O3</td>
</tr>
</tbody>
</table>

![Diagram](image)

1 trial

mini-block:
3 identical sequences in a row

Kornysheva et al. (2014). eLife.
Factorial MANOVA designs (example)

Overall

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td><img src="T1_S1.png" alt="Image" /></td>
<td><img src="T1_S2.png" alt="Image" /></td>
<td><img src="T1_S3.png" alt="Image" /></td>
</tr>
<tr>
<td>T2</td>
<td><img src="T2_S1.png" alt="Image" /></td>
<td><img src="T2_S2.png" alt="Image" /></td>
<td><img src="T2_S3.png" alt="Image" /></td>
</tr>
<tr>
<td>T3</td>
<td><img src="T3_S1.png" alt="Image" /></td>
<td><img src="T3_S2.png" alt="Image" /></td>
<td><img src="T3_S3.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Spatial

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td><img src="S1_T1.png" alt="Image" /></td>
<td><img src="S1_T2.png" alt="Image" /></td>
<td><img src="S1_T3.png" alt="Image" /></td>
</tr>
<tr>
<td>T2</td>
<td><img src="S2_T1.png" alt="Image" /></td>
<td><img src="S2_T2.png" alt="Image" /></td>
<td><img src="S2_T3.png" alt="Image" /></td>
</tr>
<tr>
<td>T3</td>
<td><img src="S3_T1.png" alt="Image" /></td>
<td><img src="S3_T2.png" alt="Image" /></td>
<td><img src="S3_T3.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Integrated

<table>
<thead>
<tr>
<th>Mean</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td><img src="Mean_S1.png" alt="Image" /></td>
<td><img src="Mean_S2.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Temporal

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td><img src="T1_S1.png" alt="Image" /></td>
<td><img src="T2_S1.png" alt="Image" /></td>
<td><img src="T3_S1.png" alt="Image" /></td>
</tr>
<tr>
<td>S2</td>
<td><img src="T1_S2.png" alt="Image" /></td>
<td><img src="T2_S2.png" alt="Image" /></td>
<td><img src="T3_S2.png" alt="Image" /></td>
</tr>
<tr>
<td>S3</td>
<td><img src="T1_S3.png" alt="Image" /></td>
<td><img src="T2_S3.png" alt="Image" /></td>
<td><img src="T3_S3.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Kornysheva et al. (2014). eLife.
Factorial MANOVA designs (example)

Ipsilateral

Contralateral

lateral  medial  medial  lateral

Overall encoding

Kornysheva et al. (2014). eLife.
Factorial MANOVA designs (example)

Kornysheva et al. (2014). eLife.
Factorial representational models (MANOVA)

**Linearity Assumption**

Patterns for different components overlap linearly

if they engage independent neuronal subpopulations

if they combine linearly to determine firing rate

*AND*

if the relationship between neural activity and BOLD is approximately linear

Experimental conditions should be similar in overall activity

*Note: mean value subtraction in analysis does not fix this!*
1. Pattern covariance matrices and squared Euclidean distance matrices capture the same information, but the former retain the baseline

2. A representational component (RC) is a group of representational features.

3. A representation can be modelled as weighted combination of RCs (one weight per RC).

4. Weighted combinations of RCs correspond to weighted combinations of representational distance matrices.

5. Component weights can be estimated using regression and tested directly (against zero) in group analyses.
Overview

• Features and representational components
• Factorial models (MANOVA)
• Linear representational models
• Nonlinear representational models
• Summary
sequence consisting of chunks

At what level are sequences represented?

Sequence

Chunk (2nd level)

Chunk

Press

2-nd order
From features to distances

**Feature**

- Fingertransitions
- Chunks
- 2nd-order chunks
- Sequences

**Pattern component**

- Covariance
- Distance
Mixture models of representations

sequence consisting of chunks

Sequence

S1: A 13 B 524 C 212 D 453
S2: D 453 C 212 B 524 A 13
S3: H 531 G 242 F 421 E 35

Sequence

Chunk (2nd level)

Chunk

Press

2-nd order
Mixture models of representations

\[ \omega_1 + \omega_2 + \omega_3 + \omega_4 = \]

The use of simple regression (OLS) assumes that distances:

- are independent
- have equal variance
- are \( \sim \) normally distributed

Is this a reasonable assumption?
Mixture models of representations

Distance 1-2

Distance 1-3

Distance 2-3

Mean
\begin{array}{ccc}
  d_{12} & d_{13} & d_{23} \\
\end{array}

Co-variance
\begin{array}{ccc}
  d_{12} & d_{13} & d_{23} \\
\end{array}
Mixture models of representations

Distance 1-2

Distance 1-3

Distance 2-3

Mean

Co-variance

$d_{12}$  $d_{13}$  $d_{23}$

$d_{12}$  $d_{13}$  $d_{23}$
Mixture models of representations

\[ \mathbf{V} = \frac{4}{R} (\mathbf{C}^T \mathbf{G} \mathbf{C} \circ \Sigma) + \frac{2}{R(R-1)} (\Sigma \circ \Sigma) \]

Signal dependent  Constant
Mixtures of representational models: IRLS estimation

1. Start with initial guess of $\eta$

2. Predict distances
   \[ \hat{d} = X\eta \]

3. Calculate variance-covariance of $d$
   \[
   G = -\frac{1}{2}HDH \\
   V = \frac{4}{R}(C^TGC \circ \Sigma) + \frac{2}{R(R-1)}(\Sigma \circ \Sigma)
   \]

4. Use in estimation
   \[ \eta = \left(X^T V^{-1} X\right)^{-1} X^T V^{-1} d \]

Until convergence

Model

Data
Mixtures of representational models: IRLS estimation

Sequence

Large chunk

Chunk

2-Finger

Sequence

Large chunk

Chunk

2-Finger

Improvements in SD

0.7%

7%

3%

16%
Linear representational models

Model comparison in simulated data
EMG vs. Natural statistics model

Correct decisions (%) vs. omega (% of noise)

- Likelihood ratio
- Cosine Angle
- Pearson
- Spearman
Overview

• Features and representational components
• Factorial models (MANOVA)
• Linear representational models
• Nonlinear representational models
• Summary
Non-linear representational models

Models in which component matrices are non-linear functions of parameters

<table>
<thead>
<tr>
<th>Tuning width</th>
<th>Tuning curves</th>
<th>Features</th>
<th>Distances</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma = 1 )</td>
<td><img src="image1.png" alt="Activity" /></td>
<td><img src="image2.png" alt="Target" /></td>
<td><img src="image3.png" alt="Distance1" /></td>
</tr>
<tr>
<td>( \sigma = 5 )</td>
<td><img src="image4.png" alt="Activity" /></td>
<td><img src="image5.png" alt="Target" /></td>
<td><img src="image6.png" alt="Distance2" /></td>
</tr>
<tr>
<td>( \sigma = 15 )</td>
<td><img src="image7.png" alt="Activity" /></td>
<td><img src="image8.png" alt="Target" /></td>
<td><img src="image9.png" alt="Distance3" /></td>
</tr>
</tbody>
</table>
Non-linear representational models

• Sometimes good linear approximations can be found (Example: AR-estimation in first-level SPMs)
• Otherwise, estimate nonlinear parameters to optimize the log-likelihood:

\[ \log p(\hat{d} | \theta) \propto -\frac{1}{2} (\hat{d} - d)^T V(d)^{-1} (\hat{d} - d) \]
Overview

• Features and representational components
• Factorial models (MANOVA)
• Linear representational models
• Nonlinear representational models
• Summary
## Representational analyses

<table>
<thead>
<tr>
<th><strong>Traditional “rank-based” RSA</strong></th>
<th><strong>Component model (OLS)</strong></th>
<th><strong>Component model (IRLS)</strong></th>
<th><strong>Component model (Bayesian)</strong></th>
<th><strong>Pattern component model</strong></th>
<th><strong>PRF-style model</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evaluation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spearman’s r</td>
<td>Linear fit</td>
<td>Likelihood</td>
<td>Marginal Likelihood</td>
<td>Marginal Likelihood</td>
<td>Crossvalidation</td>
</tr>
<tr>
<td>Kendall’s tau</td>
<td>$R^2$</td>
<td>$p(d \mid \omega, X)$</td>
<td>$p(d \mid X, \theta)$</td>
<td>$p(U \mid X, \theta)$</td>
<td></td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$d \sim N(X\omega, I\sigma^2)$</td>
<td>$d \sim N(X\omega, V(\hat{d}, \Sigma))$</td>
<td>$\omega_i \sim N(0, \theta_i)$</td>
<td>$\log(h_c) \sim N(0, \theta_i)$</td>
<td>$w \sim N(0, I\lambda^{-1})$</td>
</tr>
<tr>
<td><strong>Summary stats</strong></td>
<td>Cross-validated distance</td>
<td>LDC ($d$)</td>
<td>LDC ($d$)</td>
<td>covariance matrix ($G$)</td>
<td></td>
</tr>
<tr>
<td>Dissimilarity</td>
<td>LDC ($d$)</td>
<td>Noise estimate ($\Sigma$)</td>
<td>Noise estimate ($\Sigma$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cross-validated distance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pattern</strong></td>
<td><strong>U</strong></td>
<td><strong>U</strong></td>
<td><strong>U</strong></td>
<td><strong>U</strong></td>
<td><strong>U</strong></td>
</tr>
</tbody>
</table>
Representational component models

- Representational component models assume
  - independence of data across partitions
    -> (zero-distance is meaningful)
  - independence of feature patterns across components
  - linear overlap of patterns (within small range of variations)

- Representational component models do **NOT** assume
  - normality of the data
  - independence of distance estimates
  - linear relationship between psychological variables and BOLD
1. Large (squared Euclidean) distances are estimated with larger variability than smaller distances.
2. Distance estimates are statistically dependent in a way that is determined by the true distance structure.
3. Component weights can be estimated using iteratively reweighted least squares (IRLS), which yields better estimates than ordinary least squares (OLS) in some cases.
Thanks!
Representational component models

**Representational component model**
- Intercept is not included in fitting
- Needs to be explicitly modeled
- Predictions on ratio-scale
- Flexible factorial and combined models
- Linearity assumption (narrow range)
- Non-linearity can be modeled

**Rank-based RSA**
- Intercept is implicitly removed
- Does not contribute to model comparison
- Predictions on ordinal scale
- Single models
- Non-linearity of distances removed
The whole process

First-level GLM → Searchlight Prewhitening → Pattern estimate

Y → B → U

Distance / covariance estimation → Representational model fit → Group analysis

\( \hat{\mathbf{D}} \) → \( \hat{\mathbf{G}} \) → Component weights

Group map

\( \hat{\omega} \)

Covariance matrix

\( \hat{\omega} \)

\( \hat{\omega} \)

Group analysis

- Model coefficients can be directly tested (unbiased)
- Sometimes it is more sensible to use \( \sqrt{\hat{\omega}} \) (SD vs. variance)
Linear representational models: variance of distances

\[ \hat{d}_{ij} = u_i - u_j \quad \hat{d}_{ij} = \delta_{ij} + \varepsilon \]

Differences between patterns are measured with noise

\[ \hat{d} = \delta^{(m)} \delta^{(n)\top} = \left( \delta + \varepsilon^{(m)} \right) \left( \delta + \varepsilon^{(n)} \right)^\top \]

Squared distances are a sum of inner products, Signal with signal, signal with noise, and noise with noise

\[ \hat{d} = \delta \delta^\top + \varepsilon^{(m)} \delta^\top + \delta \varepsilon^{(n)\top} + \varepsilon^{(m)} \varepsilon^{(n)\top} \]

In the expected value, the inner products containing noise drop out

\[ E(\hat{d}) = \delta \delta^\top + \varepsilon^{(m)} \delta^\top + \delta \varepsilon^{(n)\top} + \varepsilon^{(m)} \varepsilon^{(n)\top} \]

For the variance, we obtain a part that depends on the distances, and one part that only depends on the noise.

\[ \text{var}(\hat{d}) = \text{var}(\delta \delta^\top) + \sigma^2 \varepsilon \delta \delta^\top + \sigma^2 \delta \delta^\top + \sigma^2 \varepsilon \sigma^2 \]

The covariance of \( p \) distances when doing exhaustive crossvalidation over \( R \) partitions:

\[ \text{var}(\hat{d}) = \frac{4}{R} \Delta \circ \Sigma + \frac{2P}{R(R-1)} \Sigma \circ \Sigma \]

Distance dependent \hspace{1cm} Constant

\[ \Sigma \quad \text{Within-run covariances of} \quad \hat{\delta} \]

\[ \Delta \quad \text{True inner products} \quad \langle \delta, \delta \rangle \]

\[ \circ \quad \text{Element-by-element multiplication} \]