Projection methods as dimensionality reduction tool and application in neuroscience

Nicolas Levernier
I) Experimental data and analysis with dimension reduction tools

II) Theoretical analysis of this dimension reduction tool
I) Experimental data and analysis with dimension reduction tools

II) Theoretical analysis of this dimension reduction tool
Setup of J. Epsztein and J. Koenig

[Bourboulou et al. eLife 2019]
Investigating spatial coding in the hippocampus

Setup of J. Epsztein and J. Koenig
[Bourboulou et al. eLife 2019]
Obtained data:
N time series of activation of hippocampal neurons

Setup of J. Epsztein and J. Koenig
[Bourboulou et al. eLife 2019]
Investigating spatial coding in the hippocampus

Setup of J. Epsztein and J. Koenig
[Bourboulou et al. eLife 2019]

Obtained data:
N time series of activation of hippocampal neurons

After binning and preprocessing:
matrix of firing rates, N * T
Obtained data:
N time series of activation of hippocampal neurons

After binning and preprocessing:
matrix of firing rates, N * T

Investigating spatial coding in the hippocampus

Setup of J. Epsztein and J. Koenig
[Bourboulou et al. eLife 2019]

Does the activity state code for mouse’s position in the corridor?
Two possible routes for analysis

I) Single cell analysis, averaging over laps

![Graph showing firing rate over maze position](image)
Two possible routes for analysis

I) Single cell analysis, averaging over laps

« Place cell »
Two possible routes for analysis

I) Single cell analysis, averaging over laps

T (time steps)

N (neurons)

Firing rate

N rate maps

Maze Position (cm)

« Place cell »
Two possible routes for analysis

I) Single cell analysis, averaging over laps

- **Advantages:**
  - can detect the most informative cells

- **Cons:**
  - assume a coding by place fields
  - lose of a lot of information
  - sensible to experimental issues

**Firing rate**

**N rate maps**

« Place cell »
II) Population analysis, no averaging

T (time steps) → T points in a N-dimensional space

N (neurons) → One data point
II) Population analysis, no averaging

How are this points distributed? Can this distribution be matched on position in the maze?
II) Population analysis, no averaging

- T (time steps)
- N (neurons)
- One data point

T points in a N-dimensional space

- How are these points distributed?
- Can this distribution be matched on position in the maze?

Dimensionality reduction tools

High dim

Low dim
II) Population analysis, no averaging

**Advantages:**
- no a priori about the coding
- use all the information
- no averaging over laps

**Cons:**
- signal could be lost if only a small fraction of cells are relevant

T (time steps) → T points in a N-dimensional space

How are these points distributed?
Can this distribution be matched on position in the maze?

Dimensonality reduction tools

High dim → Low dim

Congrès SFP - MC3
Nicolas Levernier
Dimensionality reduction, concept

Find a projection in a low dim space preserving the « structure » of the data
- Usual methods: linear projection, such as PCA

Activity in motor cortex of reaching monkey
[Churchland et al, Nature 2013]
- Usual methods: linear projection, such as PCA

*But only works if the data lies on a low dimension vectorial space*

Activity in motor cortex of reaching monkey
[Churchland et al, Nature 2013]
Dimensionality reduction

- Usual methods: linear projection, such as PCA

  *But only works if the data lies on a low dimension vectorial space*

- Non linear method are more and more popular.

Activity in motor cortex of reaching monkey [Churchland et al, Nature 2013]
Dimensionality reduction

- Usual methods: linear projection, such as PCA

  *But only works if the data lies on a low dimension vectorial space*

- Non linear method are more and more popular.

Head-direction cells activity
[Chandhuri et al, Nat Neurosci 2019]

Activity in motor cortex of reaching monkey
[Churchland et al, Nature 2013]
Dimensionality reduction

- Usual methods: linear projection, such as PCA

  *But only works if the data lies on a low dimension vectorial space*

- Non linear methods are more and more popular.

Head-direction cells activity  
[Chandhuri et al, Nat Neurosci 2019]

Activity in motor cortex of reaching monkey  
[Churchland et al, Nature 2013]

Same module grid-cells activity  
[Gardner et al, Nature 2022]
UMAP : a non-linear dimension reduction tool

[McInnes et al. arXiv 2020]

1) Compute all pair distances in the N-dim space

2) Find the best position in the low dim space such that the distribution of distances looks similar
UMAP: a non-linear dimension reduction tool

[McInnes et al. arXiv 2020]

1) Compute all pair distances in the N-dim space

2) Find the best position in the low dim space such that the distribution of distances looks similar

Mathematically: minimize the function

\[ \mathcal{L}_{\text{UMAP}} = - \sum_{i \neq j} v_{ij} \log(w_{ij}) + (1 - v_{ij}) \log(1 - w_{ij}) \]
UMAP: a non-linear dimension reduction tool

[McInnes et al. arXiv 2020]

1) Compute all pair distances in the N-dim space

2) Find the best position in the low dim space such that the distribution of distances looks similar

Mathematically: minimize the function

\[ L_{UMAP} = -\sum_{i \neq j} v_{ij} \log(w_{ij}) + (1 - v_{ij}) \log(1 - w_{ij}) \]

Interpretation

High dim space:

Distance in high dim space

Low dim space:

Distance in low dim space
UMAP: a non-linear dimension reduction tool

[McInnes et al. arXiv 2020]

1) Compute all pair distances in the N-dim space

2) Find the best position in the low dim space such that the distribution of distances looks similar

Mathematically: minimize the function

$$L_{UMAP} = - \sum_{i \neq j} v_{ij} \log(w_{ij}) + (1 - v_{ij}) \log(1 - w_{ij})$$

Interpretation

In the low dimension space:
- points that are close in the initial space will attract
- points that are far in the initial space will repulse

High dim space:

Low dim space:

Distance in high dim space

Distance in low dim space
UMAP: a non-linear dimension reduction tool

[McInnes et al. arXiv 2020]

1) Compute all pair distances in the N-dim space

2) Find the best position in the low dim space such that the distribution of distances looks similar

Mathematically: minimize the function

$$\mathcal{L}_{\text{UMAP}} = - \sum_{i \neq j} v_{ij} \log(w_{ij}) + (1 - v_{ij}) \log(1 - w_{ij})$$

Interpretation

In the low dimension space:
- points that are close in the initial space will attract
- points that are far in the initial space will repulse

In the low dim space, points behave like a gaz of particles with pairwise interactions

High dim space:  
Distance in high dim space  
\(v\)  
\(1\) 
Low dim space:  
Distance in low dim space  
\(w\)  
\(1\)
Application to spatial coding without visual cues

N (neurons) vs T (time steps)

Neuron 1
Neuron 2
Neuron 3
Application to spatial coding without visual cues

From dimension 64 to dimension 2
Application to spatial coding without visual cues

From dimension 64 to dimension 2
Application to spatial coding without visual cues

From dimension 64 to dimension 2

Neuron 1

Neuron 2

Neuron 3

No Object
Application to spatial coding without visual cues

Path integration

From dimension 64 to dimension 2
Application to spatial coding without visual cues
Application to spatial coding without visual cues
Application to spatial coding without visual cues

**Stops : Resetting of path integration !**
Application to spatial coding without visual cues

_Sops : Resetting of path integration !_
Application to spatial coding with visual cues
Application to spatial coding with visual cues

Objects

Graphs showing position and UMAP 1-2 with distance (cm) axes.
Cues affect the mental representation only locally
I) Experimental data and analysis with dimension reduction tools

II) Theoretical analysis of this dimension reduction tool
Interpreting UMAP projection quantitatively?

UMAP on a line in high dimension $N$
+ noise of amplitude $e$

Increase $e$
Interpreting UMAP projection quantitatively?

UMAP on a line in high dimension $N$
+ noise of amplitude $e$

What is the critical noise $e$ to detect signal?

Can we infer signal-noise ratio from UMAP representation?

Is UMAP optimal?

Statistical physics tools!
Well-defined data : pure noise

Coulomb gas with quenched charges

$$\mathcal{L}_{\text{UMAP}} = - \sum_{i \neq j} v_{ij} \log(w_{ij}) + (1 - v_{ij}) \log(1 - w_{ij})$$

$$v_{ij} = \exp\left(-\frac{d_{ij} - \rho_i}{\sigma}\right) + \exp\left(-\frac{d_{ij} - \rho_j}{\sigma}\right) - \exp\left(-\frac{2d_{ij} - \rho_i - \rho_j}{\sigma}\right)$$

$$w_{ij} = \frac{1}{1 + a \|y_i - y_j\|^2b}$$

a, b, σ : Umap parameters
Coulomb gas with quenched charges

$$\mathcal{L}_{\text{UMAP}} = - \sum_{i \neq j} v_{ij} \log(w_{ij}) + (1 - v_{ij}) \log(1 - w_{ij})$$

$$v_{ij} = \exp\left(-\frac{d_{ij} - \rho_i}{\sigma}\right) + \exp\left(-\frac{d_{ij} - \rho_j}{\sigma}\right) - \exp\left(-\frac{2d_{ij} - \rho_i - \rho_j}{\sigma}\right)$$

$$w_{ij} = \frac{1}{1 + a \|y_i - y_j\|^2b}$$

a, b, σ : Umap parameters

Uniform ν

Girko-Ginibres’s circular law.
Well-defined data: pure noise

Coulomb gas with quenched charges

\[ \mathcal{L}_{\text{UMAP}} = - \sum_{i \neq j} v_{ij} \log(w_{ij}) + (1 - v_{ij}) \log(1 - w_{ij}) \]

\[ v_{ij} = \exp \left( -\frac{d_{ij} - \rho_i}{\sigma} \right) + \exp \left( -\frac{d_{ij} - \rho_j}{\sigma} \right) - \exp \left( -\frac{2d_{ij} - \rho_i - \rho_j}{\sigma} \right) \]

\[ w_{ij} = \frac{1}{1 + a \|y_i - y_j\|^2} \]
a, b, \sigma: Umap parameters

Uniform \( \nu \)

Normally iid points in high D space

Girko-Ginibres’s circular law.
Well-defined data : binary signal

\[ \mathcal{L}_{\text{UMAP}} = - \sum_{i \neq j} v_{ij} \log(w_{ij}) + (1 - v_{ij}) \log(1 - w_{ij}) \]

**Binary signal**

Two communities 1 and 2.

\[ v_{11} = v_{22} = \nu \]
\[ v_{12} = \nu - \delta\nu \]
Well-defined data: binary signal

\[ \mathcal{L}_{\text{UMAP}} = - \sum_{i \neq j} v_{ij} \log(w_{ij}) + (1 - v_{ij}) \log(1 - w_{ij}) \]

Binary signal

Two communities 1 and 2.

\[ v_{11} = v_{22} = \nu \]
\[ v_{12} = \nu - \delta \nu \]
Well-defined data: binary signal

$$\mathcal{L}_{\text{UMAP}} = -\sum_{i \neq j} v_{ij} \log(w_{ij}) + (1 - v_{ij}) \log(1 - w_{ij})$$

**Binary signal**

Two communities 1 and 2.

$$v_{11} = v_{22} = v$$

$$v_{12} = v - \delta v$$

With noise, work in progress
Well-defined data: continuous signal

1d continuous signal

\[ x^k_i = i/N \delta_{0k} + \epsilon \eta^k_i \]

Work in progress...
Conclusion

UMAP seems to be a good way to represent our data:

- use the **whole signal**
- **no a priori** about the coding scheme
- **no averaging**
- **can detect** some experimental issues
Conclusion

UMAP seems to be a good way to represent our data:

- use the **whole signal**
- **no a priori** about the coding scheme
- **no averaging**
- **can detect** some experimental issues
Conclusion

UMAP seems to be a good way to represent our data:

- use the **whole signal**
- **no a priori** about the coding scheme
- **no averaging**
- **can detect** some experimental issues
Conclusion

UMAP seems to be a good way to represent our data:
- use the **whole signal**
- **no a priori** about the coding scheme
- **no averaging**
- **can detect** some experimental issues

We have highlighted the **robust path integration**.

On top of distance coding, some **position coding** emerge when objects are present.

Theoretical progress to use UMAP in a **quantitative** way, or to **improve** the algorithm.
Acknowledgements

Hervé Rouault

Mathilde Nordlund, Jérôme Epsztein and Julie Koenig

Remi Monasson, Massimilliano Trippa