Recitation — t-SNE

Emaad Ahmed Manzoor

January 25, 2019
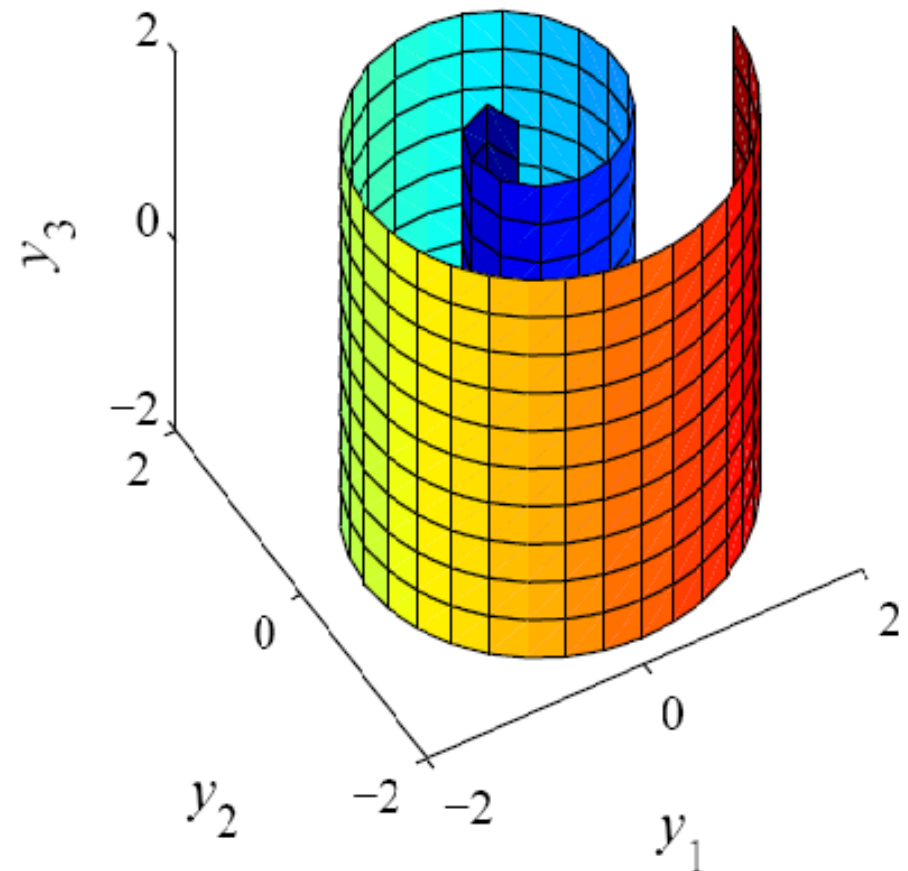
Manifold Learning

$$\mathcal{X} = \{x_1, x_2, ..., x_n \in \mathbb{R}^h\} \rightarrow \mathcal{Y} = \{y_1, y_2, ..., y_n \in \mathbb{R}^l\}$$

$$\min_{\mathcal{Y}} C(\mathcal{X}, \mathcal{Y})$$

Preserve distances

Preserve topology

Distances to Probabilities — Why?

$$p_{j|i} = \frac{\exp(- \| x_i - x_j \|^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(- \| x_i - x_k \|^2 / 2\sigma_i^2)}$$

Original data

$$q_{ij} = \frac{(1 + \| y_i - y_j \|^2)^{-1}}{\sum_{k \neq l} (1 + \| y_k - y_l \|^2)^{-1}}$$
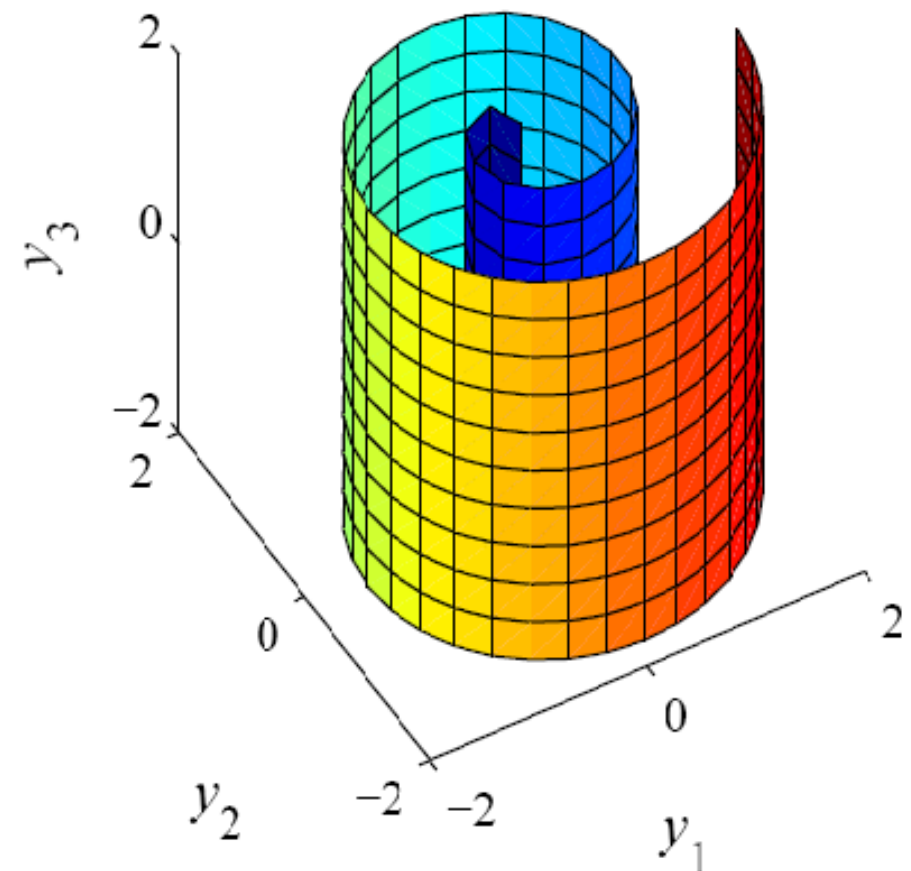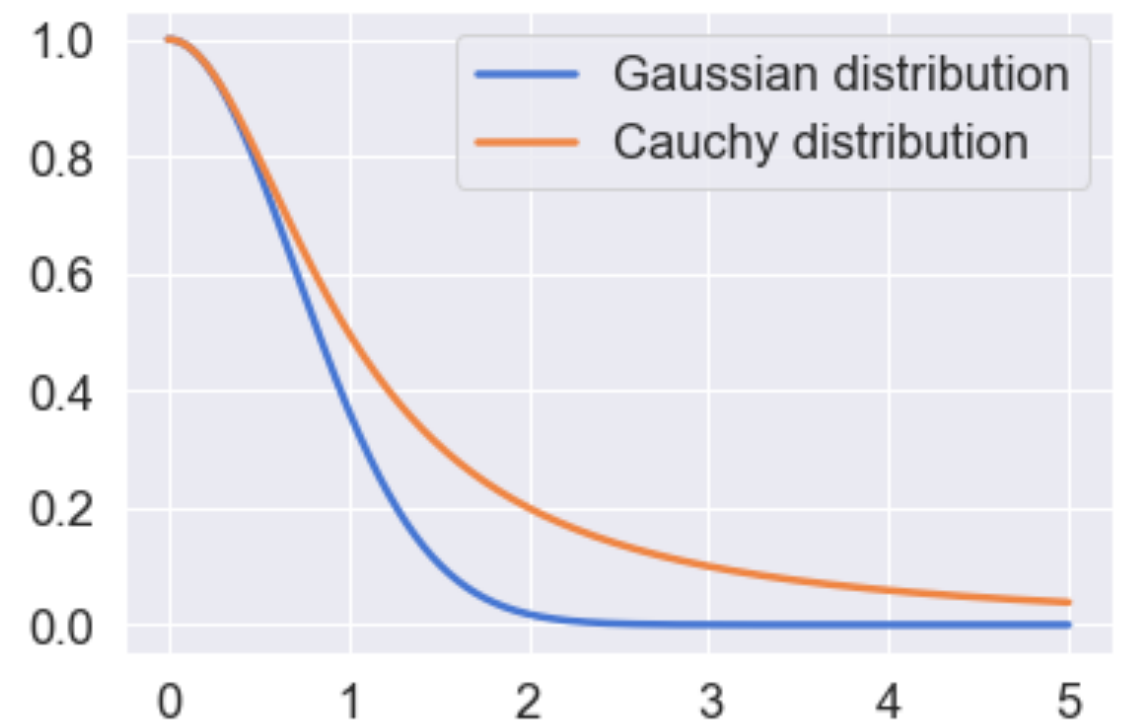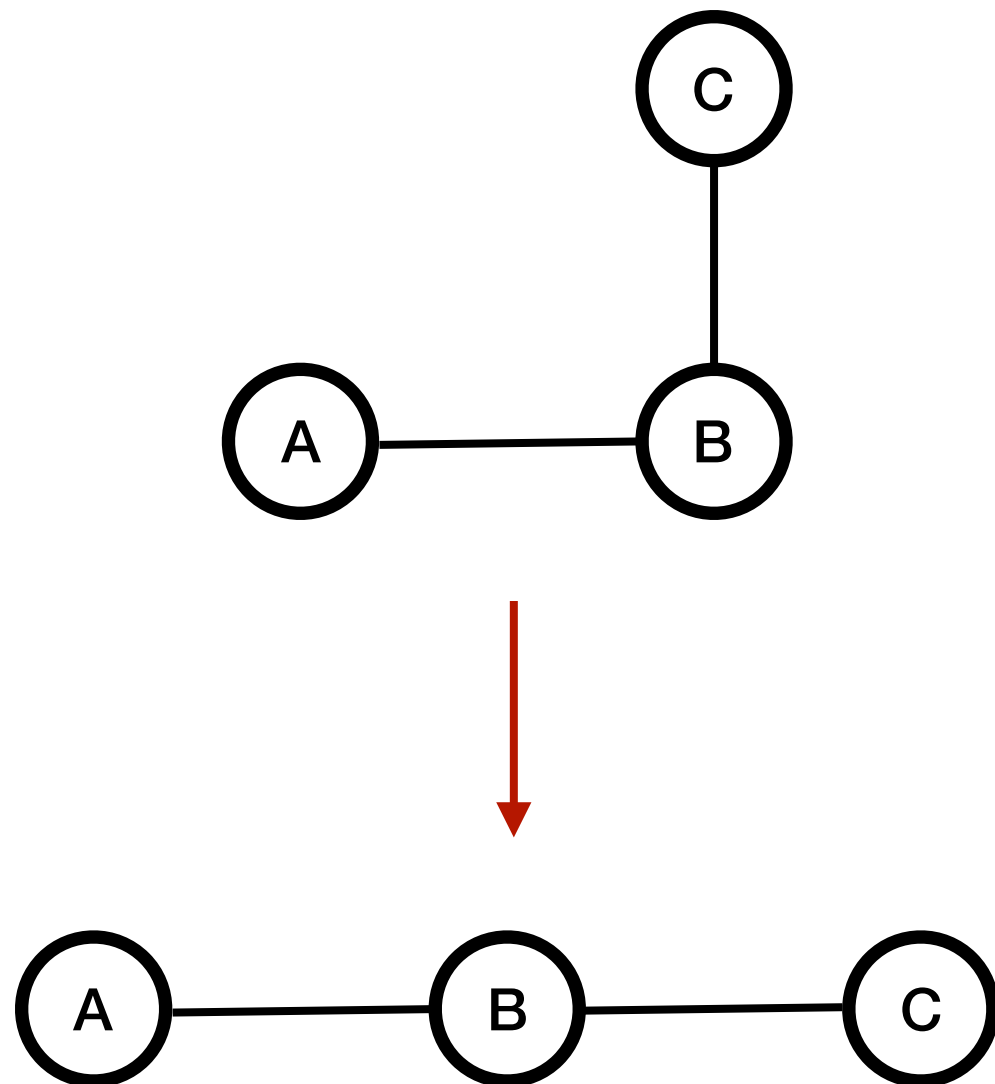
Mapped data

Uncertainty

Local vs. global topology

Perplexity & density

Why the t-Distribution?

To model local structure, allow
dissimilar points to be *far!*

https://gist.github.com/emaadmanzoor/
f20ee72b5635e1b86223d1c9fa25cc53