Thermodynamic Manifold Inference

Karthik Srinivasan

Dimensionality reduction

- Matrix Factorisation methods
- Examples: PCA, SVD, NMF
- Data is represented as a product of multiple simpler matrices
- Allows for *approximate* reconstruction of data
- Can not tell us anything about the manifold

- Manifold learning methods
- Examples: diffusion maps, Isomaps, tSNE, UMAP, Laplacian Eigenmaps
- Assumes that the data lies on a much lower dimensional embedded manifold
- Uses kernel based approaches
- Allows inference of manifold
- Can not represent the data using approximate reconstruction

Thermodynamic manifold inference

- Based on statistical physics
- Offers physically interpretable solution to dimensionality reduction
- Assumes that the data can be represented by a Gibbs-Boltzmann distribution
- It has modeled probabilities in a variety of complex systems such as such as ensembles of protein sequences, signal networking, collective firing of neurons, collective motion of birds

$$Z^{(\alpha)} = \sum_{a} \exp\left(-\sum_{k=1}^{K} \lambda_k^{(\alpha)} Y_{ka}\right)$$

$$q_a^{(\alpha)} = \frac{1}{Z^{(\alpha)}} \exp\left(-\sum_{k=1}^{K} \lambda_k^{(\alpha)} Y_{ka}\right)$$

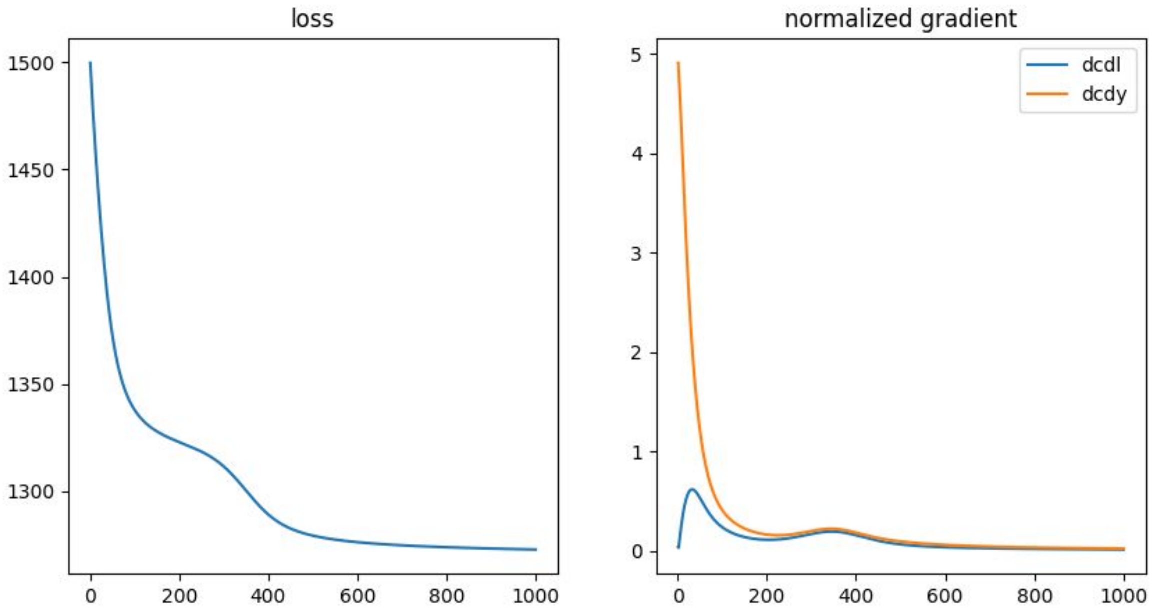
Gradient descent

$$C = \sum_{\alpha} \sum_{a} \mathbf{x}_a^{(\alpha)} \ln \frac{\mathbf{x}_a^{(\alpha)}}{q_a^{(\alpha)}}$$

$$\lambda_k^{(\alpha)} \leftarrow \lambda_k^{(\alpha)} - \eta \frac{\partial C}{\partial \lambda_k^{(\alpha)}}$$

$$Y_{ka} \leftarrow Y_{ka} - \eta \frac{\partial C}{\partial \lambda_k^{(\alpha)}}$$

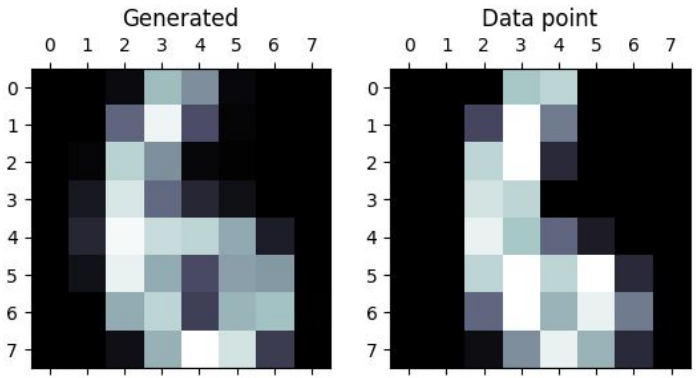
Metrics



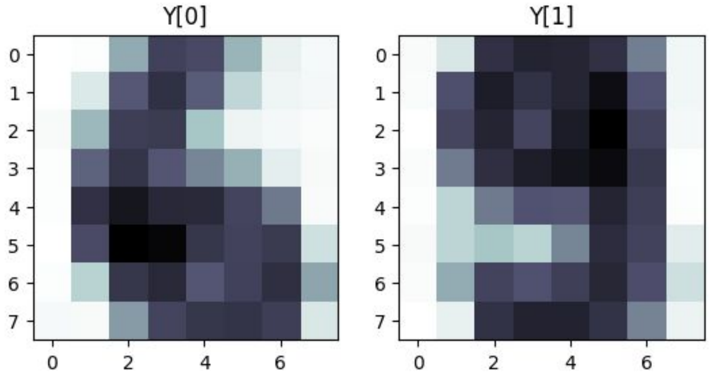
loss
1500
1450
1400
1350
1300
0
200
400
600
800
1000
normalized gradient
dcdl
dcdy
5
4
3
2
1
0
0
200
400
600
800
1000

My Results

Data reconstruction

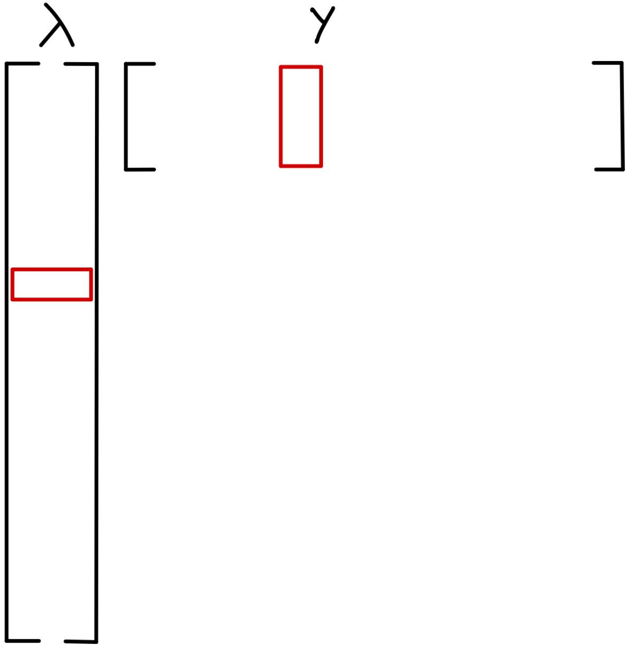


Physical interpretation of Y



Intensive and Extensive variables

- Each entry of the matrix q conveys how much the data point represents a 6 and 9 respectively.
- Each entry of Y conveys the intensity of each feature for each class.



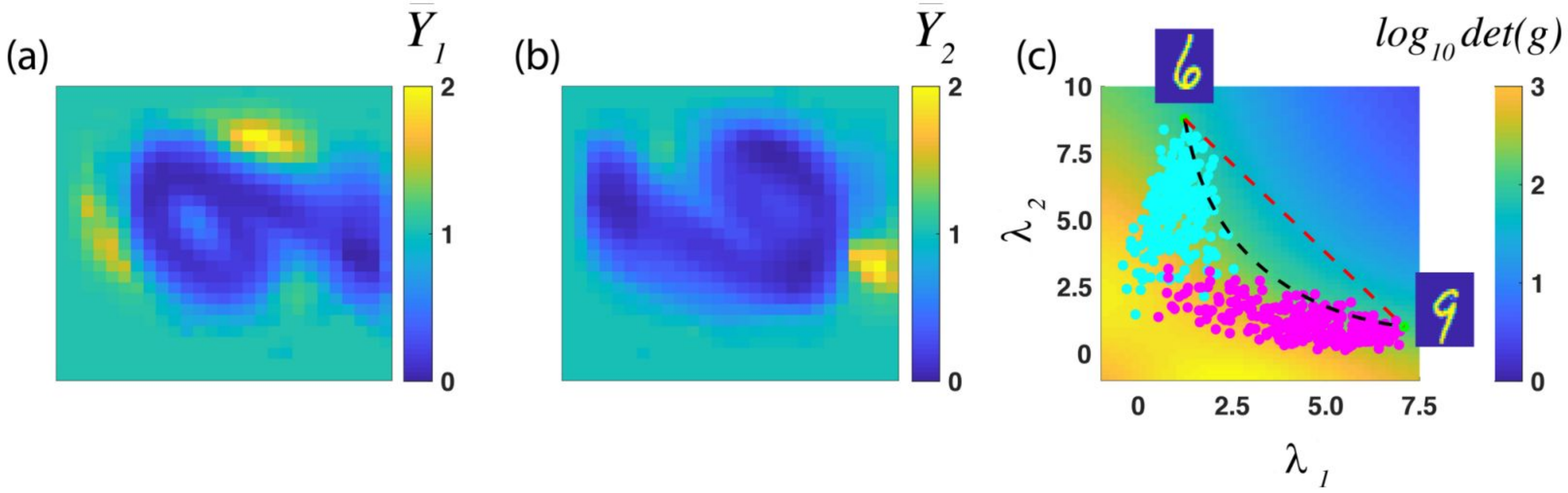
Riemannian distance metric metric

- Defined on the space of intensive variable
- The metric on the space of intensive variables defines a notion of distance between the distributions as well as the “number of points” in any given volume element

$$P \propto \int_0^T \frac{d\bar{\lambda}^{\mathrm{T}}}{dt} g(\bar{\lambda}) \frac{d\bar{\lambda}}{dt} dt$$

$$g_{ij}(\bar{\lambda}) = \int_0^\infty \langle \delta \bar{Y}_i(0) \delta \bar{Y}_j(\tau) \rangle_{\bar{\lambda}} d\tau$$

Paper results



Thank you!