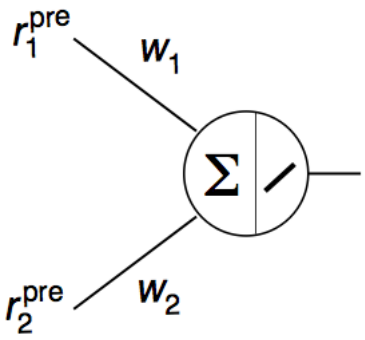
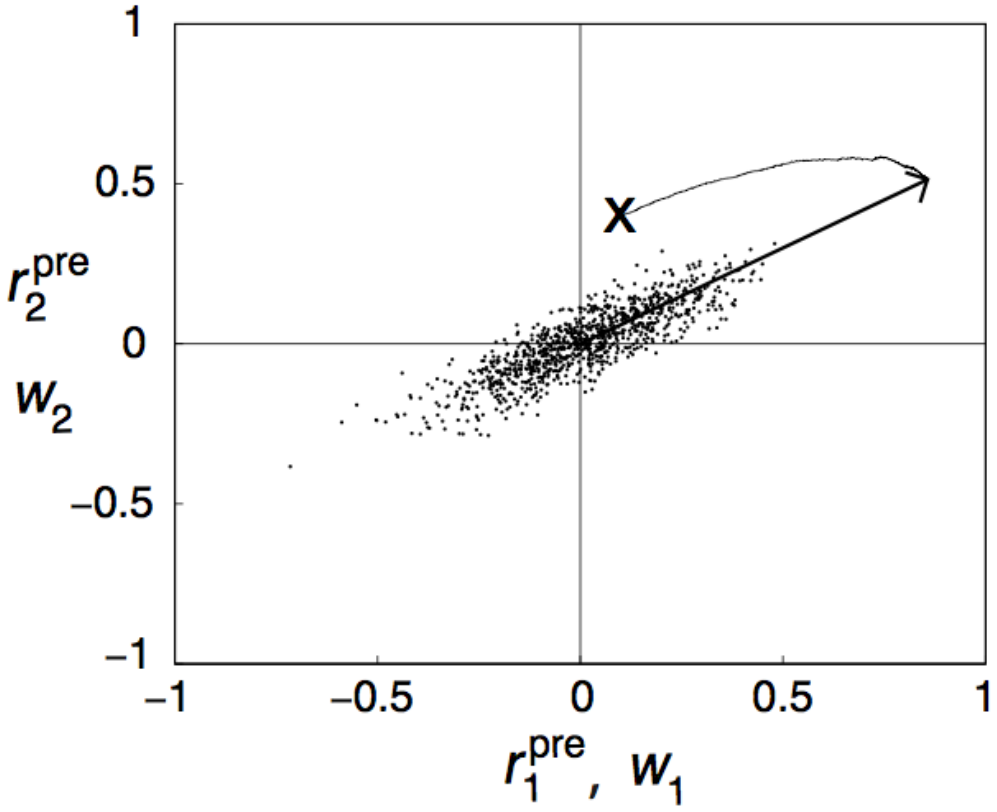
Dimensionality reduction
end
Representational learning

PCA: Principal Component Analysis

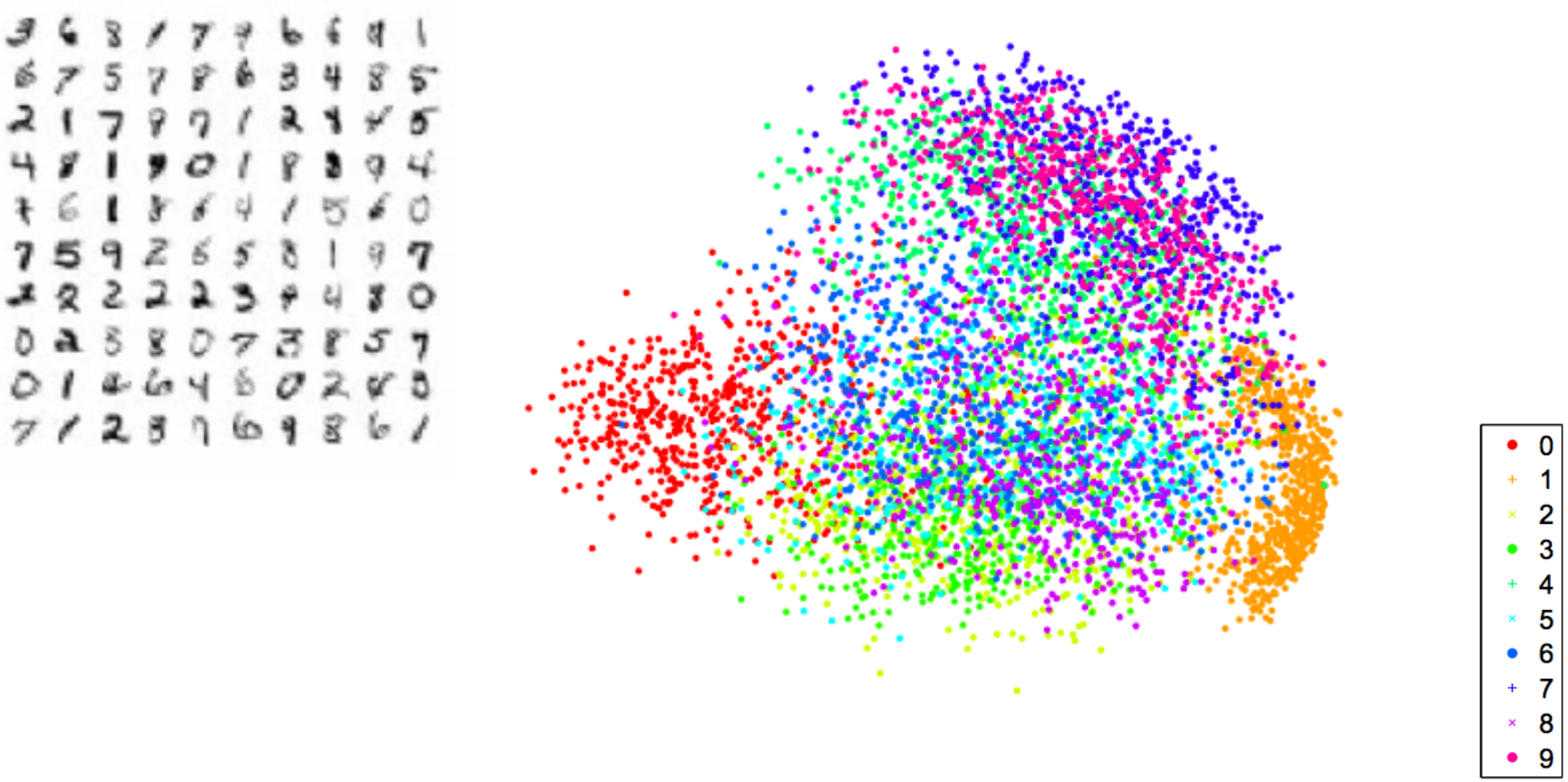




$$\Delta w_{ij} = r_i r_j - (r_i)^2 w_{ij}$$

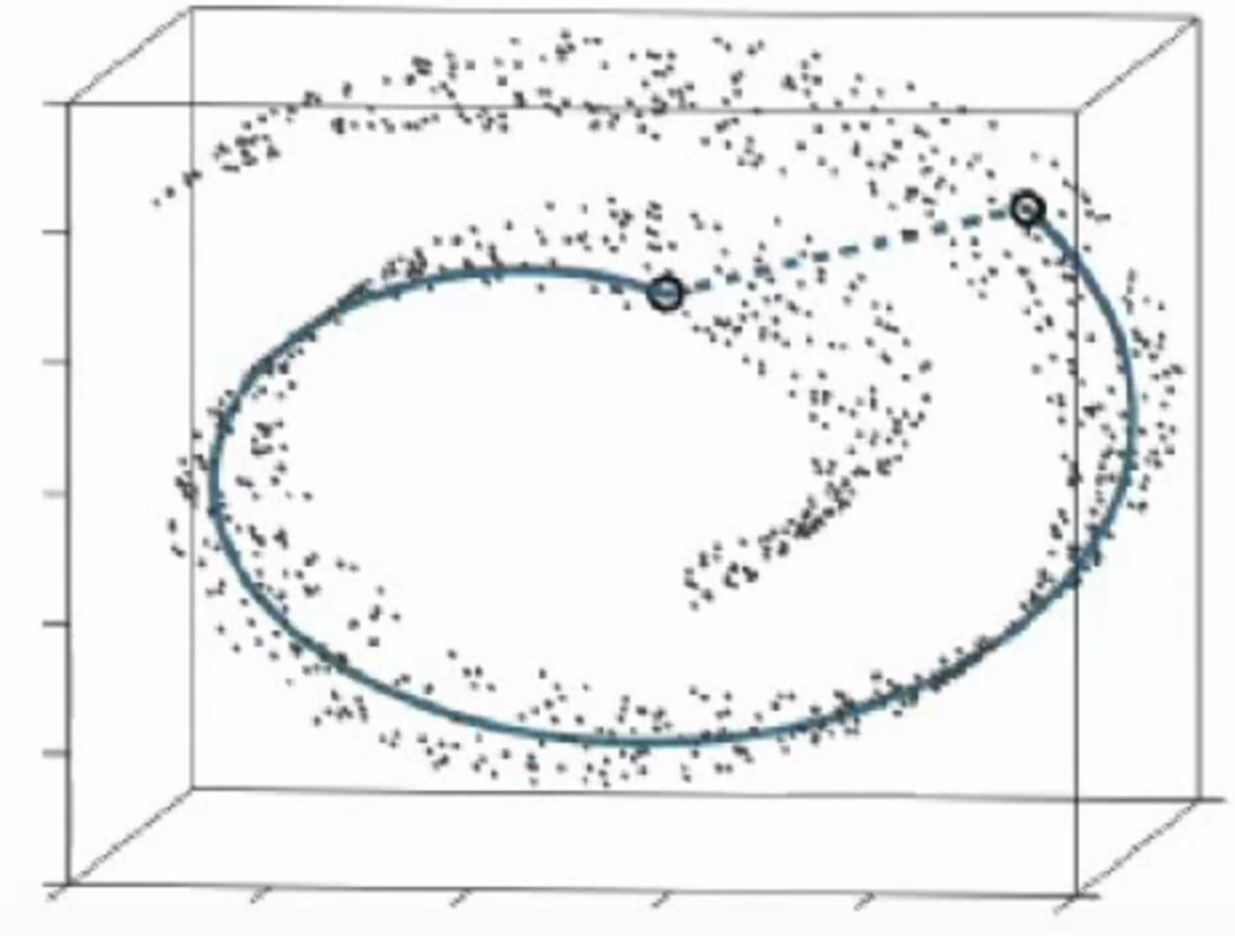
Linear transformation that maximizes variance

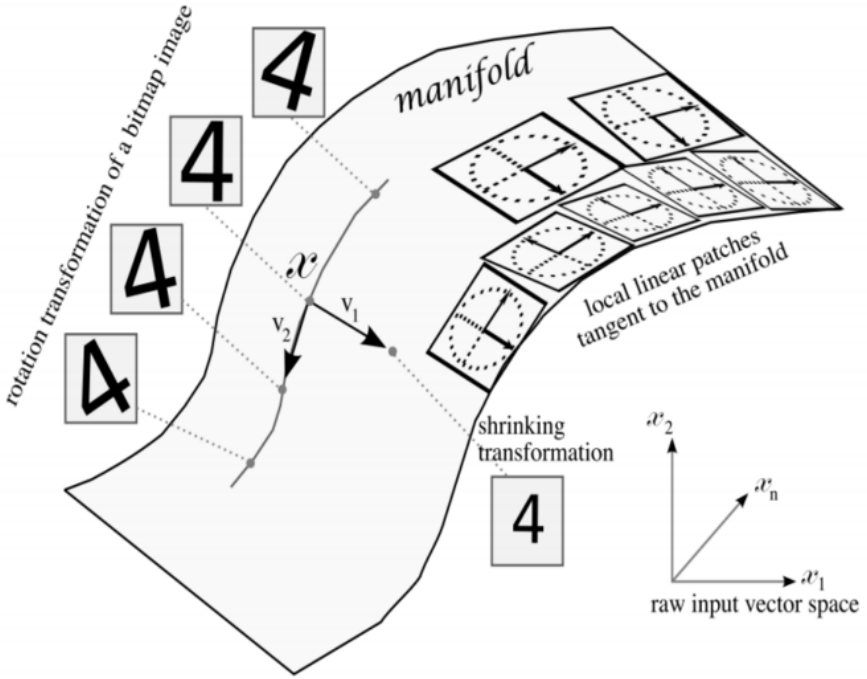
PCA of MNIST data

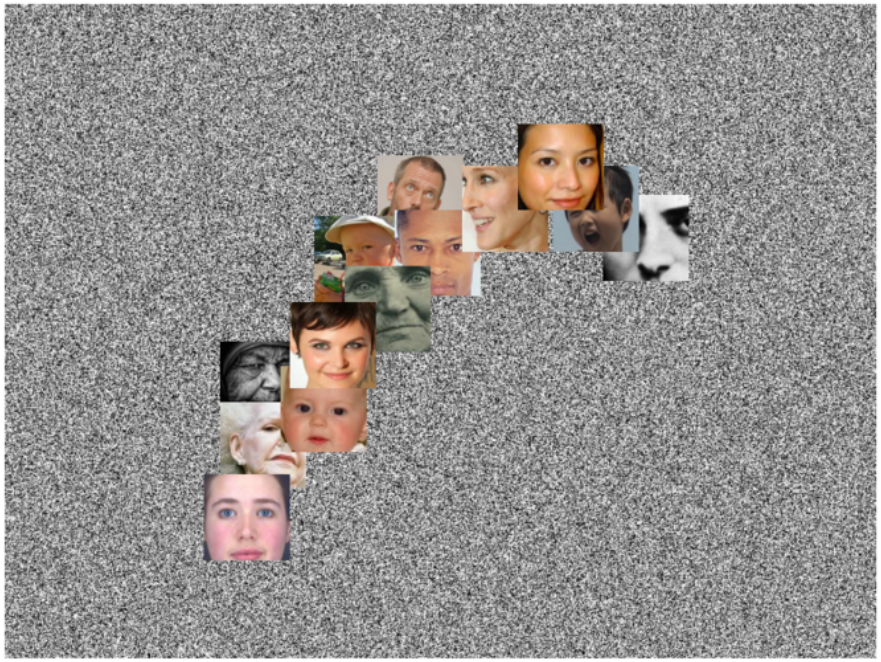


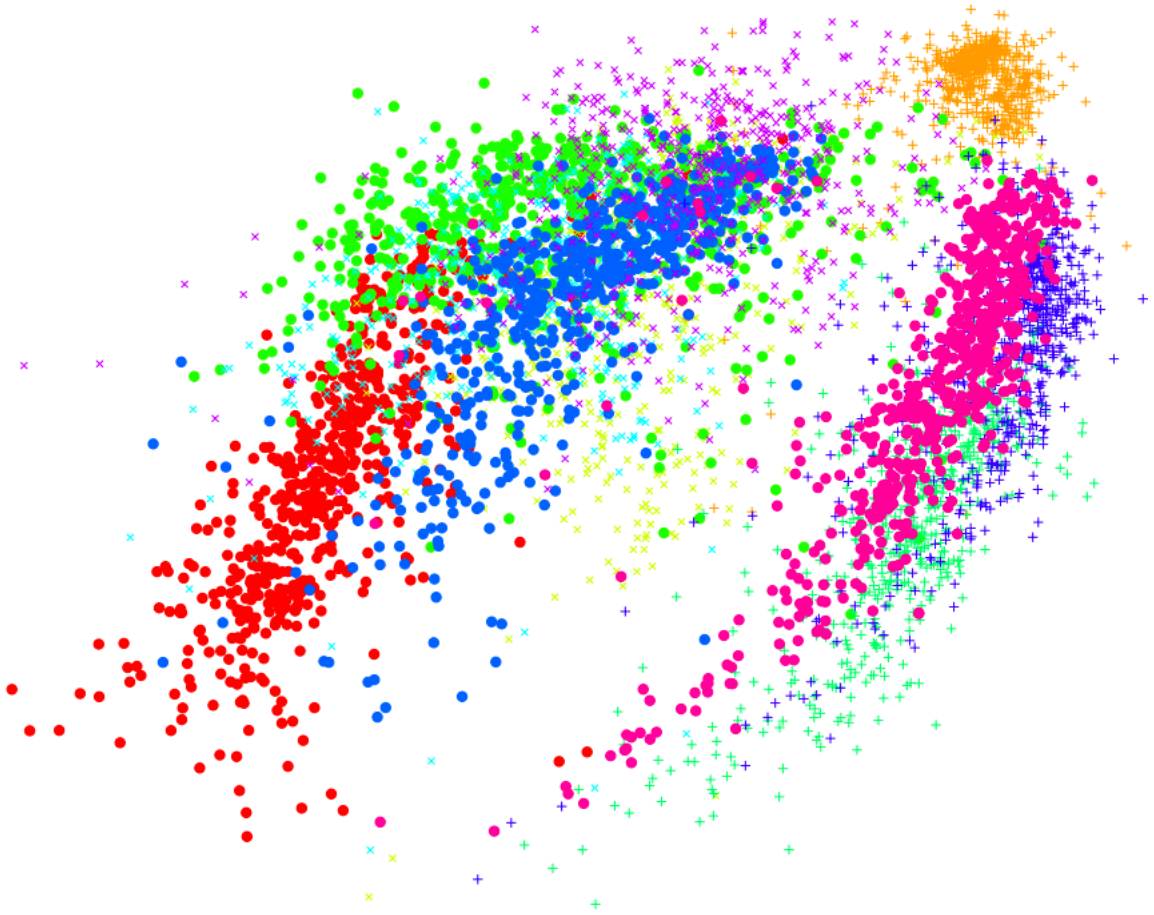
(a) Visualization by PCA.

Preserve large pairwise distances



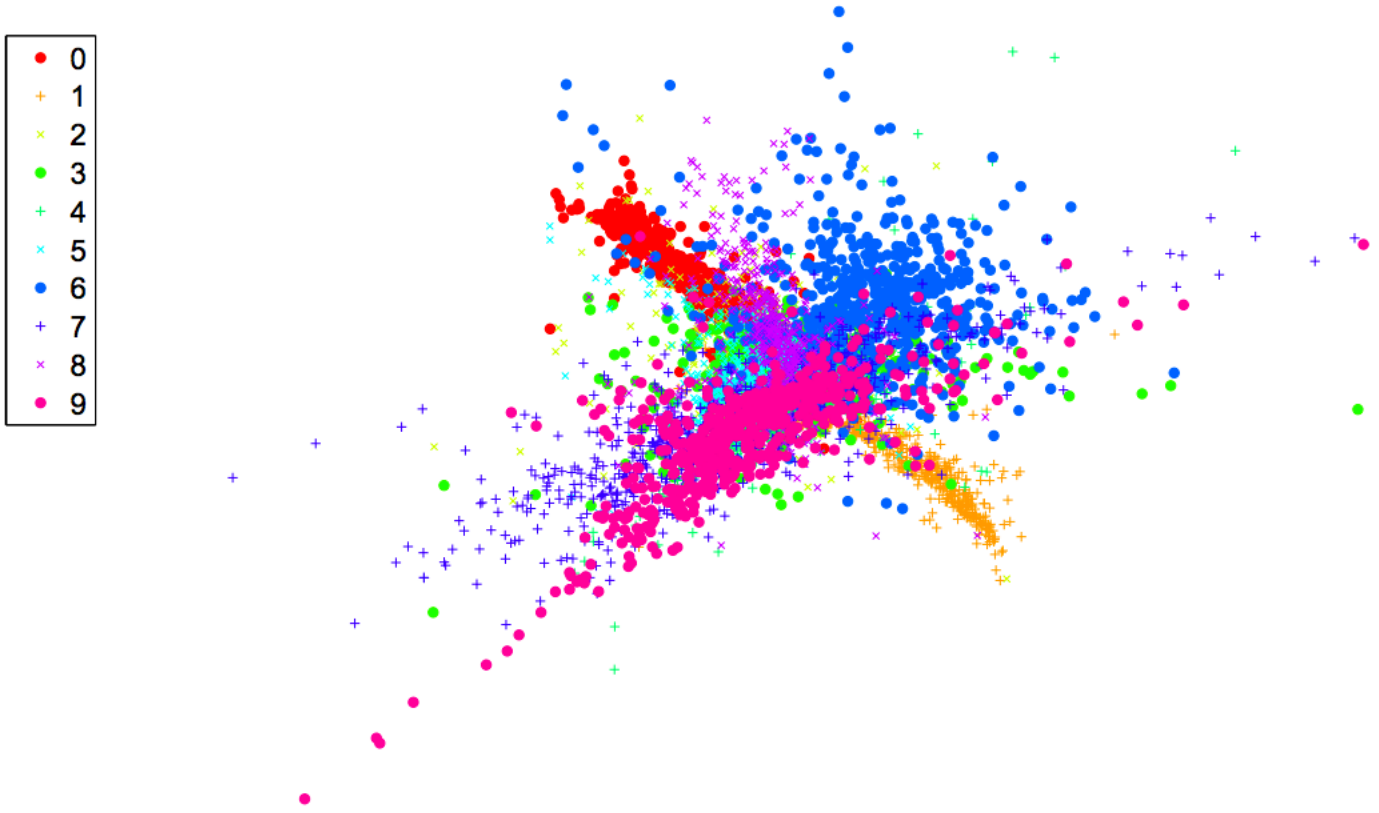

rotation transformation of a bitmap image
manifold
x
v1
v2
local linear patches
tangent to the manifold
shrinking
transformation
x2
xn
x1
raw input vector space





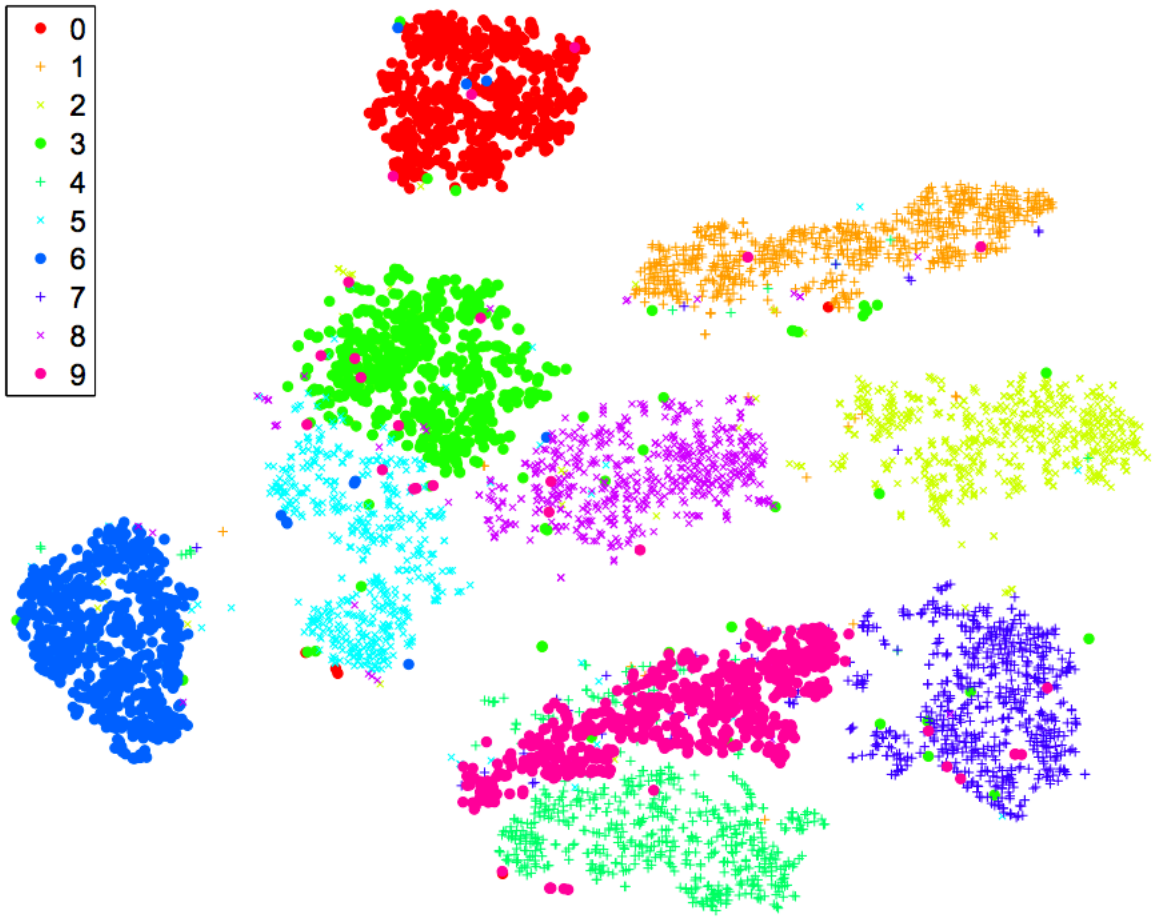
(a) Visualization by Isomap.

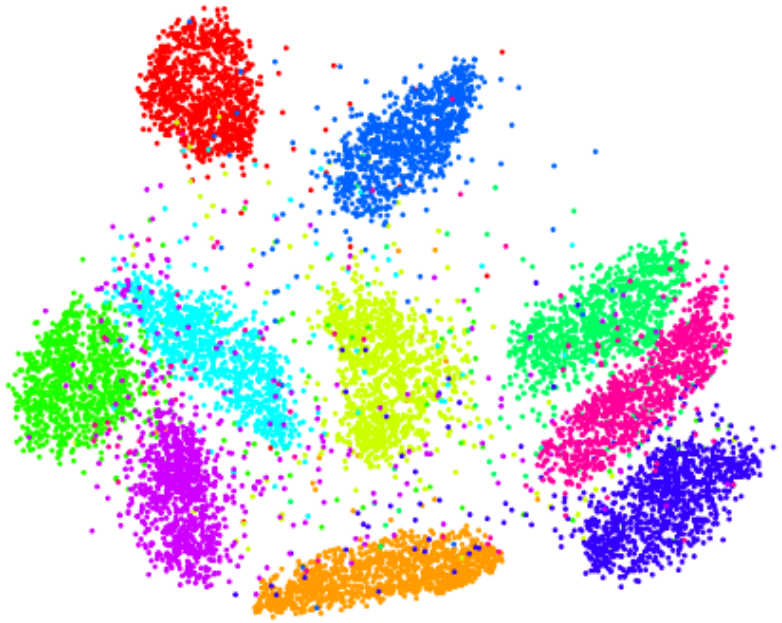
Local Linear Embedding



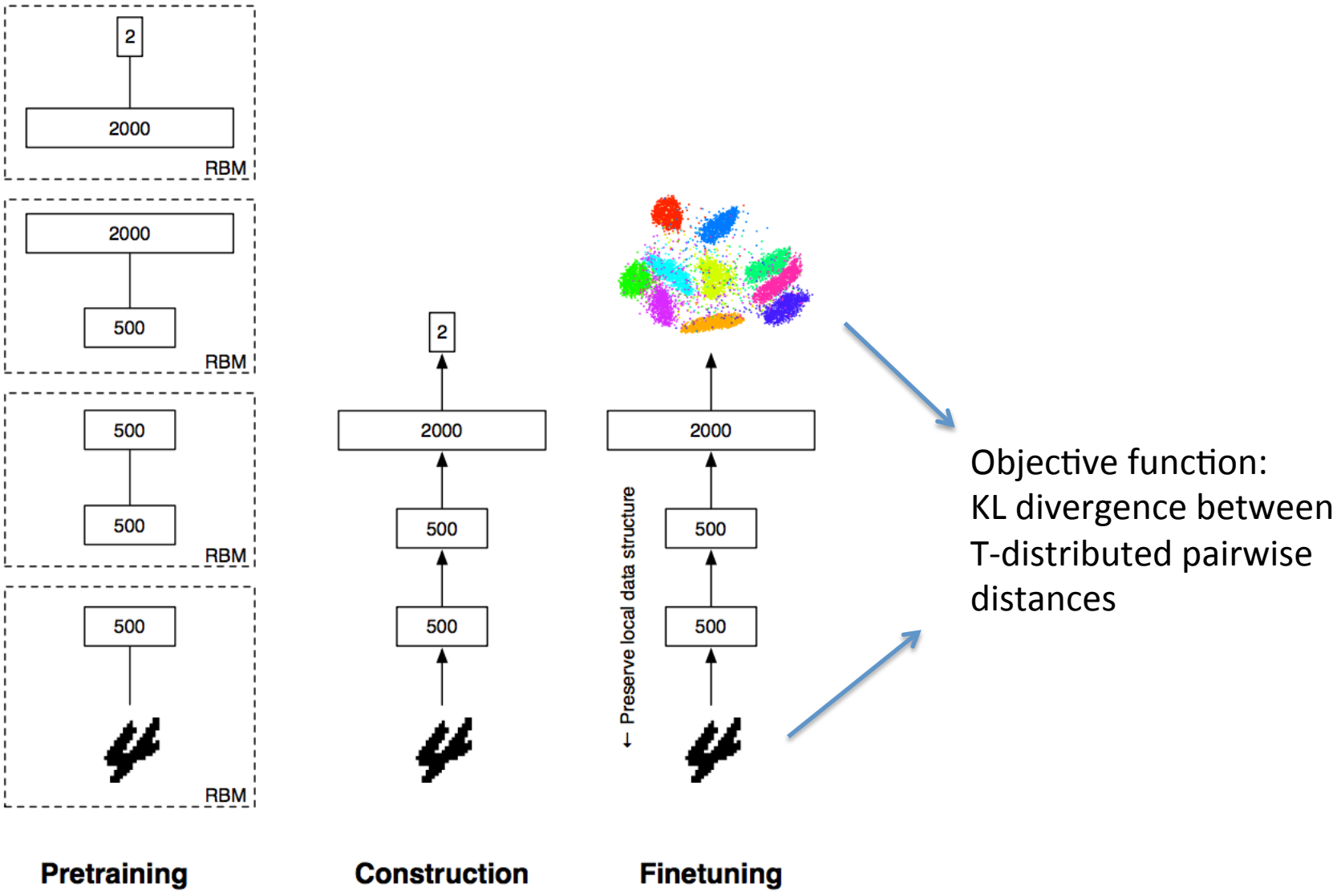
(b) Visualization by LLE.

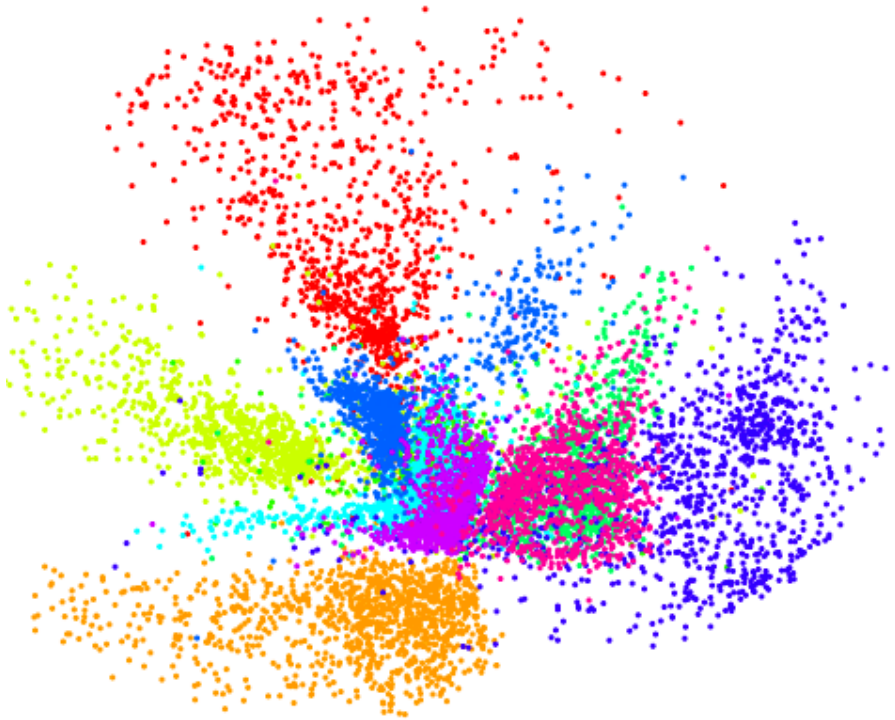
T_SNE (van der Maaten, Hinton 2008)





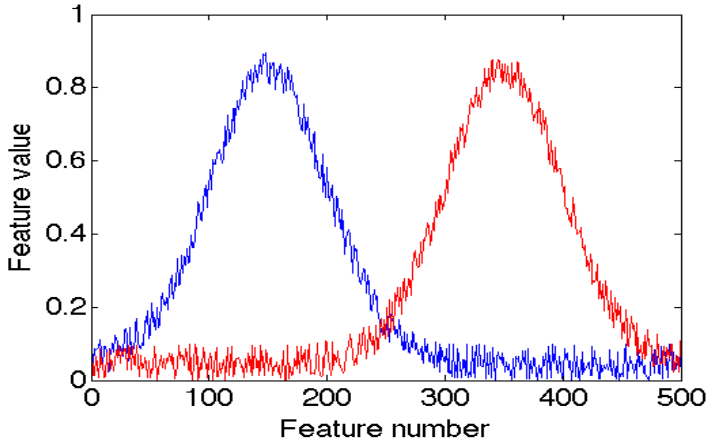
c) Visualization by parametric t-SNE.

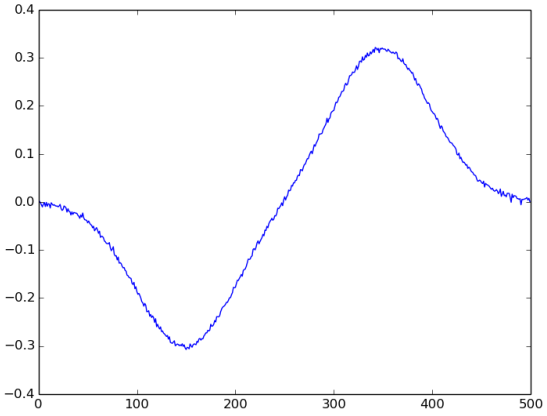

2
2000
RBM
2000
500
RBM
500
500
RBM
500
RBM
2
2000
500
500
2000
500
500
← Preserve local data structure
Objective function:
KL divergence between
T-distributed pairwise
distances
Pretraining
Construction
Finetuning

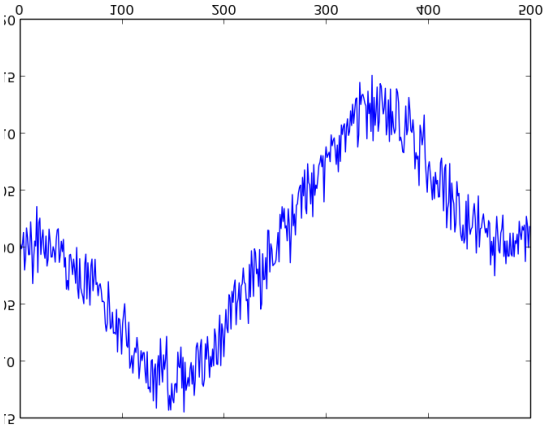


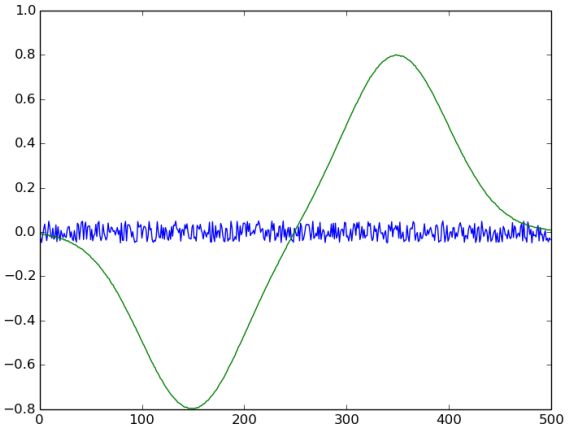
(b) Visualization by an autoencoder.

3 simple examples

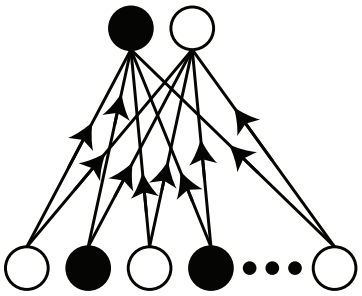




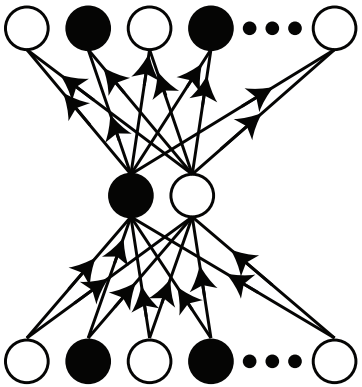
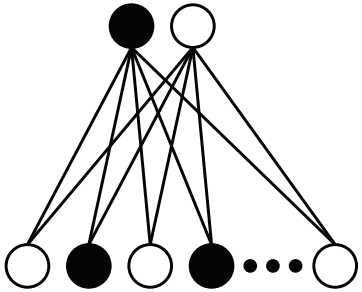




Perceptron



Autoencoder



RBM