Auto-encoders

Mladen Kolar (mkolar@chicagobooth.edu)

Dimension reduction by linear transformation

PCA is linear dimension reduction technique.

It seeks a parsimonious representation for data in a large (p) dimensional space.

So the goal is to map p-dimensional vectors x_i, for $i = 1, \ldots, n$

- assembled in an $n \times p$ design matrix X

onto an $m \ll p$ dimensional space, wherein they take on the representation z_i, for $i = 1, \ldots, n$

- assembled into the $n \times m$ design matrix Z.

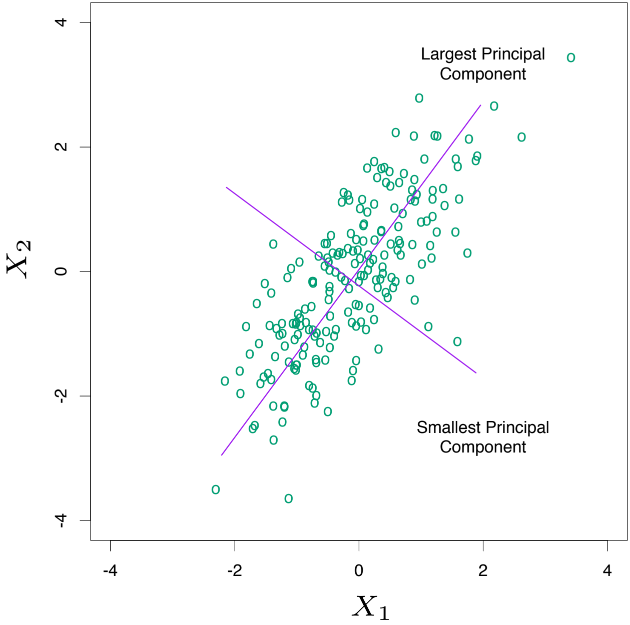
Once in the lower (m) dimensional space

- regressions are easier,
- predictions are more stable,
- and interpretation is parsimonious.

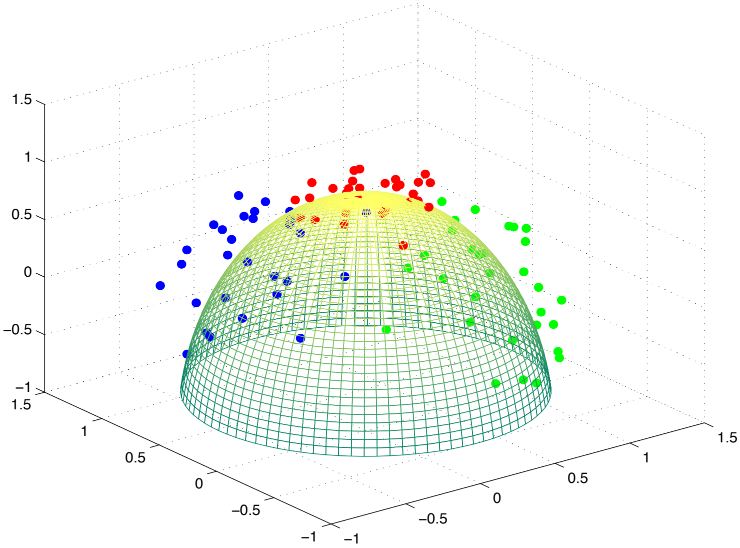
In the (linear) regression context, which is the biggest consumer of dimension reduction techniques, the fact that the procedure operates only on the x's

- ignoring the response (variable of interest), y

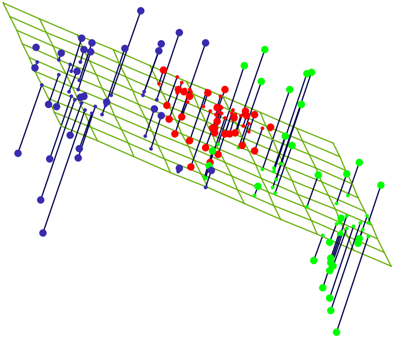
classifies PCA as form of unsupervised learning.

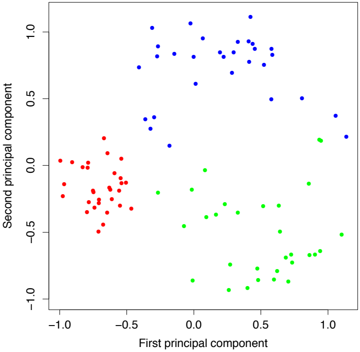

Largest Principal
Component
Smallest Principal
Component
X_1
X_2
-4
-2
0
2
4

Half-sphere data:

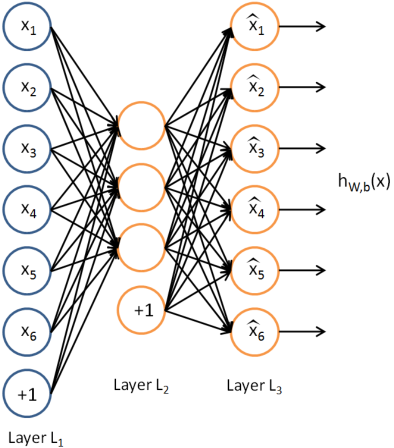


Projecting down onto the right linear subspace helps separate the colors.





Autoencoder



Network trained to reproduce its input at the output layer.
Usually tie the weights that go into and out of the hidden layer.

Autoencoder

Loss function

- For real valued inputs, try to find weights such that

$$\frac{1}{2}\sum_k (x_k - \hat{x}_k)^2$$

  is minimized

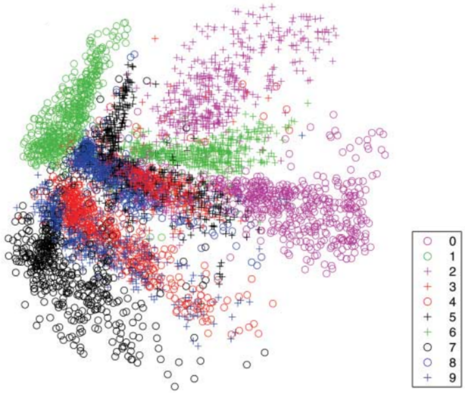
- For binary input cross entropy is used, which is similar to deviance

Fitting autoencoder

- Same tricks as before
- Greedy learning of stacked autoencoders
- https://www.cs.toronto.edu/~hinton/science.pdf

Autoencoder: Why are they useful?

Learning compressed representation of the input distribution (dimensionality reduction)



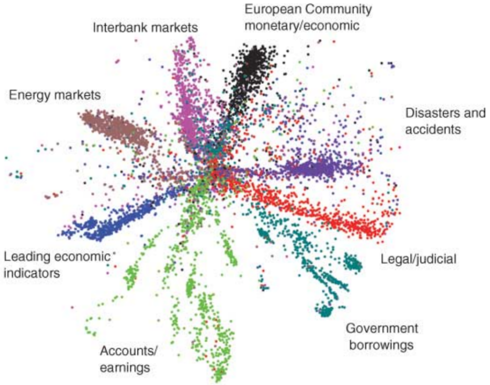
Autoencoder structure: 784 — 1000 — 500 — 250 — 2

https://www.cs.toronto.edu/~hinton/science.pdf

A comparison of methods for compressing digit images to 30 real numbers

Autoencoder: Why are they useful?

Information retrieval: 804,414 newswire stories

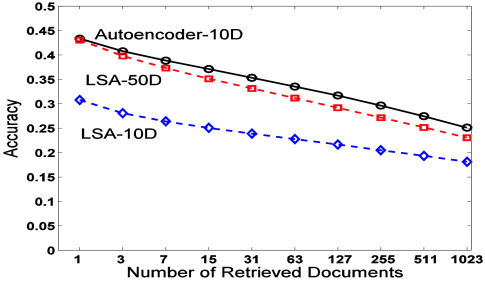


Autoencoder structure: 2000 — 500 — 250 — 125 — 2

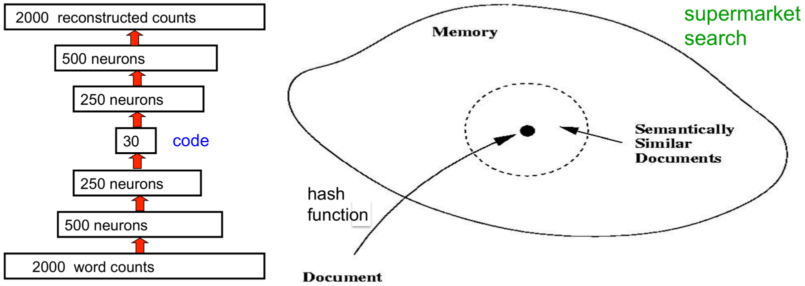
https://www.cs.toronto.edu/~hinton/science.pdf

Comparison with PCA

Retrieval performance on 400,000 Reuters business news stories

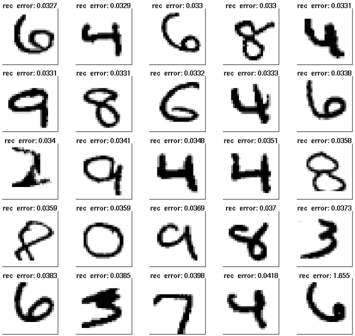


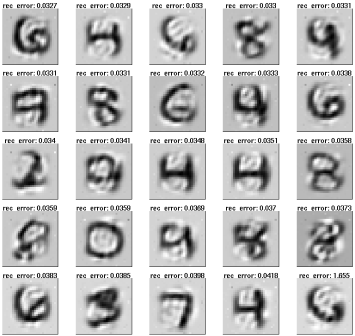
Finding binary codes for documents



Autoencoder: Why are they useful?

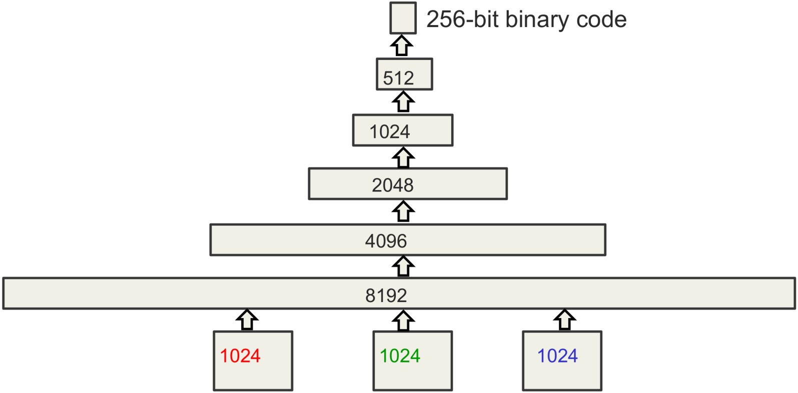
- unsupervised pretraining of weights (many unlabeled images, but only few labeled)
- anomaly detection



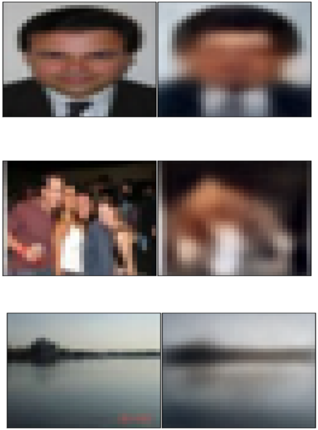


Autoencoder: Why are they useful?

Image retrieval



Reconstructions of 32x32 color images from 256-bit codes



retrieved using 256 bit codes

retrieved using Euclidean distance in pixel intensity space

retrieved using 256 bit codes

retrieved using Euclidean distance in pixel intensity space

Leftmost column is the search image.

Other columns are the images that have the most similar feature activities in the last hidden layer.