



Dimensionality Reduction

Objective



Objective

Illustrate the need for
dimensionality
reduction

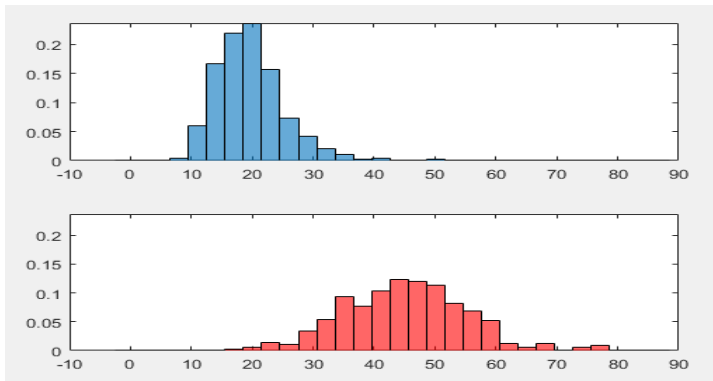
What is Dimensionality Reduction?



- | We have N data points in a high-dimensional space,
 - e.g., in the order of tens of thousands of dimensions.
- | We want to project them into some low-dimensional space,
 - e.g., in the order of tens of dimensions.
- | Why dimensionality reduction?
 - A key technique to mitigate *curse of dimensionality*

The Curse of Dimensionality

Consider histogram as a density estimator.



→ Exponentially more samples would be needed in higher-dimensional spaces for the same “resolution”.

Many Techniques for Dimensionality Reduction



| Many ways for going from a higher-dimensional space to a lower-dimensional space.

- Feature Selection achieves this by keeping only a subset of the original features/dimension.

| There are many other techniques, employing a feature mapping/projection approach.

- New features are generated (instead of selecting only from the original features).
- The underlying assumptions and/or goals of the techniques are often different.

Examples of Feature Mapping



- | Linear discriminant analysis (LDA)
- | Independent component analysis (ICA)
- | Non-negative matrix factorization (NMF)
- | Auto-encoder
- | Self-organizing maps
- | **Principal component analysis** (and its variants)