

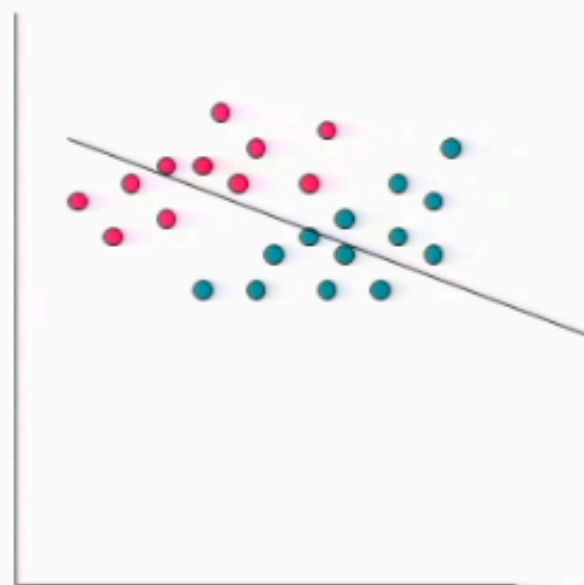
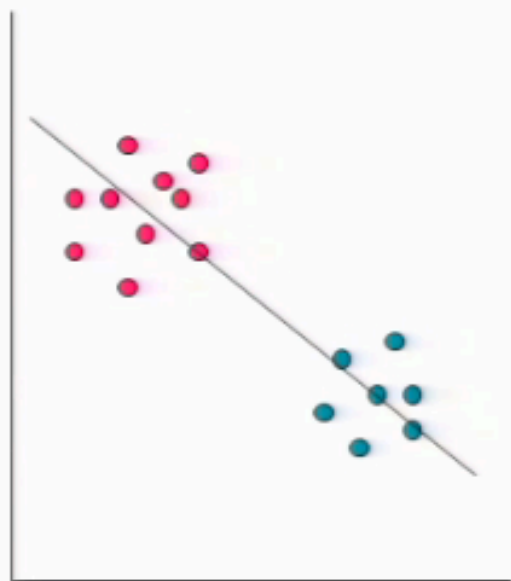
Supervised Dimensionality Reduction

Supervised Dimensionality Reduction

- Neural nets: learn hidden layer representation, designed to optimize network prediction accuracy
- PCA: unsupervised, minimize reconstruction error
 - but sometimes people use PCA to re-represent original data before classification (to reduce dimension, to reduce overfitting)
- Fisher Linear Discriminant
 - like PCA, learns a *linear* projection of the data
 - but supervised: it uses labels to choose projection

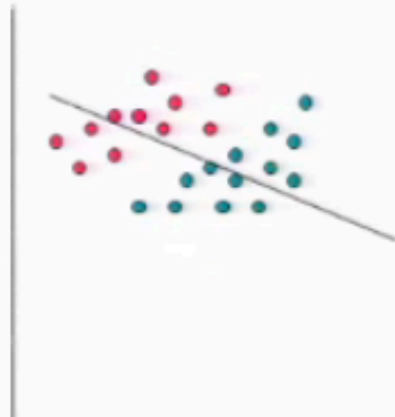
Fisher Linear Discriminant

- A method for projecting data into lower dimension to hopefully improve classification
- We'll consider 2-class case



Project data onto vector that connects class means?

Fisher Linear Discriminant



Project data onto one dimension, to help classification

$$y = \mathbf{w}^T \mathbf{x}$$

Define class means: $\mathbf{m}_i = \frac{1}{N_i} \sum_{n \in C_i} \mathbf{x}^n$

Could choose \mathbf{w} according to: $\arg \max_{\mathbf{w}} \mathbf{w}^T (\mathbf{m}_2 - \mathbf{m}_1)$

Instead, Fisher Linear Discriminant chooses: $\arg \max_{\mathbf{w}} \frac{(\mathbf{m}_2 - \mathbf{m}_1)^2}{s_1^2 + s_2^2}$

$$m_i \equiv \mathbf{w}^T \mathbf{m}_i \quad s_i^2 \equiv \sum_{n \in C_i} (x^n - m_i)^2$$