



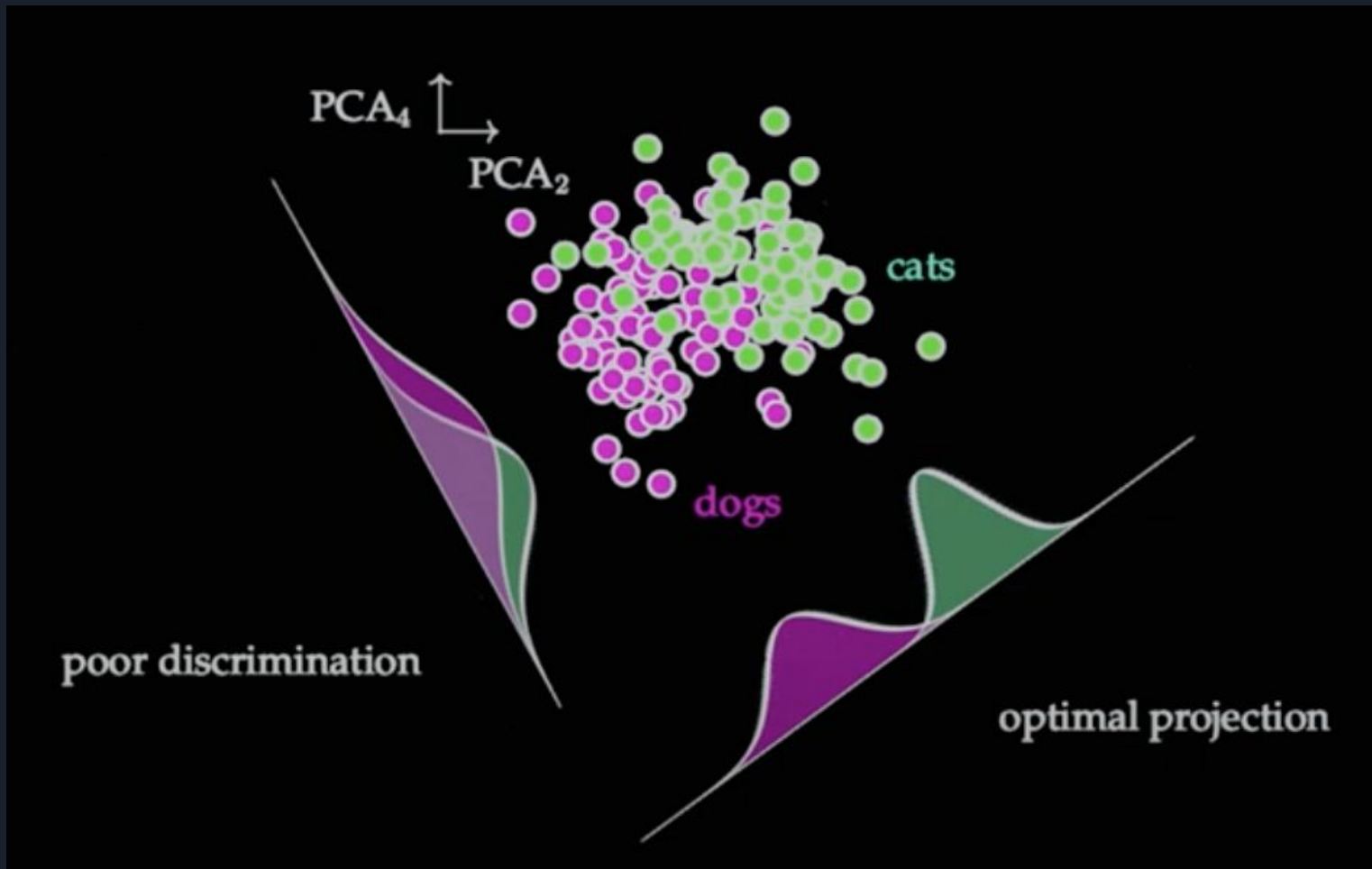
Linear Discriminant Analysis

Lenka Rozborilova
16/2 2021



Linear Discriminant Analysis (LDA)

- **supervised** learning concept that categorizes a set of data into **known** classes
- similar to Principal Component Analysis (PCA)
- aim: assign labels to the data - make an accurate classification decision
- linear technique - draw an optimal line (= decision threshold) that is discriminating between different classes in data set





Linear Discriminant Analysis (LDA)

- **Mathematics** behind (computing two quantities):
 - distance between means
 - variation of data around each mean (*scatter*)

Formal optimization

$$\mathbf{w} = \arg \max_{\mathbf{w}} \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}}$$

between- and in-class variance

$$\mathbf{S}_B = (\mu_2 - \mu_1)(\mu_2 - \mu_1)^T$$

$$\mathbf{S}_W = \sum_{j=1}^2 \sum_{\mathbf{x}} (\mathbf{x} - \mu_j)(\mathbf{x} - \mu_j)^T$$

Linear Discriminant Analysis (LDA)

- **Mathematics** behind (computing two quantities):
 - distance between means
 - variation of data around each mean (*scatter*)

Formal optimization

$$\mathbf{w} = \arg \max_{\mathbf{w}} \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}}$$

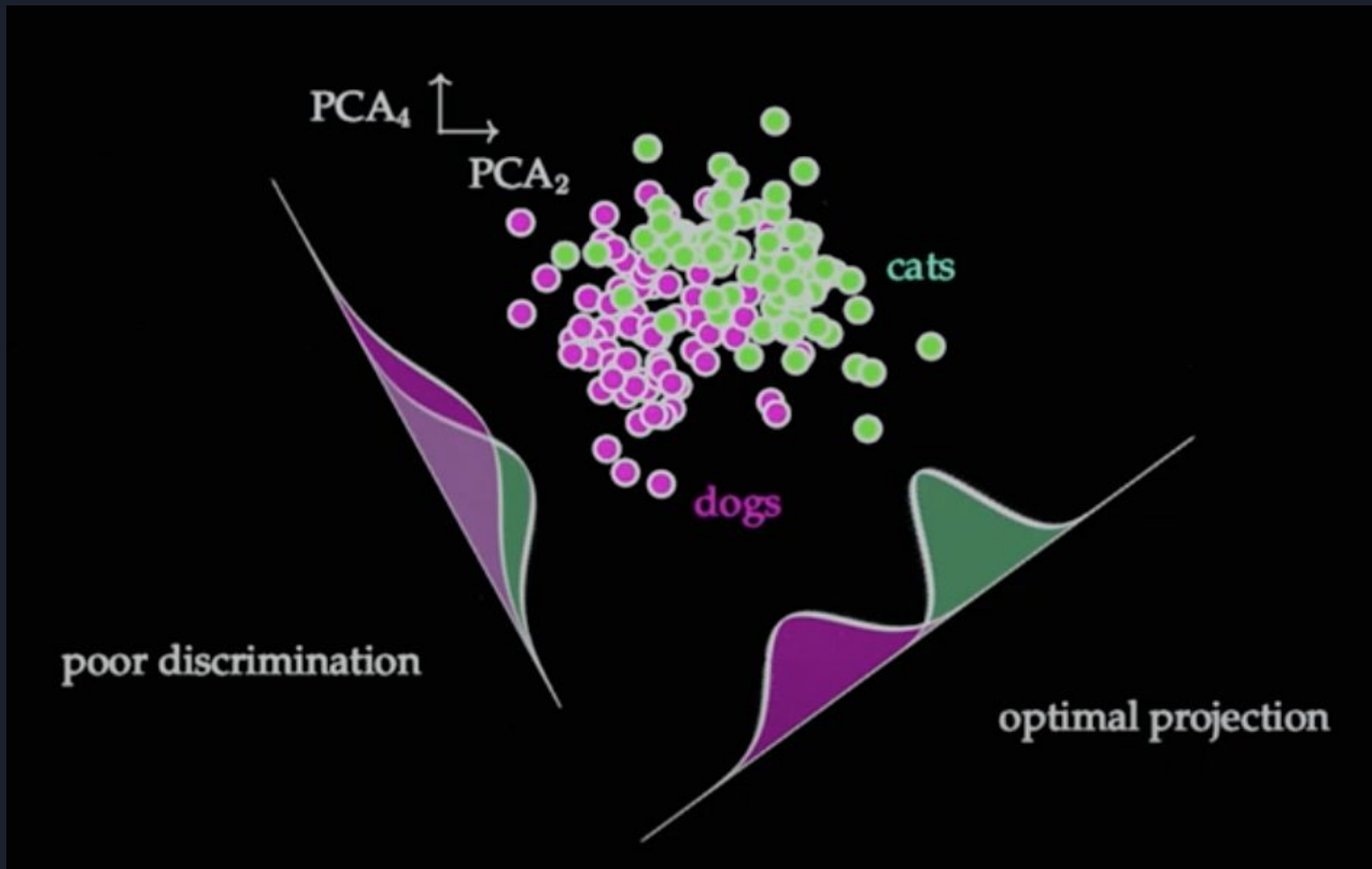
maximise

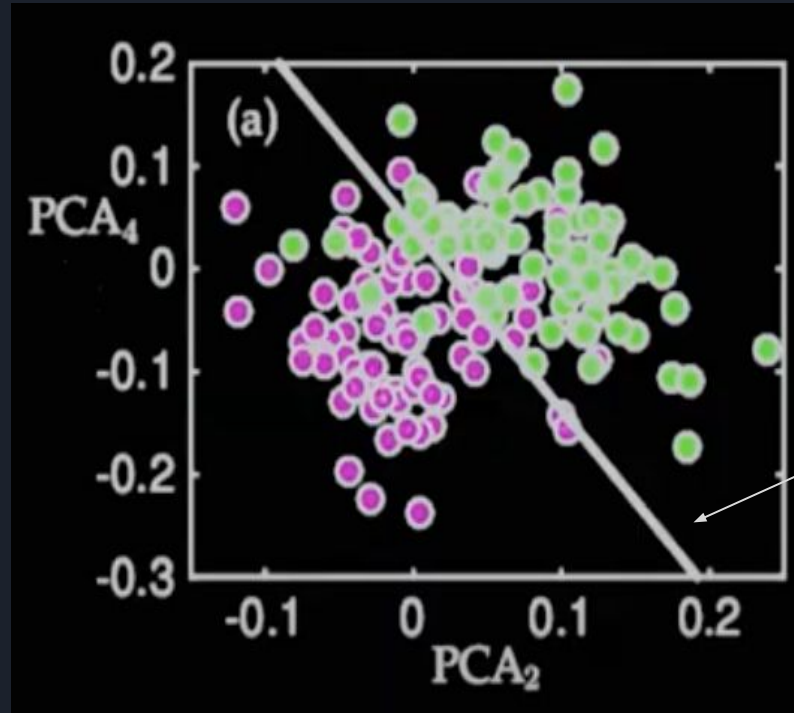
between- and in-class variance

minimise

$$\mathbf{S}_B = (\mu_2 - \mu_1)(\mu_2 - \mu_1)^T$$

$$\mathbf{S}_W = \sum_{j=1}^2 \sum_{\mathbf{x}} (\mathbf{x} - \mu_j)(\mathbf{x} - \mu_j)^T$$





Linear Discriminant Line



Linear Discriminant Analysis (LDA)

- LDA is like PCA, they both reduce the dimensions, however LDA focuses on maximising the separability among known categories