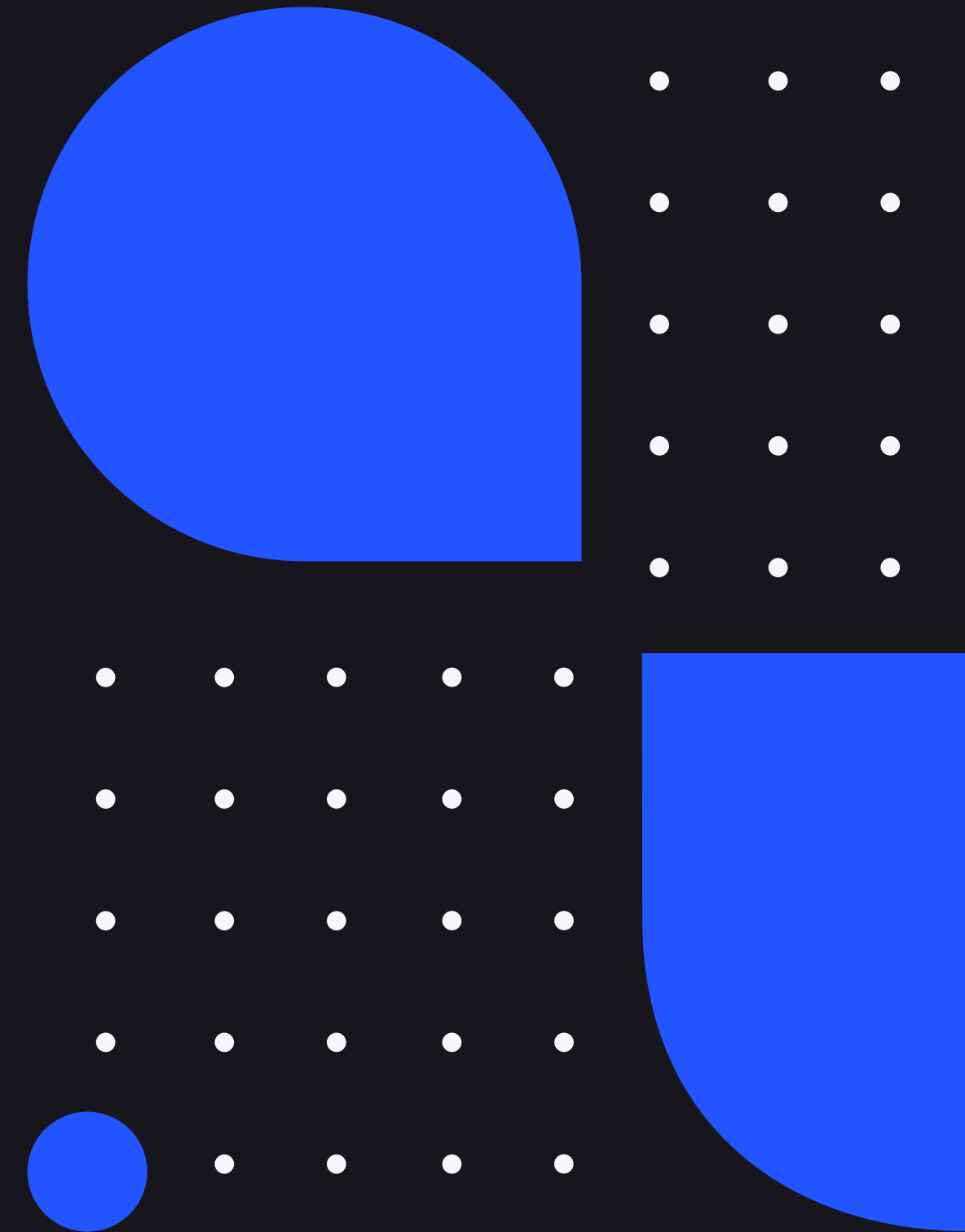


Análise Discriminante Linear

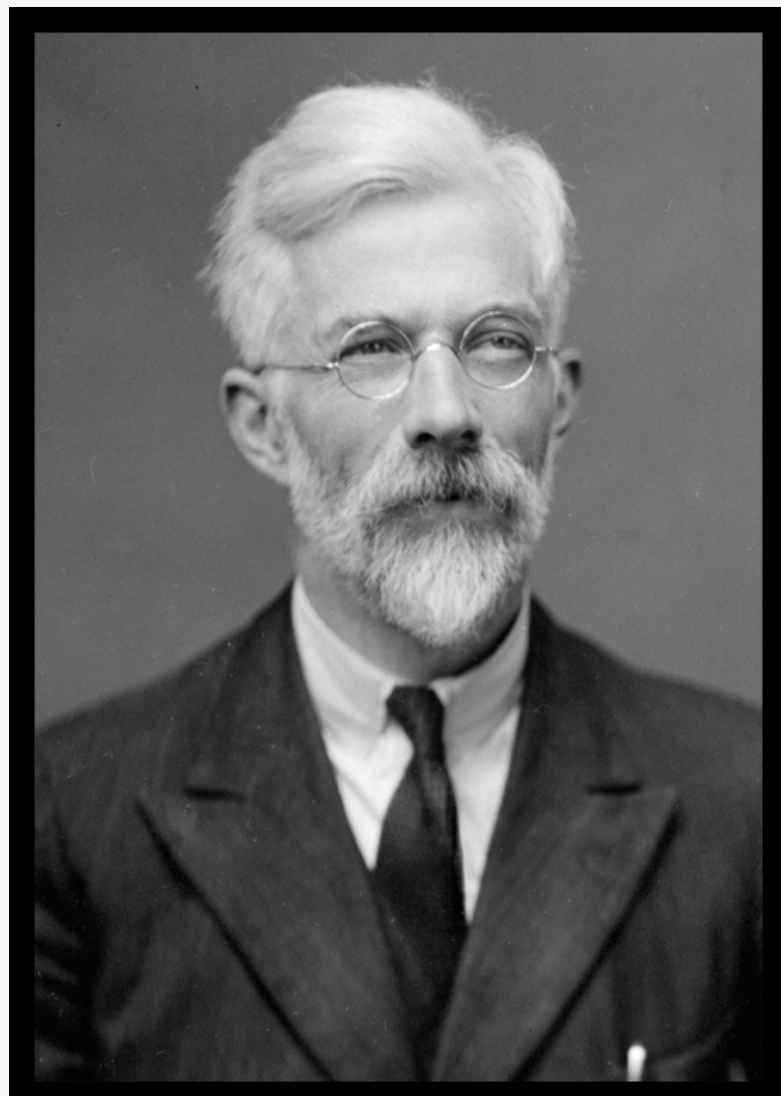
João Pedro Fontoura da Silva
Alícia Isaías Macedo

LAMFO/UnB



1. Introdução

- A Análise Discriminante Linear (LDA) foi introduzida em sua forma inicial por **Robert Fisher** em 1936
- É um método utilizado para classificação, redução de dimensão e visualização de dados
- Tem aplicações em diversos campos do conhecimento (finanças, biologia, tecnologia), sendo particularmente útil com amostras de dados multivariados



2. Classificação

Objetivo: alocar uma variável \mathbf{X} a uma de K classes

- Regra de máxima verossimilhança vs. Regra de Bayes

- **RMV**: cada classe j pode ocorrer com a mesma probabilidade

$$j = \arg \max_i f_i(\mathbf{x})$$

- **RB**: já temos as probabilidades π_1, \dots, π_K , logo alocamos para

$$j = \arg \max \pi_i f_i(\mathbf{x})$$



● **Análise discriminante linear vs. quadrática**

Assumindo uma distribuição Gaussiana multivariada, $\mathbf{X} \sim N(\mu, \Sigma)$ e seguindo a regra Bayesiana, podemos classificar \mathbf{x} na classe j que maximiza a função discriminatória

$$\delta_i(\mathbf{x}) = \log f_i(\mathbf{x}) + \log \pi_i$$

- **LDA**: assumimos igual covariância entre as classes; a fronteira de decisão é linear:

$$\delta_k(\mathbf{x}) = \mathbf{x}^T \Sigma^{-1} \mu_k - (1/2) \mu_k^T \Sigma^{-1} \mu_k + \log \pi_k$$

- **QDA**: a função discriminante é uma função quadrática:

$$\delta_k(\mathbf{x}) = -(1/2) \log |\Sigma_k| - (1/2) (\mathbf{x} - \mu_k)^T \Sigma^{-1} (\mathbf{x} - \mu_k) + \log \pi_k$$

● Simplificando o LDA



Diagonalizamos a matriz de covariância, transformando os dados de modo a obter uma matriz identidade de covariância. Em seguida, "esfericizamos" os dados (\mathbf{x}^*) e obtemos as médias das classes no novo espaço. Assim, classificamos \mathbf{x} por:

$$\delta_k(\mathbf{x}^*) = \mathbf{x}^{*T} \hat{\mu}_k - (1/2) \hat{\mu}_k^T \hat{\mu}_k + \log \hat{\pi}_k$$

Agora suponha que existam duas classes, k e l . A regra de classificação é:

$$\delta_k(\mathbf{x}^*) - \delta_l(\mathbf{x}^*) > 0$$

● Simplificando o LDA



Da regra anterior:

$$\delta_k(\mathbf{x}^*) - \delta_l(\mathbf{x}^*) > 0$$

Desenvolvendo:

$$\mathbf{x}^{*T} (\hat{\mu}_k - \hat{\mu}_l) > (1/2) (\hat{\mu}_k + \hat{\mu}_l)^T (\hat{\mu}_k - \hat{\mu}_l) - \log \hat{\pi}_k / \hat{\pi}_l$$

Na equação à esquerda, temos o comprimento da projeção ortogonal de \mathbf{x}^* na linha de segmento que une as médias das classes. No lado direito, temos a locação do centro do segmento corrigido pelas probabilidades das classes.

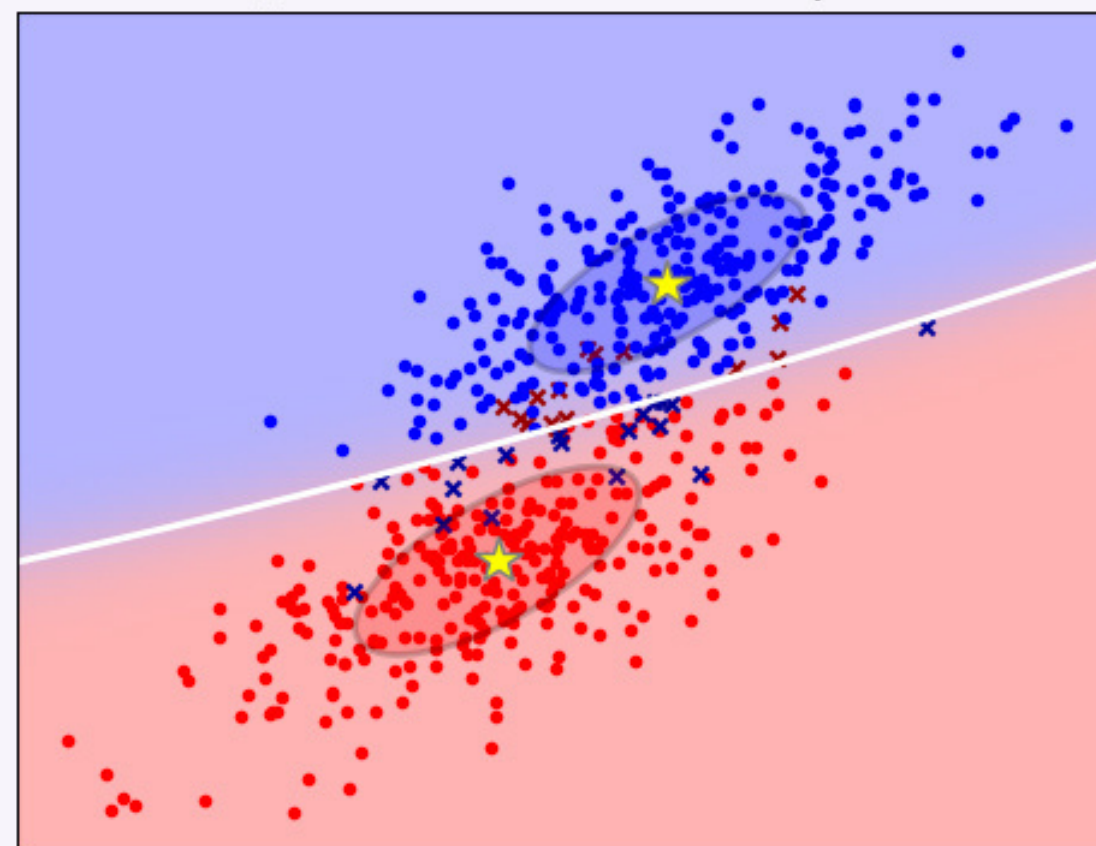
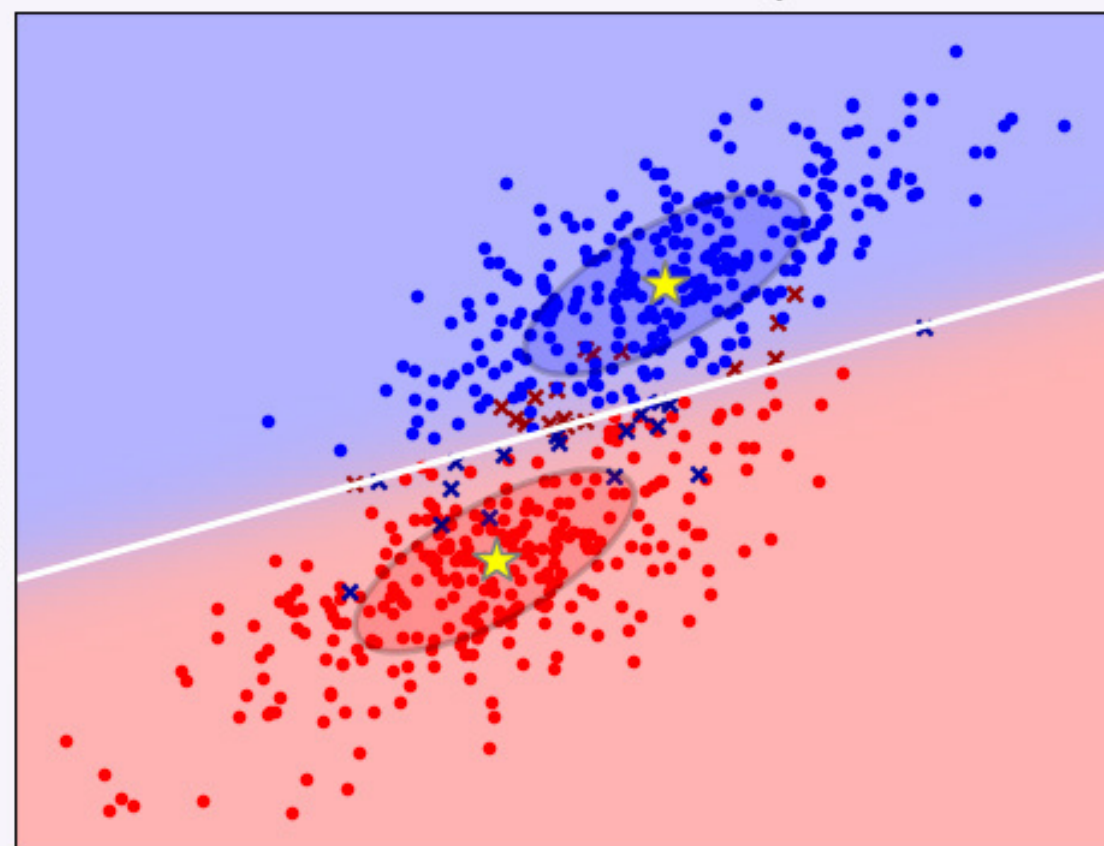
Essencialmente, **o LDA compara distâncias** entre as médias dos dados e as médias das classes, e os classifica junto àquela que apresenta **a média mais próxima**.

Linear Discriminant Analysis vs Quadratic Discriminant Analysis

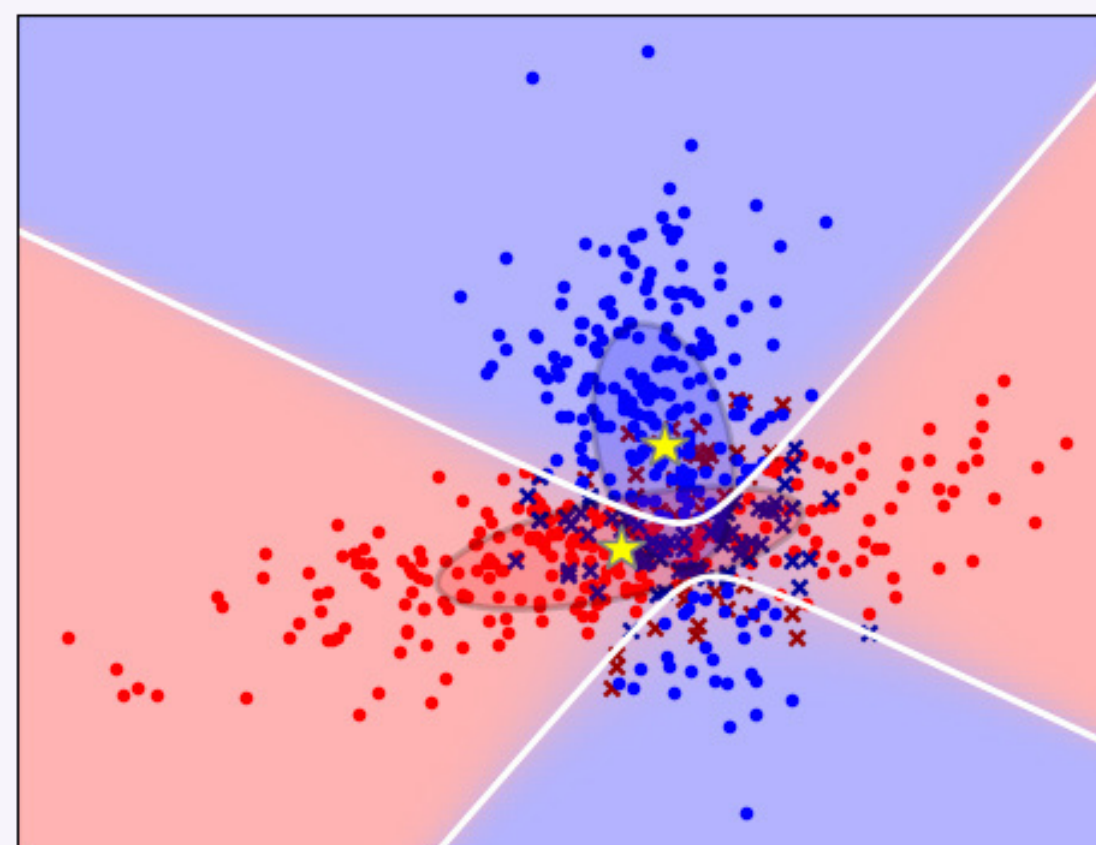
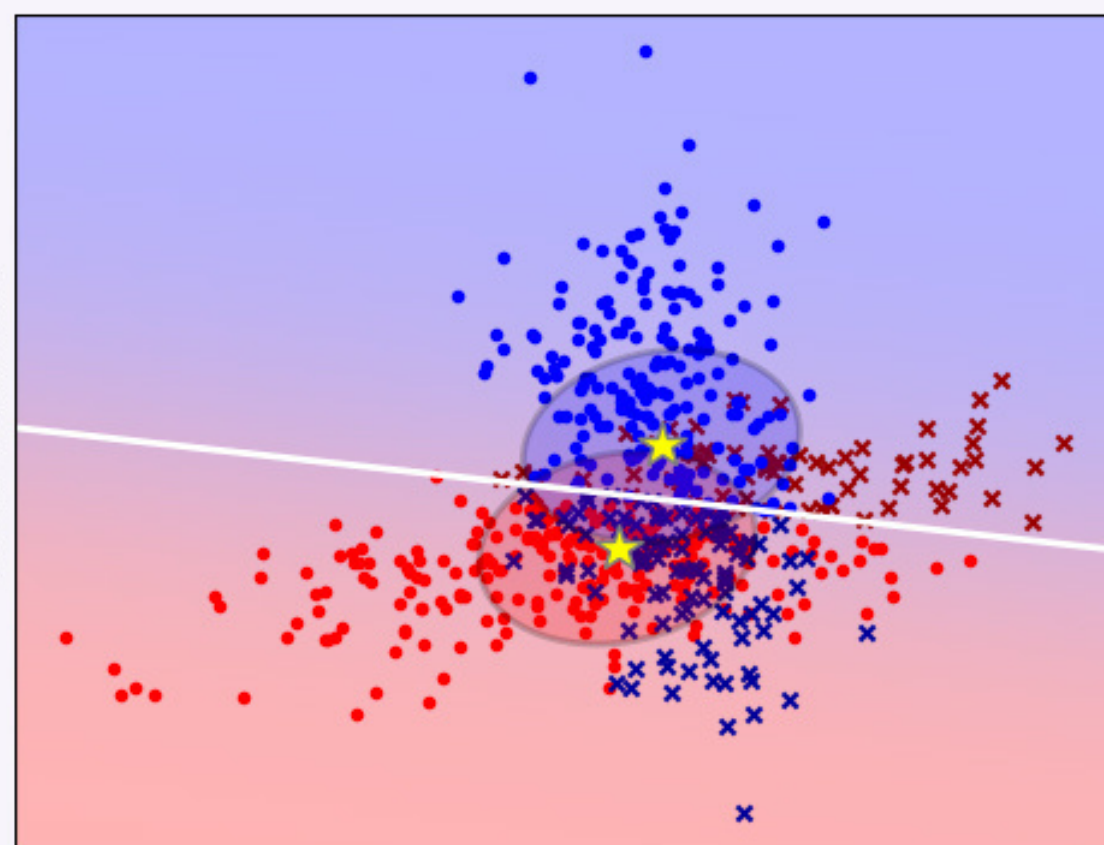
Linear Discriminant Analysis

Quadratic Discriminant Analysis

Data with
fixed covariance




Data with
varying covariances



Fonte: scikit-learn.org



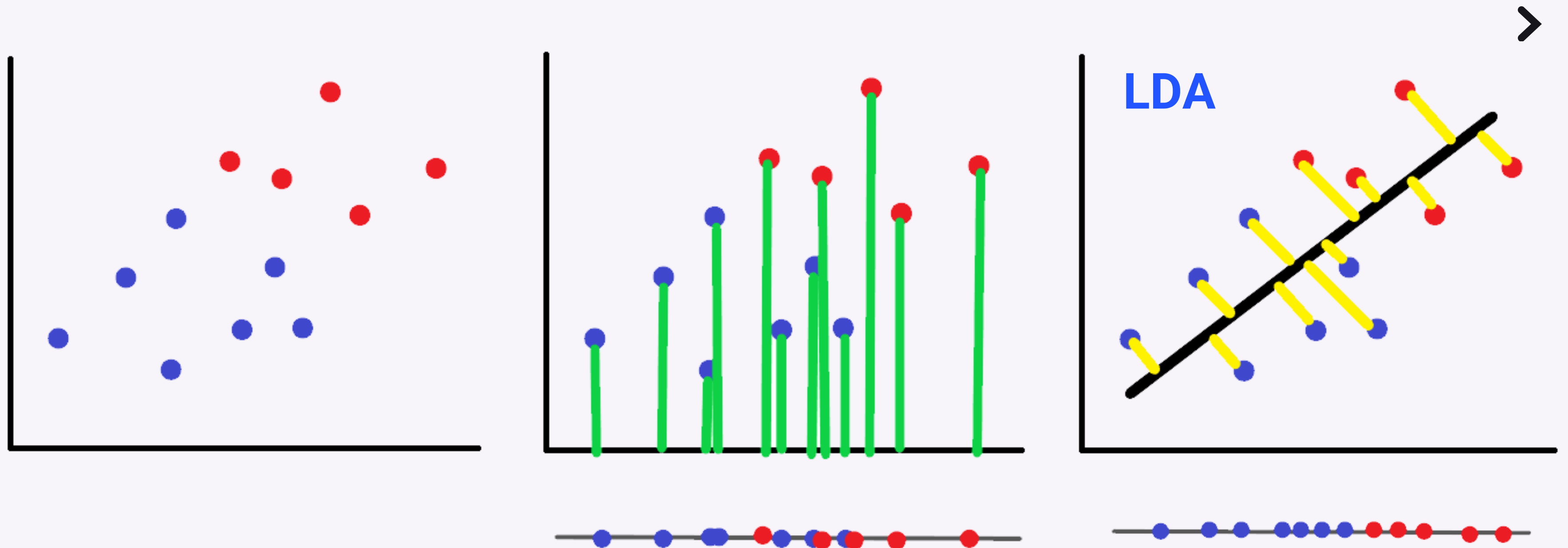


3. Redução de dimensão

Objetivo: facilitar a visualização e tratamento de dados com várias dimensões

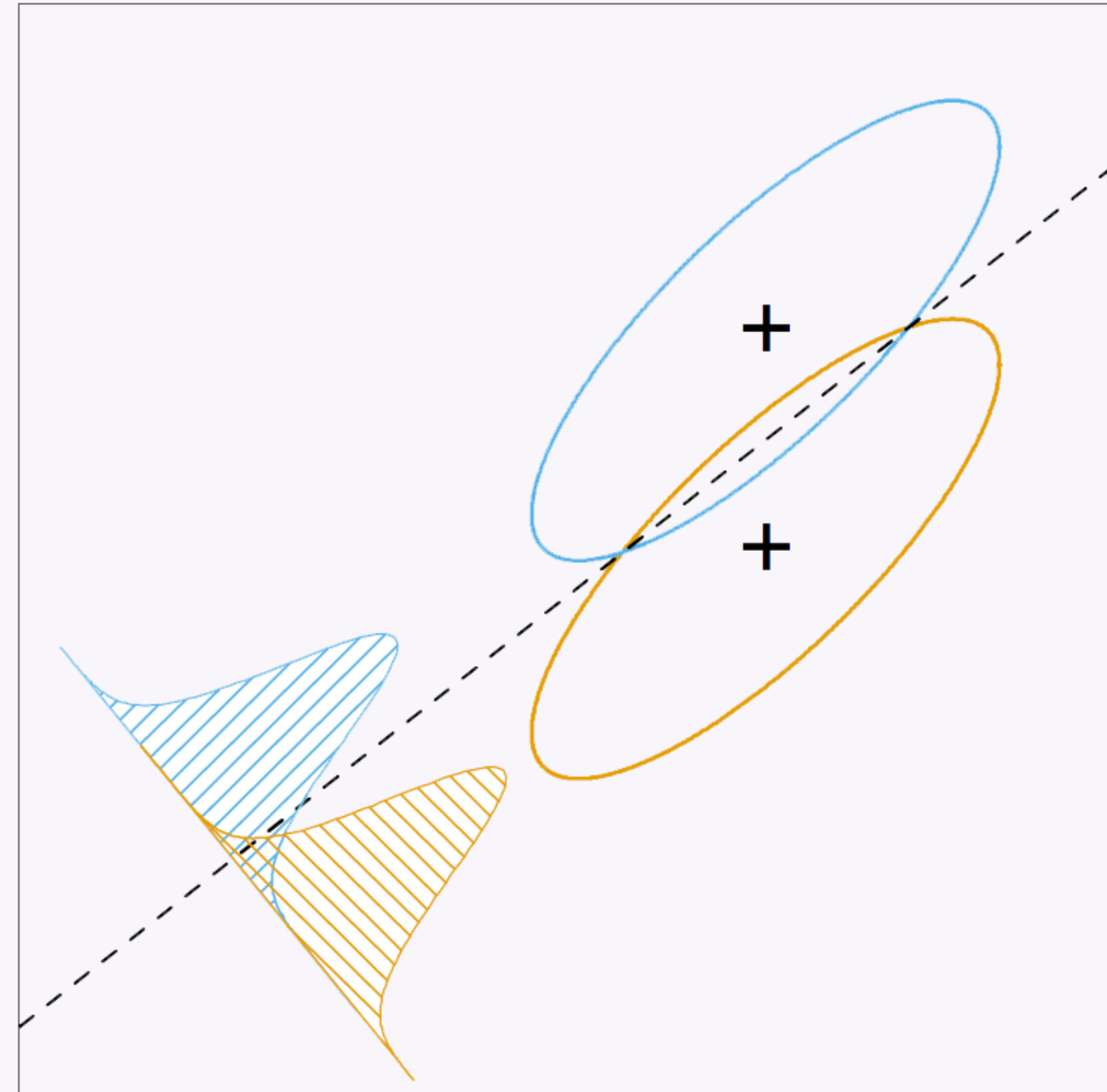
- Reduzimos as dimensões do problema ao projetar os dados ortogonalmente no subespaço dado pelas médias das classes
- Encontramos um subespaço ótimo quando a média das classes dos dados esfericizados têm máxima separação, em termos de variância





O **LDA** cria um novo eixo e projeta os dados ortogonalmente de forma a **minimizar a variância** e **maximizar a distância entre as médias das classes**

Fonte: yangxiao Zhou.github.io



O que se espera é que a **variância entre classes** seja maximizada relativamente à **variância intra-classe**.



4. Avaliação

- **Vantagens:** método simples de classificação, ferramenta robusta e de fácil interpretação, visualização e modelagem
- **Desvantagem:** linearidade pode não ser adequada para determinadas amostras de dados



Referências



- Linear and Quadratic **Discriminant Analysis**. **Scikit Learn**. Disponível em: <https://scikit-learn.org/stable/modules/lda_qda.html>
- **Brownlee, Jason**
Linear Discriminant Analysis for Machine Learning. **Machine Learning Mastery**, 2016. Disponível em: <<https://machinelearningmastery.com/linear-discriminant-analysis-for-machine-learning/>>
- **Sawla, Srishti**
Linear Discriminant Analysis. Medium Data Sciene, 2018. Disponpivel em: <<https://medium.com/@srishtisawla/linear-discriminant-analysis-d38decf48105>>
- **Schlagenhauf, Tobias**
Linear Discriminant Analysis (LDA). Python Machine Learning Tutorial. Disponível em: <https://www.python-course.eu/linear_discriminant_analysis.php>
- **Xiaozhou, Yang**
Linear Discriminant Analysys, Explained. **Xiaozhou's Notes**, 2019. Disponível em: <<https://yangxiao Zhou.github.io/data/2019/10/02/linear-discriminant-analysis.html>>



Obrigado!