

Redução de dimensionalidade

Jones Granatyr



PCA (Principal Component Analysis)

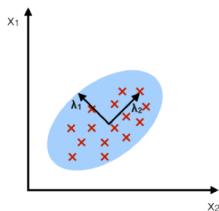
- Seleção de características x Extração de características
- Um dos principais algoritmos de aprendizagem de máquina não supervisionada
- Identifica a correlação entre variáveis, e caso haja uma forte correlação é possível reduzir a dimensionalidade
- Das **m** variáveis independentes, PCA extrai **$p \leq m$** novas variáveis independentes que explica melhor a variação na base de dados, sem considerar a variável dependente
- O usuário pode escolher o número de **p**

LDA (Linear Discriminant Analysis)

- Além de encontrar os componentes principais, LDA também encontra os eixos que maximizam a separação entre múltiplas classes
- É supervisionado por causa da relação com a classe
- Das m variáveis independentes, LDA extrai $p \leq m$ novas variáveis independentes que mais separam as classes da variável dependente

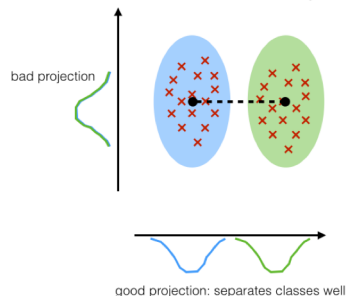
PCA:

component axes that maximize the variance



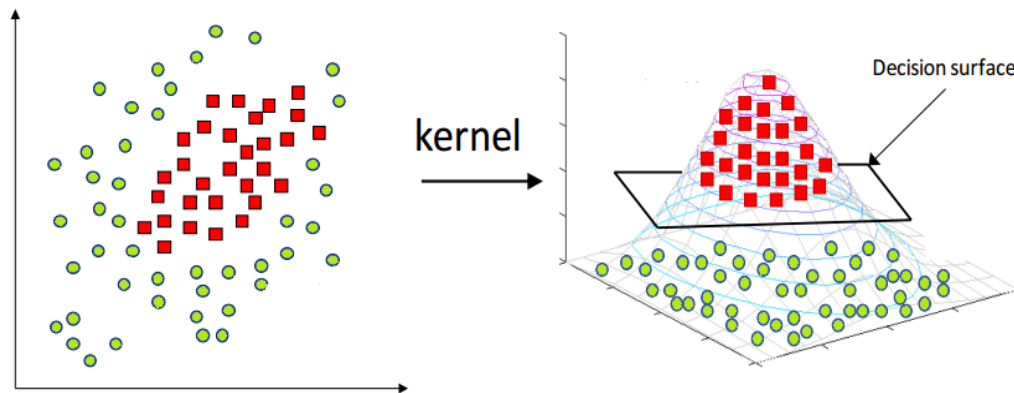
LDA:

maximizing the component axes for class-separation



Kernel PCA

- PCA e LDA são utilizados quando os dados são linearmente separáveis
- Kernel PCA é uma versão do PCA que os dados são mapeados para uma dimensão maior usando o ***kernel trick***
- Os componentes principais são extraídos dos dados com dimensionalidade maior



Conclusão

