

Redução de Dimensionalidade – PCA e LDA

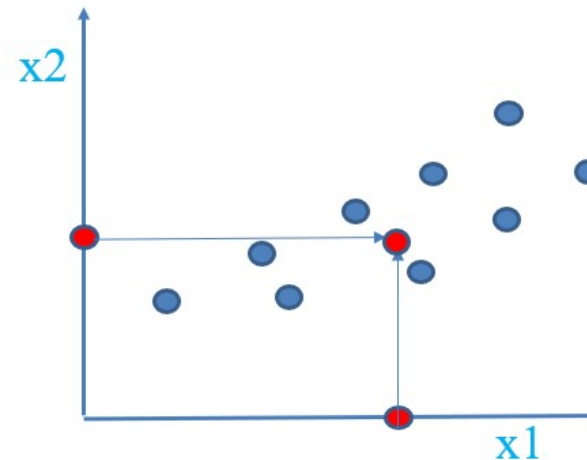
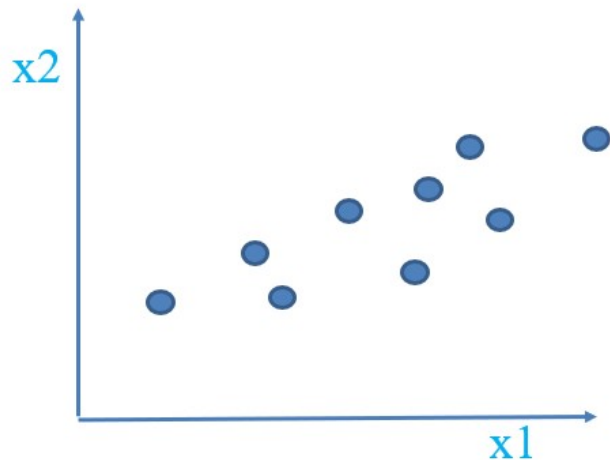
Prof. Gustavo Willam Pereira



INSTITUTO FEDERAL
Sudeste de Minas Gerais

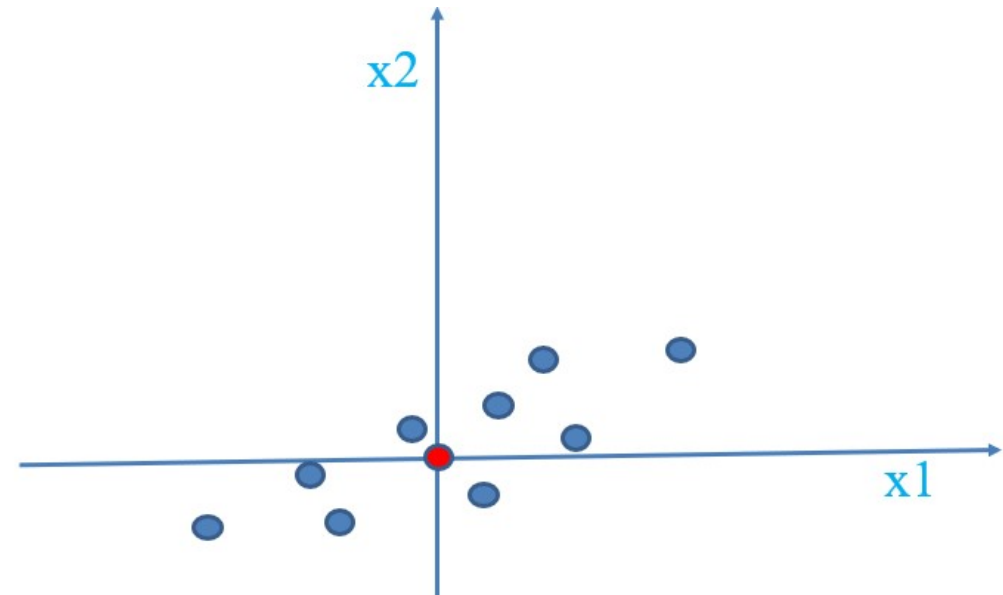
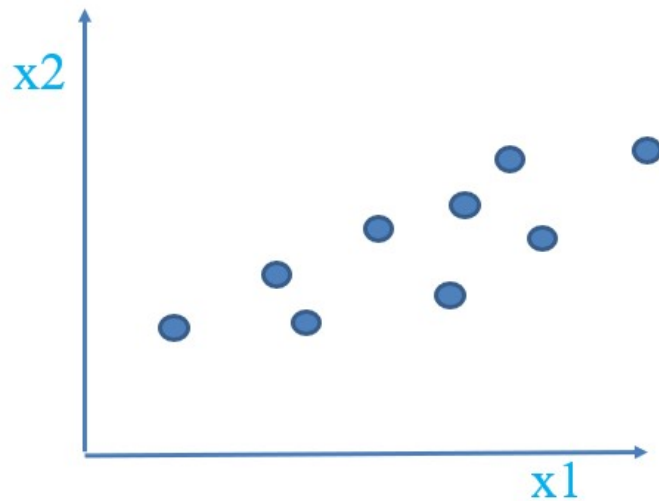
PCA - Análise de Componentes Principais

- Considere um modelo de ML com 2 variáveis (X_1 , X_2).
- Considere agora a média desses 2 valores de X_1 e X_2 . E plotando o ponto médio no gráfico.



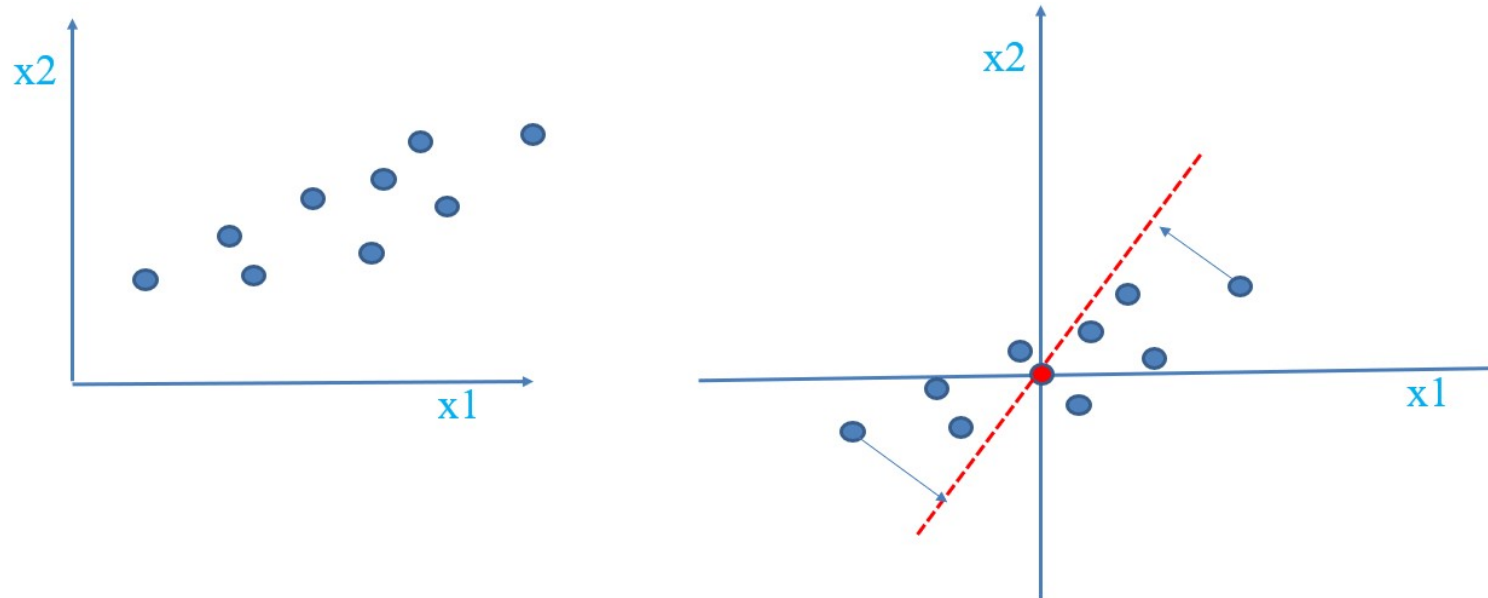
PCA - Análise de Componentes Principais

- Considere agora deslocar esses dados para a origem considerando média zero para X_1 e X_2



PCA - Análise de Componentes Principais

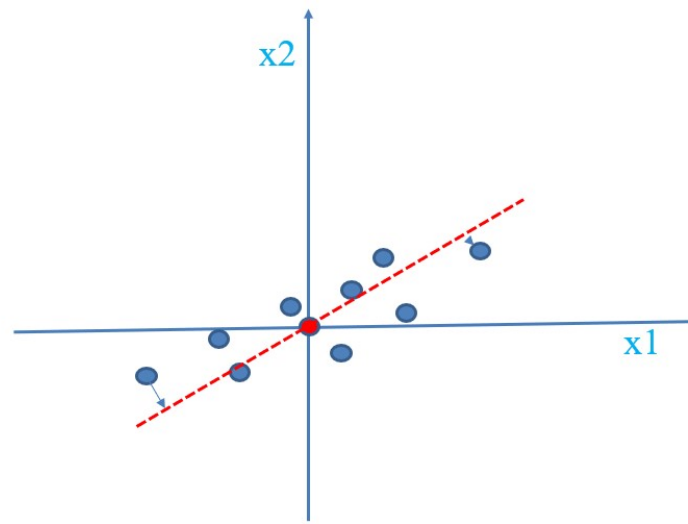
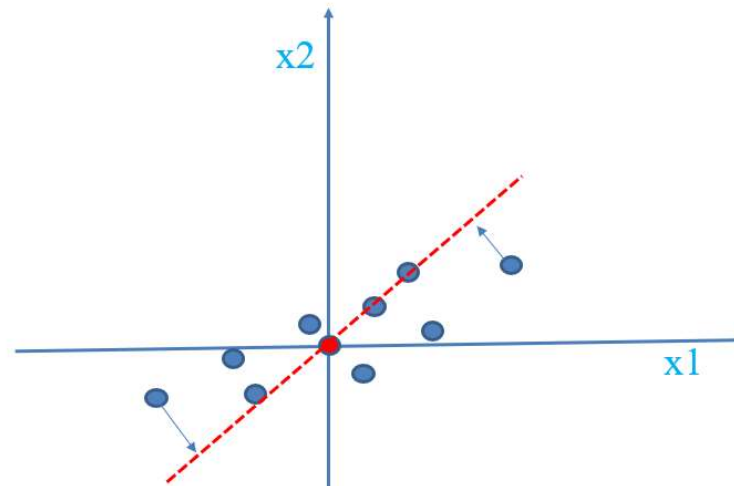
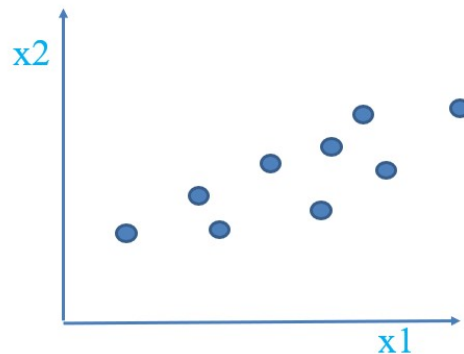
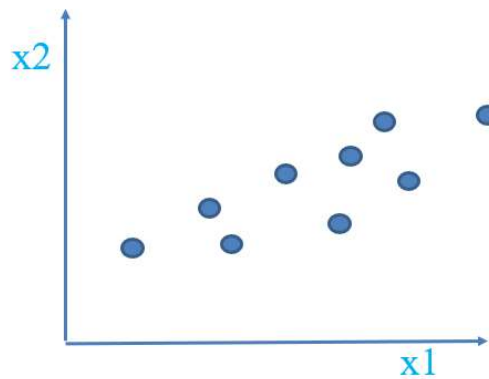
- Vamos criar um modelo linear passando entre os pontos do gráfico.



- Projetando as distâncias de todos os pontos até a reta poderíamos obter os desvios (erros).

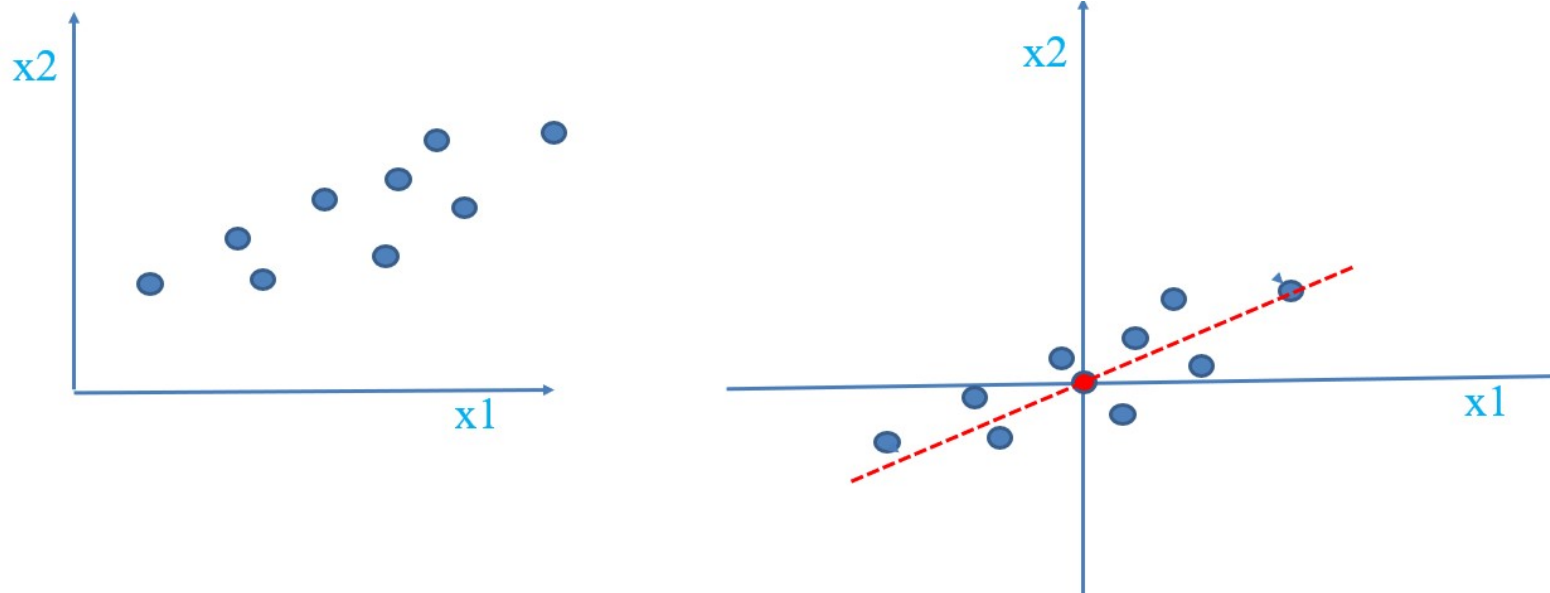
PCA - Análise de Componentes Principais

- Minimizando as distâncias entre os pontos do gráfico e a reta teríamos.



PCA - Análise de Componentes Principais

- Minimizando as distâncias entre os pontos do gráfico e a reta teríamos.

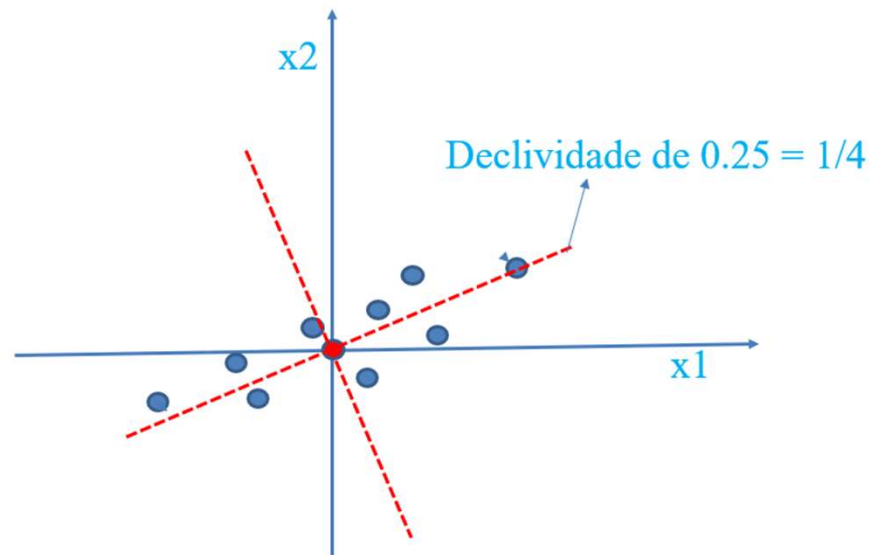
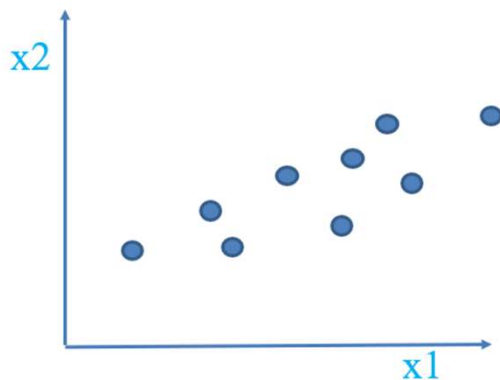


- Minimizando as distâncias até encontrar a distância mínima para os pontos.

PCA - Análise de Componentes Principais

- Imagine que a declividade seja de 0,25.
- Ou seja para cada unidade deslocada na X1 eu teria que deslocar 0,25 unidades na X2.
- Poderíamos ter uma nova variável PC1, que é uma combinação linear entre X1 e X2

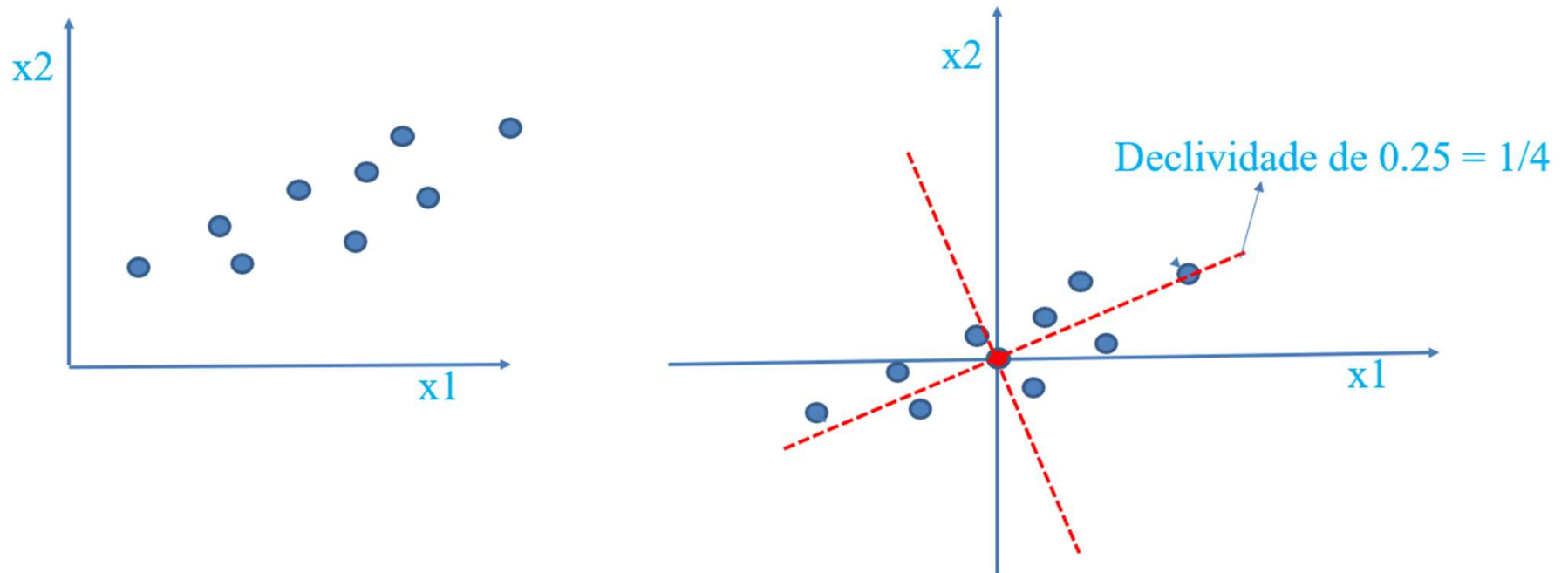
$$Pc1 = 4.x1 + 1.x2 \text{ (combinação linear)}$$



PCA - Análise de Componentes Principais

- Veja que PC1 explica uma determinada variação do modelo.
- Poderíamos ter um outro modelo linear (uma reta perpendicular a PC1), que explicaria uma certa variabilidade entre as variáveis X_1 e X_2 .

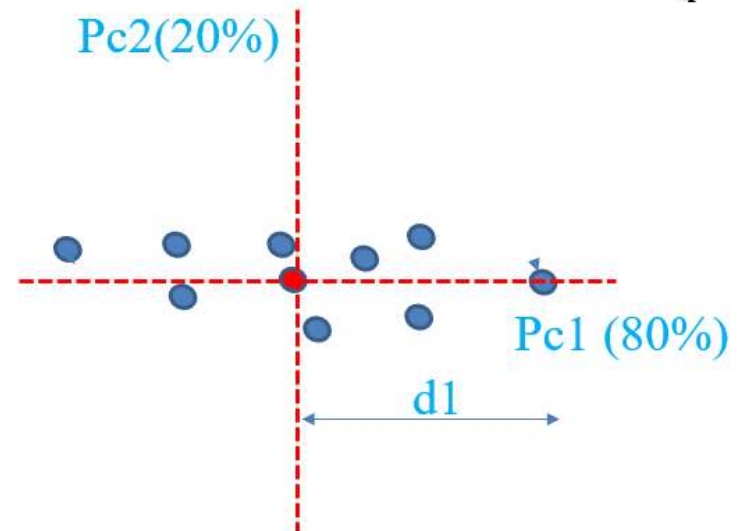
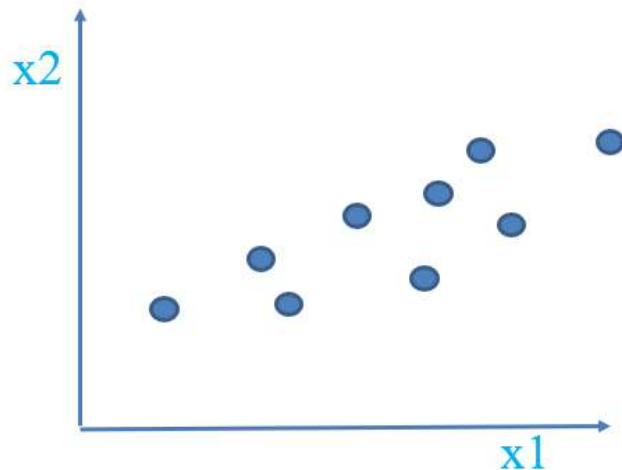
$$Pc1 = 4.x1 + 1.x2 \text{ (combinação linear)}$$



PCA - Análise de Componentes Principais

- Poderíamos eliminar X_1 e X_2 e plotar o gráfico em função de PC1 e PC2.

$$s^2 = \frac{\sum_{i=1}^N (X - \bar{X})^2}{N - 1}$$
$$s^2(pc1) = \frac{d1^2 + d2^2 \dots}{n - 1} = 20$$
$$s^2(pc2) = 5$$

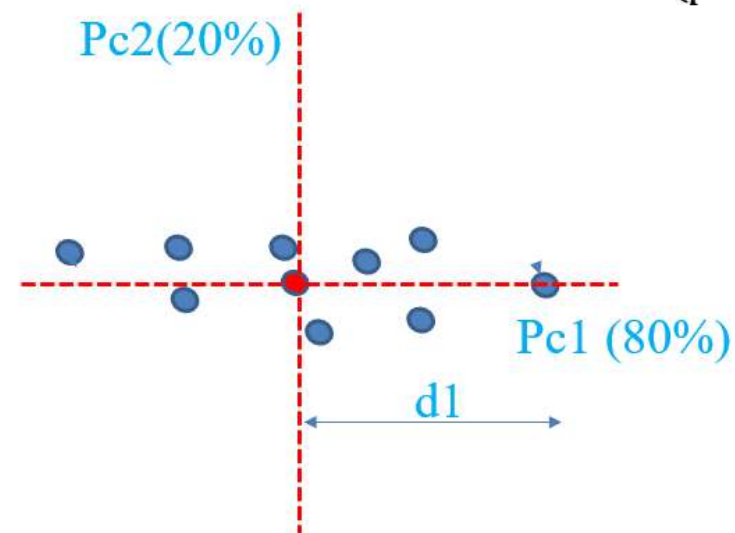
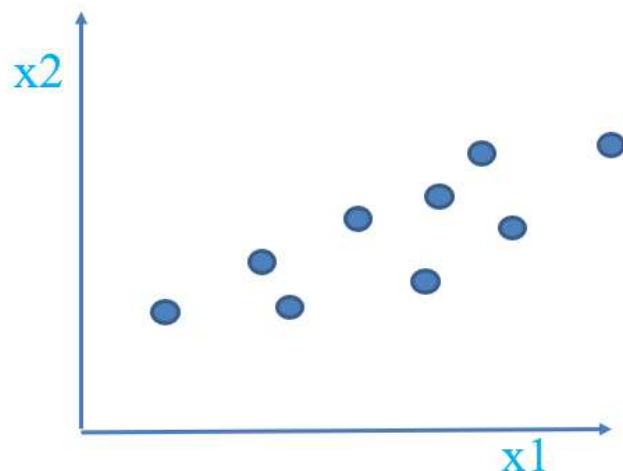


PCA - Análise de Componentes Principais

- Veja que 80% da variabilidade dos dados estão sendo explicados pela PC1 e 20% pela PC2.
- Para calcular essa variabilidade utilizando a variância (Est. I).
- Como a média é zero (ponto vermelho), poderíamos eliminar \bar{X}

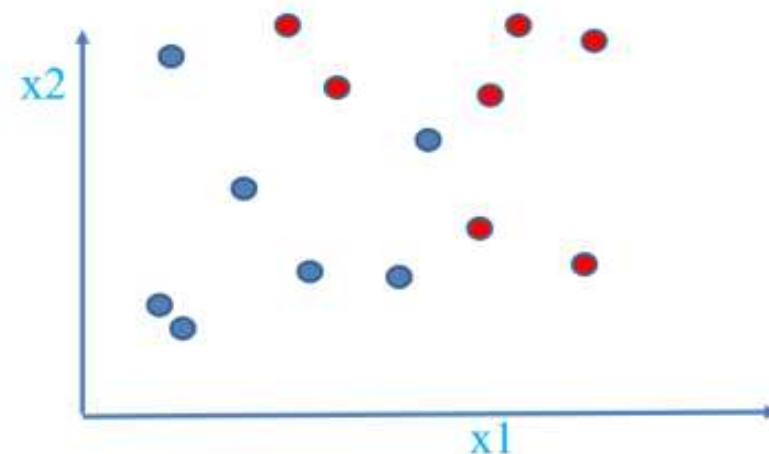
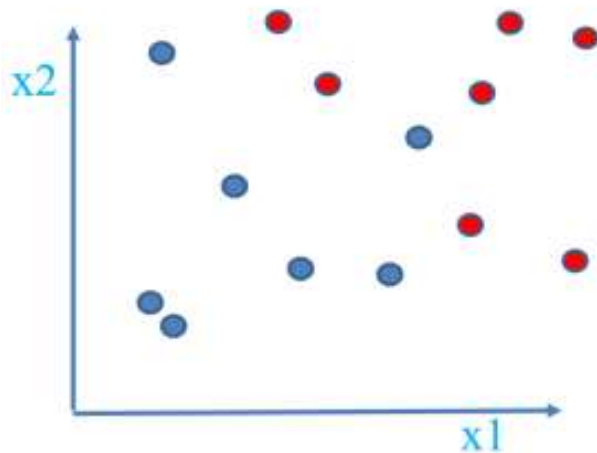
$$s^2 = \frac{\sum_{i=1}^N (X - \bar{X})^2}{N - 1}$$
$$s^2(pc1) = \frac{d1^2 + d2^2 \dots}{n - 1} = 20$$

$$s^2(pc2) = 5$$



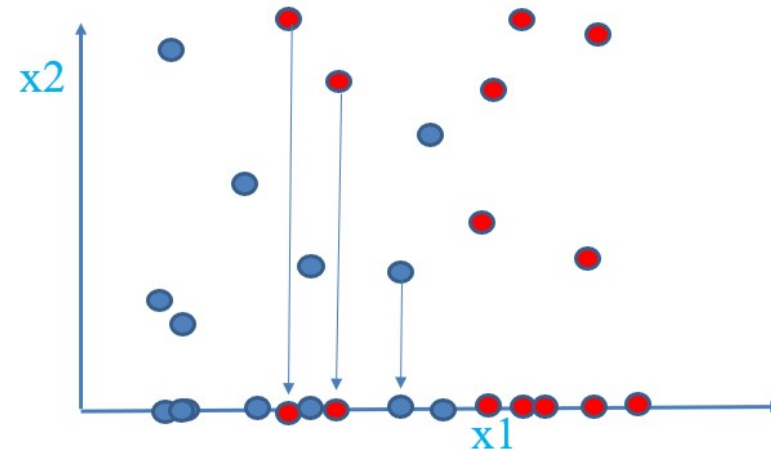
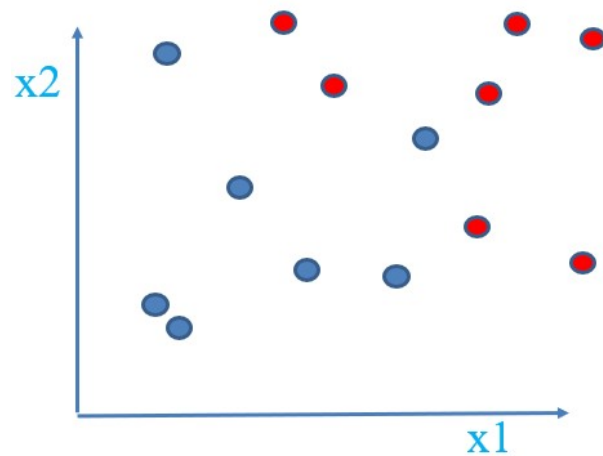
LDA – Linear Discriminante Análise

- Semelhante a PCA, entretanto na LDA temos conhecimento das classes (Y), enquanto que na PCA não temos conhecimento das classes. Então a PCA é não supervisionado enquanto que a LDA é supervisionado.



LDA – Linear Discriminante Análise

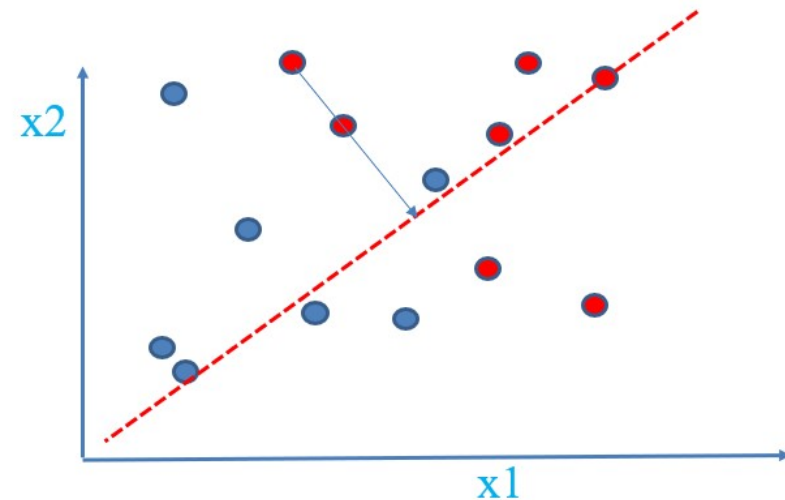
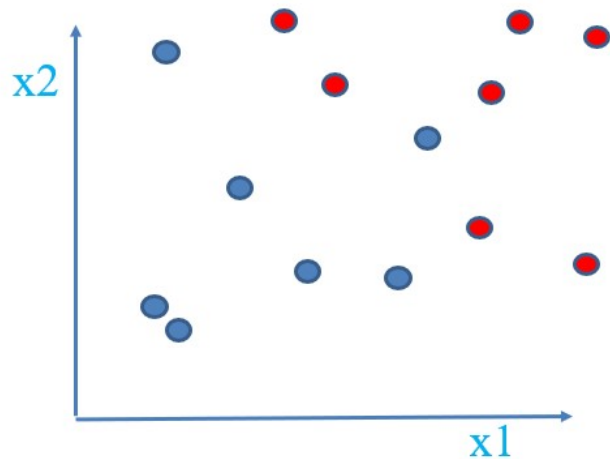
- Projetando todos os pontos na X_1 teríamos.



- E se quiséssemos separar (discriminar) estes dados ? Separando as classes vermelhas das azuis.
- Escolhendo um limiar, teríamos uma taxa de erro, pois não seria possível separar todas as vermelhas das azuis.

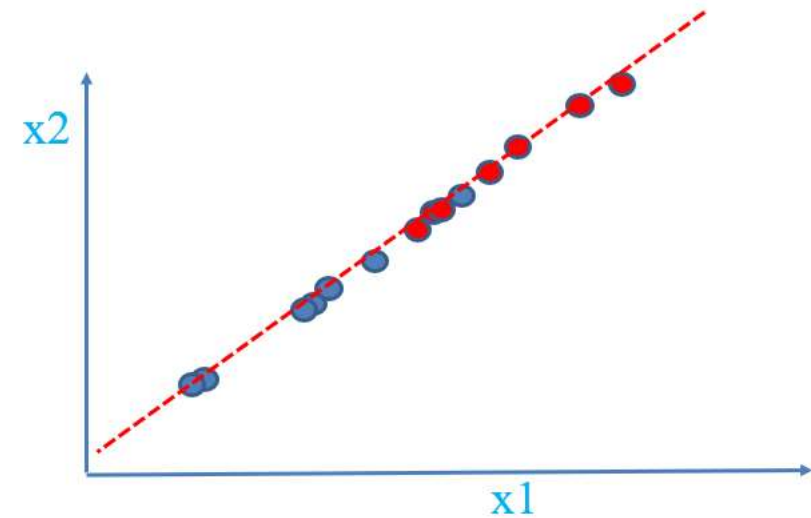
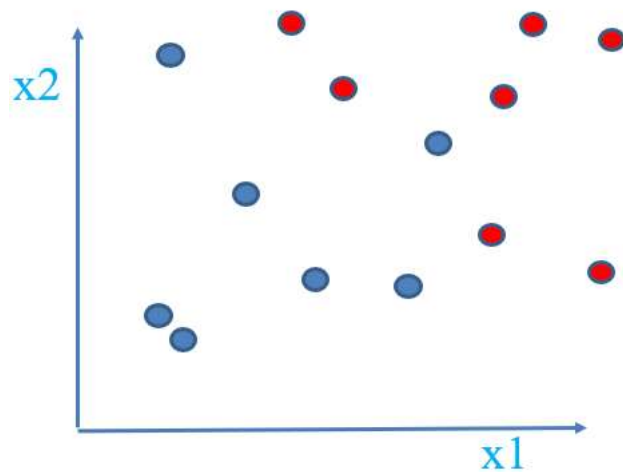
LDA – Linear Discriminante Análise

- Mas se fizéssemos diferente, traçando um modelo linear (combinação linear) como o objetivo de maximizar a discriminação das classes ?



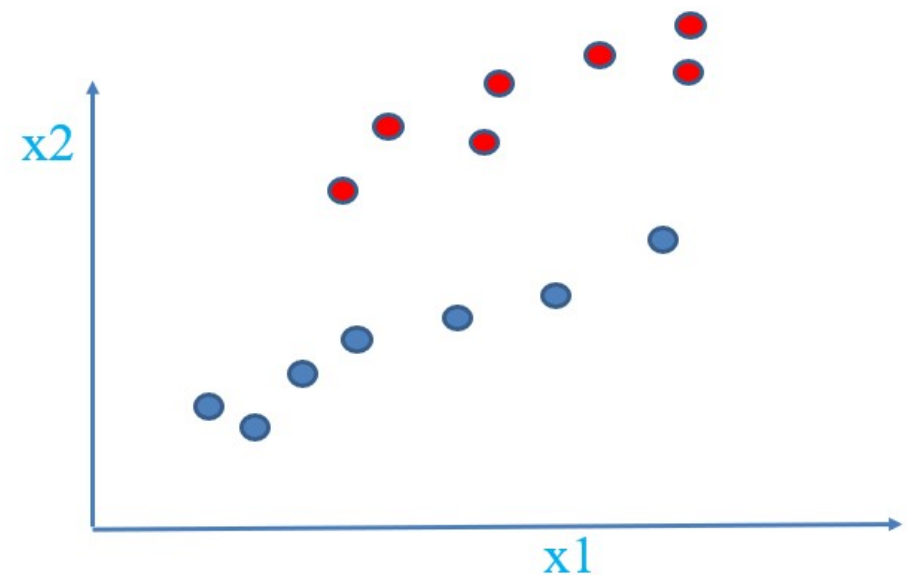
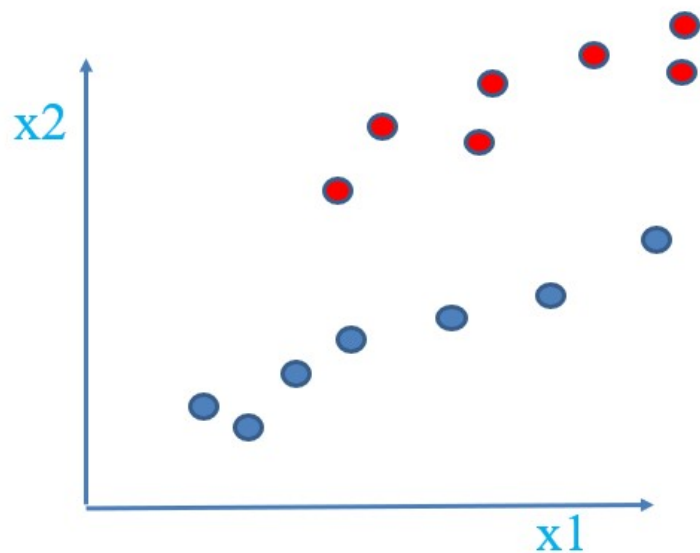
LDA – Linear Discriminante Análise

- Veja que agora teríamos apenas um erro ao discriminar as classes.



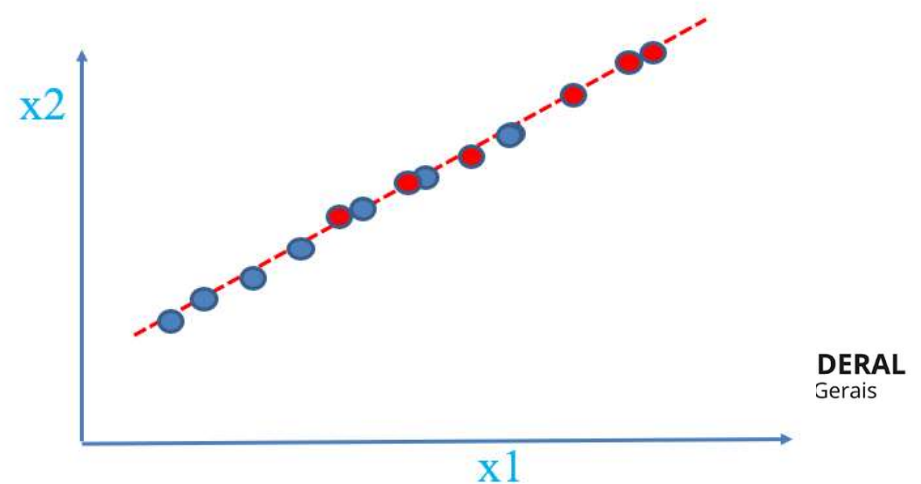
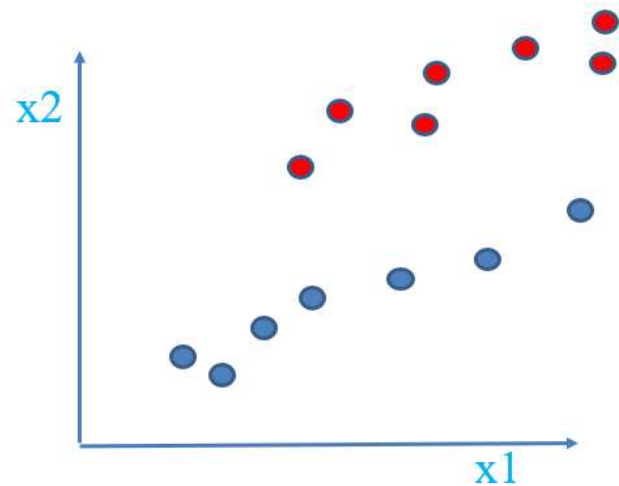
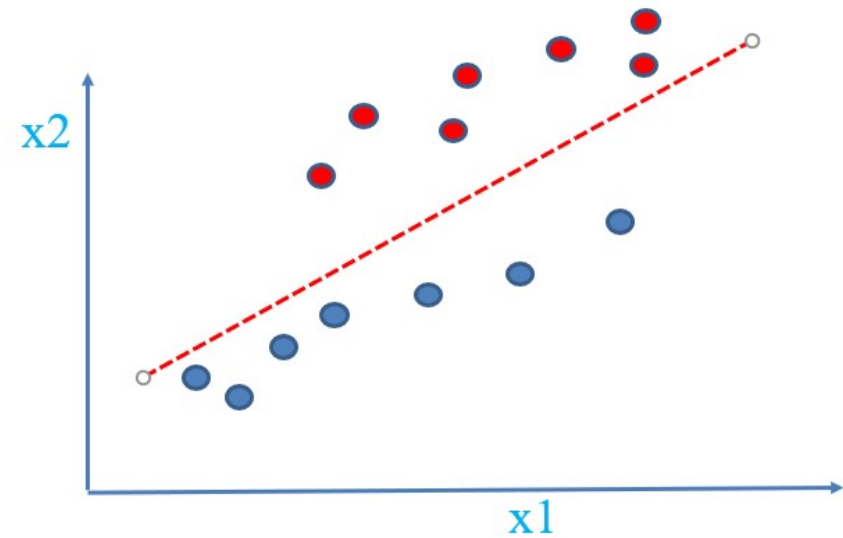
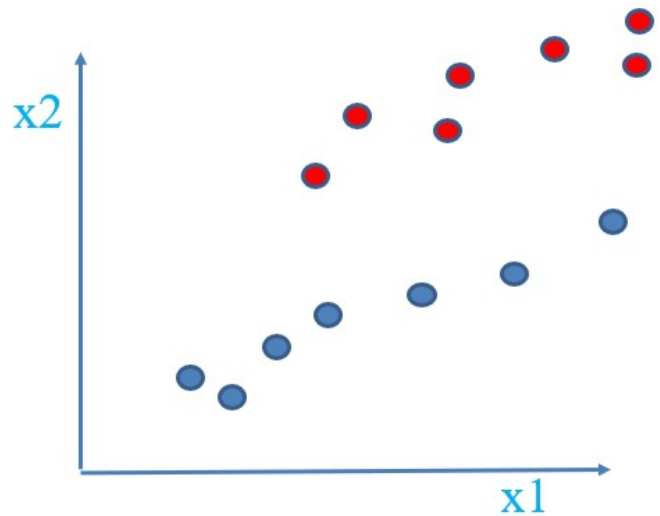
LDA – Linear Discriminante Análise

- Agora veja esses outro modelo.
- Como você definiria um modelo para discriminar as classes ?



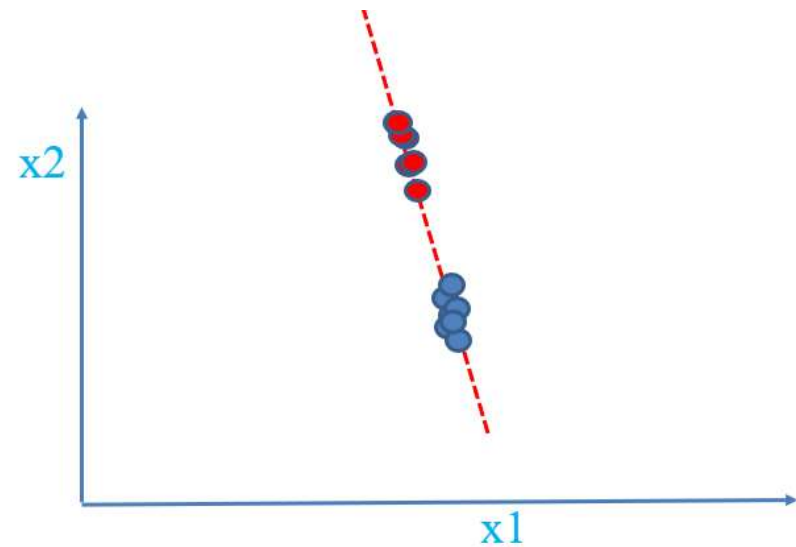
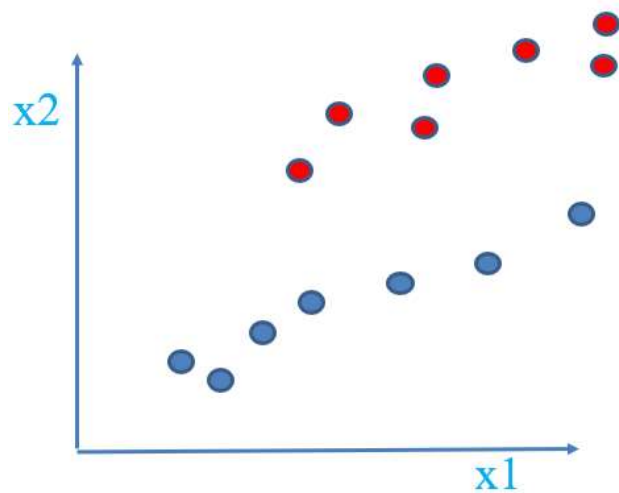
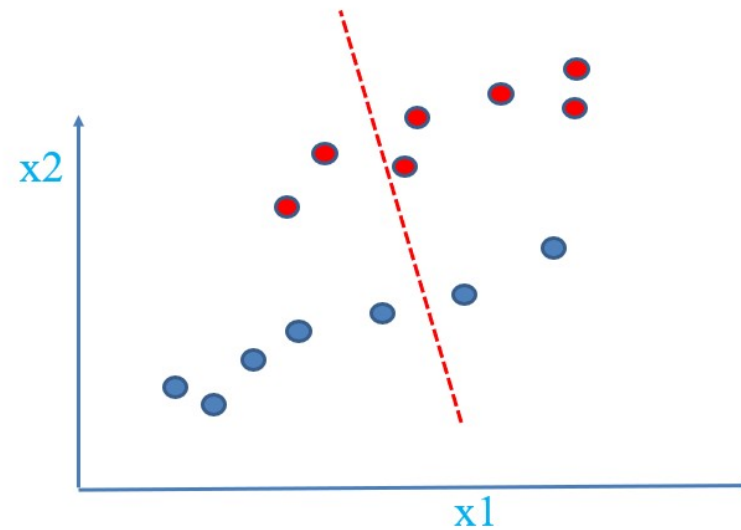
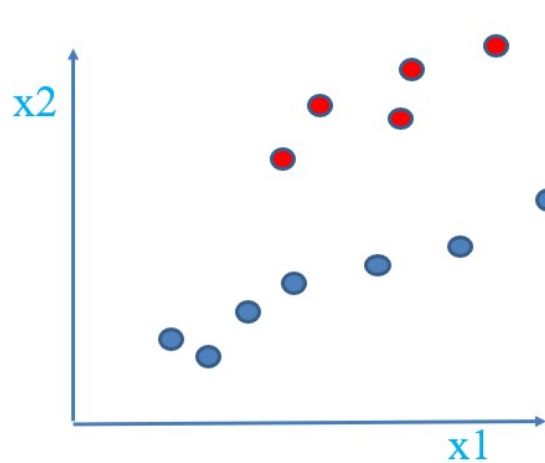
LDA – Linear Discriminante Análise

- Traçando uma reta nessa direção



LDA – Linear Discriminante Análise

- Traçando uma reta nessa outra direção



PCA e LDA - Conclusões

- A PCA tem como objetivo gerar uma combinação linear entre as variáveis que vá em direção a máxima variância dos dados
- A LDA tem como objetivo gerar uma combinação linear entre as variáveis que permite ter a máxima separação (discriminar) entre as classes.
- Vamos agora ver um código Python para PCA e LDA



INSTITUTO FEDERAL
Sudeste de Minas Gerais