

Inferencia Estadística y Reconocimiento de Patrones

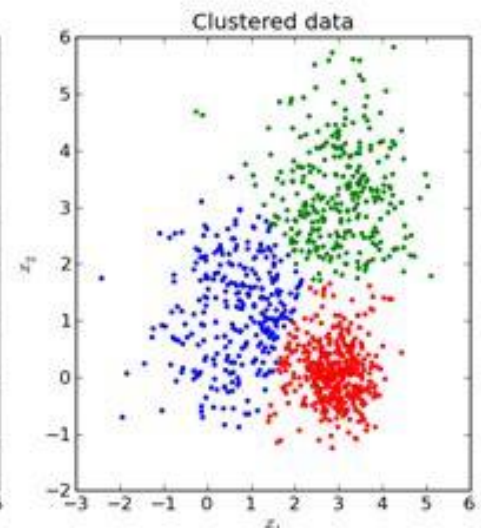
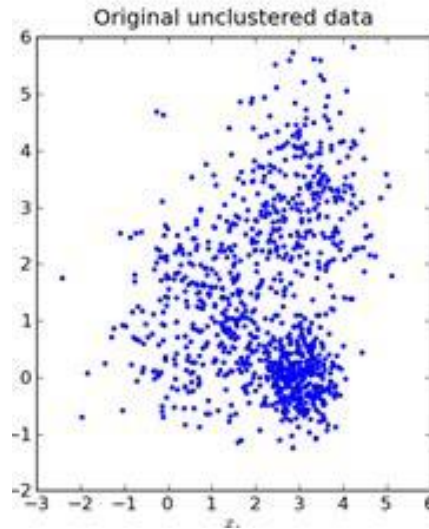
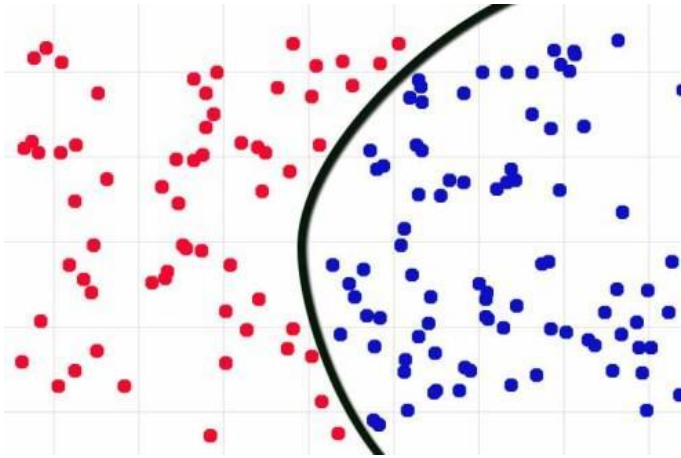
UNaB, Lic. Cs. de Datos, 2021 cuat. 2

Sebastián Pedersen (sebastian.pedersen (at) unab (punto) edu (punto) ar)

Componentes Principales

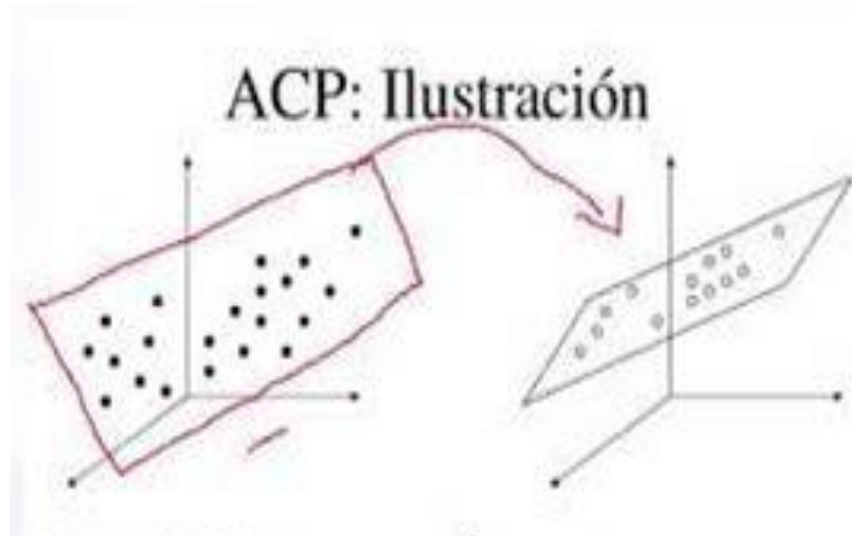
Reducción de la dimensión

Todo muy lindo con los gráficos, pero cuando los datos tienen más variables se hace imposible visualizarlos

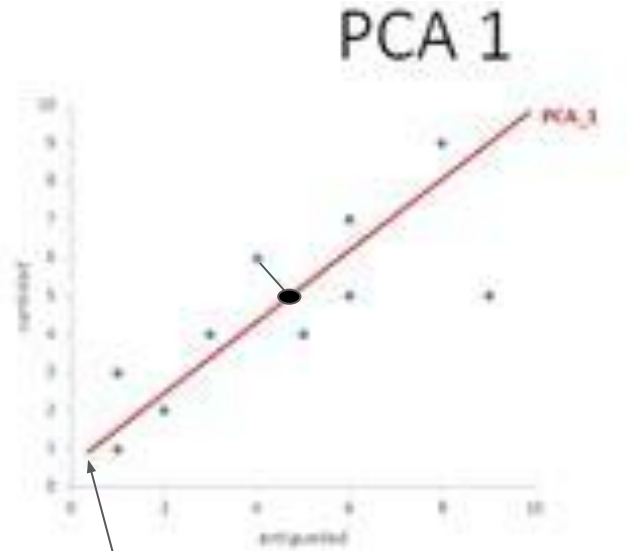


Componentes Principales

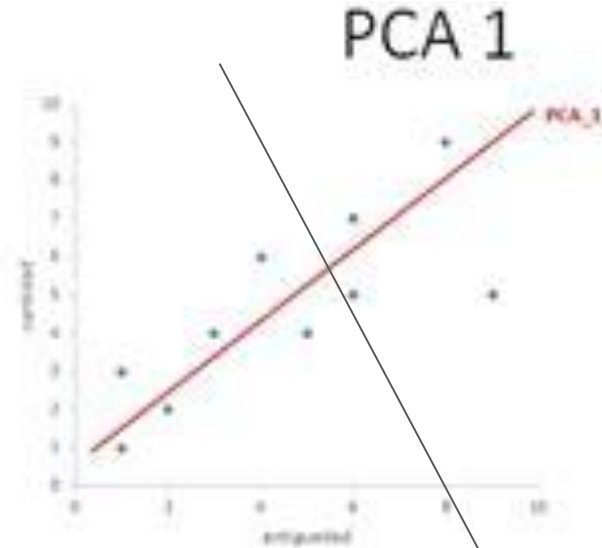
- Objetivo: proyectar los datos originales de muchas dimensiones a dos dimensiones, intentando preservar la mayor cantidad de información.



Componentes principales: busca direcciones de máxima varianza.



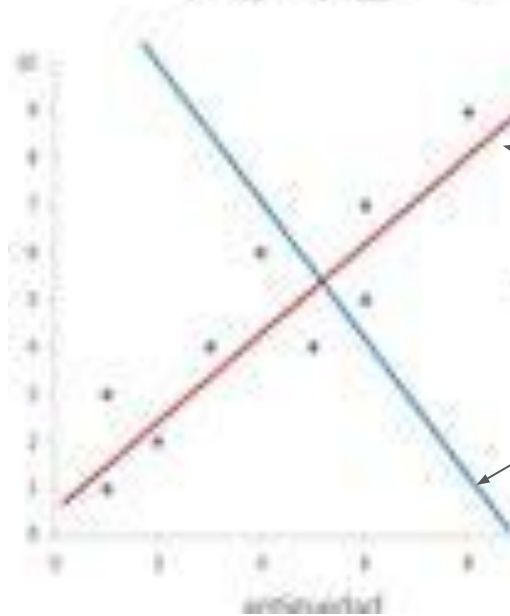
$$a.X1+b.X2 = Y1$$



¿Qué pasa si proyecto sobre esta recta?

Componentes principales: busca direcciones de máxima varianza.

PCA 1 + PCA 2



$$Y1 = a.X1 + b.X2$$

$$Y2 = c.X1 + d.X2$$

Ahora puedo reducir la dimensión de mis datos sin perder mucha información



Misma idea para pasar de 3, 4 o muchas dimensiones a 2 (o incluso a 1)

Referencias

- Bishop, Pattern Recognition and Machine Learning, sección 12.1.
- Hastie, Tibshirani, Introduction to Statistical Learning, sección 12.2.