

## Análisis de componentes principales

---

## PCA

---

- *una técnica estadística*
- *ámplicamente aplicada en descubrimiento de patrones en altas dimensiones*
- *opera a base de los vectores propios*

[http://www.cs.otago.ac.nz/cosc453/student\\_tutorials/principal\\_components.pdf](http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf)

## Operación

---

- *Obtener datos y ponlos en una matriz*
- *Restar la media en cada dimensión*
- *Calcular la matriz de covarianza*
- *Calcular sus valores y vectores propios*
- *Ordénalos y elige los mayores para formar un vector de características y transpónlo*
- *Multiplica la matriz de datos con este vector*

Hagan esto usando **numpy.linalg**

---

# Self-organizing map

---

## SOM

---

- *para la visualización de datos en alta dimensión*
- *reducción de dimensiones (típicamente a uno o dos) y agrupando elementos similares*

<http://davis.wpi.edu/~matt/courses/soms/>

## Pseudocódigo

---

- *Inicializa una rejilla*
- *En un ciclo donde  $t: (0, 1)$ :*
  - *Elige una muestra al azar*
  - *Elige el nodo que mejor corresponde*
  - *Escala sus vecinos*
  - *Incrementa  $t += dt$*

## Correspondencia

---

- *Con una medida de similitud*
  - *Algo simple tipo distancia Euclideana*

# Escalamiento

---

- *Acercamiento de los valores de los vecinos a la muestra colocada*
- *Hay que inventar una vecindad*
- *Además hay que inventar una función para la magnitud de cambio*
- *Este es el paso de “aprendizaje”*

Hagan en Python para agrupar colores (datos 3D)

---

*El ejemplo del tutorial está en Java*

## Máquinas de vectores de soporte

---

- *Máquinas de núcleo (kernel machines)*
- *Para aprendizaje a máquina*
- *Para detección de patrones*
- *A base de similitud entre datos*
  - *Usando los productos internos entre los datos individuales*

## SVM

---

<http://www.support-vector.net/icml-tutorial.pdf>

## Función de núcleo

---

- *Para datos no necesariamente linealmente separables*
- *Posiblemente en dimensionalidad muy alta*
- *En vez de manipular los vectores, sacamos un escalar por cada producto interno*

<http://www.svcl.ucsd.edu/courses/ece271B-F09/handouts/DPK.pdf>

## Estructura general

---

- *Una máquina de aprendizaje general*
- *Una función de núcleo específico al problema*
  - *El perceptron es de hecho un ejemplo de una función de núcleo*

Exploren diferentes funciones de núcleo

---

*en su código...*