

## Práctica 6: Análisis de componentes principales

### Objetivo

El objetivo es utilizar técnicas de reducción de dimensión basadas en el análisis de componentes principales. En la primera parte, utilizaremos datos sintéticos de distribuciones normales multivariadas en 3D. Segundo, utilizaremos SVD para comprimir una imagen. Una vez que nos hayamos familiarizado con el uso de PCA y SVD, evaluaremos el uso de PCA para la reconstrucción de dígitos y su clasificación. Para ello deberéis tener preparado el clasificador que utilizasteis con MNIST.

**Estudio previo** (es necesario prepararlo por escrito, antes de acudir a la práctica)

1. Repasa las transparencias de clase y estudia las funciones auxiliares proporcionadas para esta práctica. Repasa los conceptos de valores singulares/propios y vectores singulares/propios y su interpretación geométrica.

### Desarrollo de la práctica

Copia a tu directorio de trabajo los ficheros proporcionados, y comprueba que funcionan correctamente en Matlab.

1. Datos sintéticos 3D (script P61.m y fichero P61.mat): realiza el análisis de componentes principales del dataset X utilizando la descomposición en valores y vectores propios (función eig). Puedes seguir las instrucciones que aparecen como comentarios en el programa.
2. Compresión de imágenes: utiliza svd para comprimir una foto de tu cara (script p62.m). Para simplificar, utilizad una imagen en escala de grises. Puedes seguir las instrucciones que aparecen como comentarios en el programa.
3. PCA sobre MNIST: En este apartado vamos a utilizar el mismo conjunto de imágenes que se utilizó en la práctica 5. Deberás usar el algoritmo de clasificación realizado en esa práctica, así que, si no lo terminaste, hazlo antes de empezar esta. Además, puedes utilizar la misma función del apartado 2 para visualizar las componentes (script P63.m).
  - a. En primer lugar vamos a visualizar los distintos números sobre las dos componentes principales. Para ello, muestra la representación en las dos componentes principales de cada imagen con un color diferente para cada dígito. ¿Qué conclusión sacas de esta figura?

- b. En base al análisis del punto a, selecciona dos dígitos que se puedan clasificar usando solo estas dos componentes y dos que no. Usa el clasificador de la práctica anterior con estas dos componentes y determina si tu análisis era correcto. ¿Varían los resultados si sólo consideras los dígitos seleccionados y no todas las clases?

**A entregar** (en Moodle, dentro de un fichero .zip)

- **Programas P61.m, P62.m y P63.m**, junto con las funciones auxiliares que hayas programado y el fichero con tu imagen, que vaya mostrando por pantalla los resultados de todos los apartados.
- **Memoria de la práctica en un fichero P6.pdf ó P6.doc**. En la portada debe aparecer tu foto, nombre completo, y NIP. En las subsiguientes incluye los resultados de todos los apartados, su interpretación y las conclusiones que hayas obtenido.