Práctica 3: Aplicación de la descomposición de valores singulares para el cálculo de aproximaciones polinómicas.

Antes de implementar la práctica, responded a las siguientes preguntas.

1. ¿Dado un conjunto de n puntos, existe siempre un polinomio de grado m < n-1 que pase por dichos puntos?
2. ¿Se te ocurre una manera en la cual el cálculo de los valores singulares de una matriz permita calcular un polinomio de regresión adecuado?
3. ¿Como podrías transformar el problema de encontrar un polinomio de regresión en un problema de producto de matrices?
4. ¿Crees que, además, es posible calcular el error de regresión a partir de los valores singulares?

En general, los métodos de regresión polinómica que hemos visto consisten en el cálculo analítico del mínimo de una función de error. Dicho calculo pasaba por calcular el gradiente de cierta función en igualarlo a 0. En esta práctica vamos a ver que todos esos pasos no son realmente necesarios.

Apartado 1: Considerar el siguiente conjunto de datos:

i	x_i	y_i
1	0	1
2	0.25	1.2840
3	0.5	1.6487
4	0,75	2.1170
5	1	2.7183

Vamos a buscar el polinomio $P_2(x) = a_0 + a_1x + a_2x^2$ que minimiza el error al aproximar los datos anteriores usando la norma estándar de R^n. Es decir, vamos a usar una regresión cuadrática.

a) Vamos a escribir el problema de regresión como un problema de interpolación sobredimensionado asumiendo que el polinomio $P_2(x)$ pasa por todos los puntos $\mathbf{x_i,y_i}$. Para ello escribiremos el problema en la forma **Xa = y** donde **X** es una matriz 5 x 3, **a** es un vector 3 x 1 e **y** es un vector 5x1.
b) Calculad la descomposición en valores singulares $\mathbf{X = U\ S\ V^t}$ de la matriz **X** que habéis encontrado en el apartado anterior. Utilizad los apuntes de clase para obtener los valores singulares de **X** y para construir las matrices **U** de 5 x5, **S** de 5 x 3 y **V** 3x3.

Apartado 2: Se puede demostrar (de hecho, no es muy difícil), que los valores **a** = ($\mathbf{a_0}$, $\mathbf{a_1}$ **y** $\mathbf{a_2}$) del polinomio que buscamos se pueden obtener mediante el siguiente producto de matrices

a=Vz

El vector **z** = ($\mathbf{c_1/s_1, c_2/s_2, c_3/s_3}$), donde los valores $\mathbf{s_i}$ corresponden a los valores singulares de la matriz **A** y los valores $\mathbf{c_i}$ corresponden a los tres primeros valores del vector obtenido al multiplicar la matriz $\mathbf{U^t}$ por **y** = ($\mathbf{y_1, y_2, y_3, y_4, y_5}$).

Apartado 3: Se puede demostrar también (es más fácil aun) que el error cometido con la regresión obtenida para los valores es $\boldsymbol{E} = \sqrt{\boldsymbol{c_4}^2 + \boldsymbol{c_5}^2}$, calculad el error y comparadlo con el cálculo manual del error (es decir, comparad el valor E con el valor obtenido al calcular $\sqrt{\sum_{i=1}^{5}(P_2(x_i) - y_i)^2}$

Ejercicio opcional:

a) ¿Qué obtenemos cuando aplicamos el método de los apartados 1 y 2 a la construcción de un polinomio de grado 4?
b) Comprueba que, efectivamente el polinomio obtenido se corresponde en este caso a un polinomio de interpolación en lugar de a un polinomio de regresión.