Tarea 6

Fecha de publicación: Octubre 2, 2020

Fecha de entrega: Domingo 11 de octubre de 2020.

Ejercicio 1 (10 puntos).
Programar el método de mínimos cuadrados para ajustar un polinomio de grado n a un conjunto de puntos en el plano $(x_1, y_1), ..., (x_m, y_m)$.

1. Escriba una función que devuelva la transpuesta de una matriz. La función recibe como argumentos el apuntador a una matriz $\boldsymbol{A}$, su número de filas m y su número de columnas n. La salida de la función es $\boldsymbol{A}^\top$.
   Si usan una librería y esta función ya esta implementada, puede usarla.

2. Escriba una función que calcule el producto de dos matrices. La función recibe como argumentos el apuntador a una matriz $\boldsymbol{A}$, su número de filas m, su número de columnas n, el apuntador a una matriz $\boldsymbol{B}$ y su número de columnas q. La salida de la función es el producto de las matrices $\boldsymbol{AB}$. Suponemos que el producto se puede realizar.
   Si usan una librería y esta función ya esta implementada, puede usarla.

3. Escriba la función que calcula la solución de mínimos cuadrados. La función recibe como argumentos el apuntador a matriz $\boldsymbol{A}$, su número de filas m, su número de columnas n, y el apuntador a un vector $\boldsymbol{b}$. Use las funciones anteriores para calcular $\boldsymbol{A}^\top\boldsymbol{A}$. Calcule $\boldsymbol{A}^\top\boldsymbol{b}$ y la solución del sistema $\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^\top\boldsymbol{b}$ usando la factorización de Cholesky. Haga que la función devuelva un apuntador a $\boldsymbol{x}$ si el sistema fue resuelto o `NULL` si no.

4. Escriba un programa que reciba desde la línea de comandos

   - un entero d que se usará para generar una partición de un intervalo,
   - el nombre de un archivo de texto que corresponde a una tabla con los datos de

puntos 2D,

$$\begin{array}{cc} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_m & y_m \end{array}$$

- el entero n que corresponde al grado del polinomio $p(x) = c_n x^n + c_{n-1}x^{n-1} + ... + c_1 x + c_0$ que se quiere ajustar a ese conjunto de datos usando el enfoque de mínimos cuadrados. Es decir, encontrar los coeficientes $c_0, ..., c_n$ que son solución del problema de minimización

$$\min_{c_0,...,c_n} \sum_{i=1}^{m} [p(x_i) - y_i]^2 = E.$$

Construya la matriz $\boldsymbol{A}$ y el vector $\boldsymbol{b}$ correspondiente a este problema. Use la función del inciso anterior para calcular los coeficientes del polinomio. Haga que el programa imprima los valores de los coeficientes del polinomio $c_0, ..., c_n$ y el valor del error E.

Calcule el rango de valores de las abscisas,

$$x_{\min} = \min_i x_i, \quad x_{\max} = \max_i x_i.$$

Genere los puntos de una partición del intervalo $[x_{\min}, x_{\max}]$:

$$z_j = x_{\min} + \frac{x_{\max} - x_{\min}}{d} j, \quad j = 0, 1, ..., d.$$

Genere un archivo de texto que tenga dos columnas, en el que cada fila tiene los valores

$$z_j \qquad p(z_j)$$

para que pueda usar este archivo para gráficar el polinomio en el intervalo.

Puede usar la función de Horner de la tarea 2 para evaluar el polinomio.

5. Pruebe el programa con los archivos que aparecen en el ZIP `datosTarea06.zip`. Como sugerencia use primero alguna herramienta de graficación para ver la distribución de los puntos y pueda decidir cual es el grado del polinomio que va utilizar para ajustar el conjunto de datos.
Use el archivo proporcionado junto con el que genera programa para generar una gráfica que muestre tanto los puntos dados como la gráfica del polinomio que ajusta los datos para que visualmente puede evaluar el ajuste realizado. Agregue al reporte las gráficas generadas en cada caso y escriba un comentario sobre el ajuste obtenido.

Dos comentarios:

- Note que las abscisas de los puntos (x_i, y_i) no tienen que cumplir que $x_1 < x_2 < ... < x_m$. Pueden estar desordenados estos valores y por eso es que no podemos usar los puntos $(x_i, p(x_i))$ para generar la gráfica del polinomio uniéndolos con una línea poligonal. Por eso que tenemos que generar la partición del intervalo.
- No se trata de encontrar el polinomio que reduzca el error E lo más que se pueda, porque entre más grande es el grado del polinomio, éstos tienden a oscilar, crecer o decrecer muy rápido y si el propósito es usar el polinomio para hacer una predicción esos comportamientos hacen que no lo podamos usar. Es mejor buscar el polinomio de mejor grado que describa el comportamiento de los datos, es decir, el modelo más simple.