IBIO 2240: PROGRAMACIÓN CIENTÍFICA

# Profesor: Luis Felipe Giraldo T. Tarea #3

Universidad de los Andes 2023-I

Para cada problema, soluciones incompletas harán que la calificación del punto entero sea 0.0.

1. (20 puntos) Considere el siguiente problema de optimización:

donde 𝑥 = [𝑥1, 𝑥2]𝑇.

* 1. Encuentre el gradiente de 𝑓(𝑥) con respecto a 𝑥.
  2. Encuentre la Hessiana de 𝑓(𝑥) con respecto a 𝑥.
  3. A mano, resuelva cuatro iteraciones del método de Newton con una condición inicial 𝑥(0) = [5 , − 1]𝑇.

Para aplicar el método de Newton, necesitamos calcular la primera y segunda derivada de la función f(x):

Ahora, para la iteración k del método de Newton, encontramos una aproximación x(k+1) mediante la fórmula:

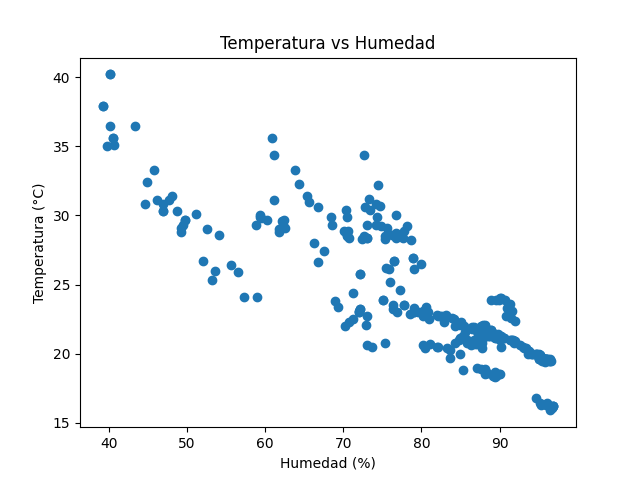
x(k+1) = x(k) - [Hf(x(k))]^-1 [∇f(x(k))]

Usando la condición inicial x(0) = [5, -1]T, podemos encontrar x(1):

Ahora podemos encontrar x(2) usando x(1) como condición inicial:

Continuamos de la misma manera para encontrar x(3) y x(4):

1. (40 puntos) Las mediciones de temperatura T (unidades en °C) y humedad relativa R (unidades en %) en un invernadero real de un cultivo de rosas se dan en el archivo greenhouse.txt. La primera columna es temperatura y la segunda es humedad.
   1. Grafique en un scatter plot la temperatura vs humedad. Analice esta gráfica.



**Figura 1.** Scatter plot de Temperatura (°C) vs Humedad (%) con los datos del invernadero.

Se nota una relación inversamente proporcional entre la temperatura y la humedad relativa en el invernadero; la tendencia de los datos muestra que, a menor humedad relativa, la temperatura tiende a aumentar, y a mayor humedad relativa, la temperatura tiende a disminuir. Se puede asumir que los comportamientos de las variables tienen una relación inversa puesto que el invernadero desea mantener una sensación térmica dentro de un determinado rango. Dado a que la sensación térmica aumenta cuando la temperatura y la humedad relativa aumentan, si se busca controlar esta característica dentro de un rango determinado, una de las dos variables deberá de disminuir su valor si la otra aumenta.

* 1. Asuma que la relación entre humedad y temperatura está dada por la ecuación 𝑇 = 𝛽0 + 𝛽1𝑅. Utilizando los datos en greenhouse.txt, estime 𝛽0 y 𝛽1 usando mínimos cuadrados.

Los betas asociados a la ecuación son: 𝛽0 = 46.40624408527903°C y 𝛽1 = -0.2835039292500497°C/%

Chart, scatter chart

Description automatically generated

**Figura 2.** Regresión lineal de la relación de temperatura y humedad usando el método de mínimos cuadrados.

En la fig. 2 se observa el resultado de linealizar la relación entre temperatura y humedad por el método de mínimos cuadrados. En esta se muestra que los betas encontrados son coherentes con la tendencia que sigue el modelo, por lo cual corresponden a una aproximación acertada de los datos reales.

* 1. Utilizando el modelo obtenido en el enunciado b), defina una regla para determinar si una medición nueva de la pareja (T, R) es anómala o no. Es decir, si tenemos una medición nueva de temperatura y humedad con los sensores del invernadero, podemos utilizar el modelo 𝑇 = 𝛽0 + 𝛽1𝑅 para saber qué tanto las mediciones nuevas siguen ese modelo. Si se desvían demasiado, podríamos considerar esas mediciones como anómalas.

Se definieron dos reglas para determinar si una medición es anómala. La primera regla consiste en obtener el error cuadrático total de la regresión. Una vez calculado, el valor resultante se divide por el número de mediciones (n), lo cual da un error promedio por medición con respecto a la regresión (E). Posteriormente, se calcula el error cuadrático de la medición, interpretado como la diferencia entre el valor real de la nueva temperatura medida y el valor obtenido al reemplazar su dato de humedad en el modelo, y si este error de la nueva medición es mayor al error promedio para una sola medición, esta medición se considera como anómala.

La segunda regla consiste en encontrar el máximo error presente (M) en una de las mediciones utilizadas para entrenar el modelo. Se calcula el error cuadrático de cada medición utilizada para entrenar el modelo, y el error más grande entre los calculados se toma como criterio para determinar si una nueva medición es anómala o no. Entonces, si el error cuadrático de una nueva medición es mayor al error máximo encontrado previamente, esta medición se considera como anómala.

* 1. Se tienen mediciones nuevas de las parejas temperatura y humedad en el archivo datosNuevos.txt. Utilizando la regla obtenida en el enunciado c) determine cuáles parejas de mediciones son anómalas y cuáles son consideradas como típicos (normales). Grafique en un scatter plot de temperatura vs humedad estos puntos, y pinte los anómalos con un color y los considerados como típicos con otro color. Identifique el patrón en la detección que se puede identificar en la gráfica.

Chart, scatter chart

Description automatically generated

**Figura 3.** Scatter plot Temperatura (°C) vs Humedad (%) de mediciones nuevas típicas y anómalas determinadas por la primera regla.

Chart, scatter chart

Description automatically generated

**Figura 4.** Scatter plot Temperatura (°C) vs Humedad (%) de mediciones nuevas típicas y anómalas determinadas por la segunda regla.

1. (40 puntos) Considere los datos con las parejas (x,y) en el archivo datosToy.txt.
   1. Asuma que estos datos están relacionados a través de la relación 𝑦 = 𝛽0 + 𝛽1𝑥. Usando mínimos cuadrados, estime los parámetros 𝛽0 y 𝛽1. Grafique los datos en un scatter plot (*y* vs *x*) y grafique la recta estimada. Analice los resultados.

Chart, scatter chart

Description automatically generated

En la regresión lineal, se obtuvieron los valores de beta1=6.3466 y beta2=-5.2480, lo que indica que la recta estimada es y=6.3466-5.2480x. En la gráfica obtenida, se puede observar que la recta se ajusta más o menos a los datos, aunque los datos de la relación entre x y y resultan en una forma con curva hace que la recta de regresión no vaya tan acorde a las curvas que se presentan, lo que sugiere que un modelo polinómico de grado mayor podría ajustarse mejor a los datos.

* 1. Asuma que estos datos están relacionados a través de la relación 𝑦 = 𝛽0 + 𝛽1𝑥 + 𝛽2𝑥2. Usando mínimos cuadrados, estime los parámetros 𝛽0 , 𝛽1, y 𝛽2. Grafique los datos en un scatter plot (*y* vs *x*) y grafique la parábola estimada. Analice los resultados.

Gráfico

Descripción generada automáticamente

Para la regresión cuadrática, los resultados indican que la relación entre la variable independiente y la variable dependiente no es lineal, sino que tiene una forma de parábola. La ecuación de regresión cuadrática muestra que la variable dependiente se puede predecir como una función cuadrática de la variable independiente. Se obtuvieron los valores de beta0=5.0580, beta1=-5.2480 y beta2=3.8581, lo que indica que la parábola estimada es y=5.0580-5.2480x+3.8581x^2. En la gráfica obtenida, se puede observar que la parábola se ajusta un poco mejor a algunos datos.

* 1. Asuma que estos datos están relacionados a través de la relación 𝑦 = 𝛽0 + 𝛽1𝑥 + 𝛽2𝑥2 + 𝛽3𝑥3. Usando mínimos cuadrados, estime los parámetros 𝛽0 , 𝛽1, 𝛽2, y 𝛽3. Grafique los datos en un scatter plot (*y* vs *x*) y grafique el polinomio estimado. Analice los resultados.

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Para la regresión cúbica, los resultados indican que la relación entre la variable independiente y la variable dependiente es aún más compleja que en la regresión cuadrática, y la forma de la curva se ajusta mejor a los datos que los modelos anteriores. La ecuación de regresión cúbica muestra que la variable dependiente se puede predecir como una función cúbica de la variable independiente. Se obtuvieron los valores de beta0=5.0580, beta1=1.3177, beta2=3.8581 y beta3=-10.9210, lo que indica que el polinomio estimado es y=5.0580+1.3177x+3.8581x^2-10.9210x^3. En la gráfica obtenida, se puede observar que el polinomio se ajusta muy bien a los datos.