IMT2220, Cálculo para ciencia de datos, 2023-2
Tarea 3

Fecha entrega: 23 de octubre de 2023

Instrucciones

Pueden discutir los problemas con sus compañeros. Sin embargo, sus entregas, tanto la parte escrita como los códigos, deben ser individuales. **No está permitido** copiar las respuestas de alguien más ni dejar que otros copien sus respuestas. El hacerlo se reflejará con la nota mínima en la evaluación (1.0).
Su entrega debe estar compuesta por un único archivo .pdf **escrito en LaTeX** que incluya todas las respuestas y todos los gráficos que se están pidiendo, junto con un archivo .zip que incluya todos los códigos que produjeron y utilizaron durante esta tarea. Estos dos archivos deben ser subidos en Canvas. Si hay más de una línea en un gráfico, utilice distintos estilos de línea o distintos colores para diferenciarlas e incluya leyendas. No se olvide de nombrar todos sus ejes y líneas en las leyendas. **Cada gráfico debe ser comentado** (un gráfico sin un análisis del mismo no es una respuesta apropiada).

1. (15 puntos) El operador laplaciano Δ, se define en coordenadas cartesianas (en 2 dimensiones) por:

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$$

donde u es una función al menos 2 veces continuamente diferenciable. Este operador es de gran utilidad al modelar fenómenos físicos mediante ecuaciones diferenciales parciales (o EDPs). En particular, este operador aparece en ecuaciones importantes tales como: Poisson, Helmholtz, calor y onda acústica. En estas aplicaciones, a veces tenemos geometría circular, ya sea en el dominio que queremos considerar o en simplificaciones características del modelo (por ejemplo, una fuente de calor que se propaga de manera radial). Por estos motivos, es beneficioso saber como se escribe el Laplaciano en coordenadas esféricas ($\Delta_{r,\theta}$). Demuestre que se cumple alguna de las siguientes expresiones para el laplaciano en coordenadas polares:

$$\begin{aligned}\Delta_{r,\theta} u &= \frac{1}{r} \cdot \frac{\partial}{\partial r}\left(r \frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} \\ &= \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2}\end{aligned}$$

donde:

$$x = r\cos(\theta)$$
$$y = r\sin(\theta)$$

Solution: Comenzamos viendo que:

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial u}{\partial \theta}\frac{\partial \theta}{\partial x}$$

Entonces:

$$\begin{aligned}\frac{\partial^2 u}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial u}{\partial \theta}\frac{\partial \theta}{\partial x}\right)\\ &= \left(\frac{\partial^2 u}{\partial r^2}\frac{\partial r}{\partial x} + \frac{\partial^2 u}{\partial \theta \partial r}\frac{\partial \theta}{\partial x}\right)\frac{\partial r}{\partial x} + \frac{\partial u}{\partial r}\frac{\partial^2 r}{\partial x^2} + \left(\frac{\partial^2 u}{\partial r \partial \theta}\frac{\partial r}{\partial x} + \frac{\partial^2 u}{\partial \theta^2}\frac{\partial \theta}{\partial x}\right)\frac{\partial \theta}{\partial x} + \frac{\partial u}{\partial \theta}\frac{\partial^2 \theta}{\partial x^2}\end{aligned}$$

Notamos que necesitamos conocer las derivadas parciales de r, θ respecto a x, y (dado que la segunda derivada de u es análoga porque solo hicimos regla de la cadena antes). Para esto, tenemo que armar el sistema de ecuaciones que hace esto:

$$\begin{aligned}1 &= \cos(\theta)\frac{\partial r}{\partial x} - r\sin(\theta)\frac{\partial \theta}{\partial x}\\ 0 &= \sin(\theta)\frac{\partial r}{\partial x} + r\cos(\theta)\frac{\partial \theta}{\partial x}\\ 0 &= \cos(\theta)\frac{\partial r}{\partial y} - r\sin(\theta)\frac{\partial \theta}{\partial y}\\ 1 &= \sin(\theta)\frac{\partial r}{\partial y} + r\cos(\theta)\frac{\partial \theta}{\partial y}\end{aligned}$$

Resolviendo, obtenemos que:

$$\begin{aligned}\frac{\partial r}{\partial x} &= \cos(\theta) & \frac{\partial \theta}{\partial x} &= -\frac{\sin(\theta)}{r}\\ \frac{\partial r}{\partial y} &= \sin(\theta) & \frac{\partial \theta}{\partial y} &= \frac{\cos(\theta)}{r}\\ \frac{\partial^2 r}{\partial x^2} &= \frac{\sin^2(\theta)}{r} & \frac{\partial^2 \theta}{\partial x^2} &= \frac{2\cos(\theta)\sin(\theta)}{r^2}\\ \frac{\partial^2 r}{\partial y^2} &= \frac{\cos^2(\theta)}{r} & \frac{\partial^2 \theta}{\partial y^2} &= -\frac{2\sin(\theta)\cos(\theta)}{r^2}\end{aligned}$$

Con esto, basta con reemplazar y se demustra lo pedido.

2. Considere el sistema de coordenadas $(u, v) = \boldsymbol{f}(x, y)$ como función de las coordenadas cartesianas (x, y):

$$u = \frac{x^3 + y^3}{2}$$
$$v = \frac{x^3 - y^3}{2}$$

(a) (5 puntos) ¿Bajo que condiciones existe las funciones inversas $x = x(u, v)$ e $y = y(u, v)$?

(b) (5 puntos) Donde sea posible, calcule dichas inversas.

(c) (5 puntos) Ahora, considere $u, v = \boldsymbol{g}(r, \theta)$, es decir, en términos de las coordenadas polares:

$$u = \frac{r^3}{2}(\cos^3(\theta) + \sin^3(\theta))$$
$$v = \frac{r^3}{2}(\cos^3(\theta) - \sin^3(\theta))$$

¿Cuándo es verdad que r, θ son funciones de u, v?

Solution:

1. Calculamos el Jacobiano correspondiente:

$$\left|\frac{\partial(u, v)}{\partial(x, y)}\right| = \left|\begin{bmatrix} \frac{3}{2}x^2 & \frac{3}{2}y^2 \\ \frac{3}{2}x^2 & -\frac{3}{2}y^2 \end{bmatrix}\right| = -\frac{9}{4}x^2y^2 - \frac{9}{4}x^2y^2 = -\frac{9}{2}x^2y^2$$

Luego, podemos encontrar una única inversa en todos lados salvo el origen, lugar donde el determinante se anula.

2.

$$x = \sqrt[3]{u + v}$$
$$y = \sqrt[3]{u - v}$$

3. De la misma forma que en el item (a) (y asumiendo que :

$$\begin{aligned}
\left|\frac{\partial(u, v)}{\partial(r, \theta)}\right| &= \left|\begin{bmatrix} \frac{3}{2}r^2(\cos^3(\theta) + \sin^3(\theta)) & \frac{3}{2}r^3(-\cos^2(\theta)\sin(\theta) + \sin^2(\theta)\cos(\theta)) \\ \frac{3}{2}r^2(\cos^3(\theta) - \sin^3(\theta)) & \frac{3}{2}r^3(-\cos^2(\theta)\sin(\theta) - \sin^2(\theta)\cos(\theta)) \end{bmatrix}\right| \\
&= -\frac{9}{4}r^5(\cos^3(\theta) + \sin^3(\theta))\sin(\theta)\cos(\theta)(\cos(\theta) + \sin(\theta)) \\
&\quad + \frac{9}{4}r^5(\cos^3(\theta) - \sin^3(\theta))\sin(\theta)\cos(\theta)(\cos(\theta) - \sin(\theta)) \\
&= \ldots \text{ (Muchos pasos de simplificación)} \\
&= -\frac{9}{8}r^5\sin(2\theta)^2
\end{aligned}$$

Luego, esto se hace 0 siempre que $r = 0$ o bien $\theta \in \{0, \pi/2, \pi, 3\pi/2, 2\pi\}$. Es decir, siempre que pase esto no tenemos asegurada la existencia de una única inversa.

3. (15 puntos) Determine y clasifique los puntos críticos de la función $f : \mathbb{R}^3 \to \mathbb{R}$ definida por:

$$f(x, y, z) = (x^2 + y^2 + z^2) \ln(1 + x^2 + y^2 + z^2)$$

Solution: Esta función es el caso particular de la función vista en la tarea 2:

$$f(\boldsymbol{x}) = \|\boldsymbol{x}\|_2^2 \ln(1 + \|\boldsymbol{x}\|_2^2)$$

en 3 variables, así que su gradiente está dado por las entradas:

$$\nabla f(\boldsymbol{x})_j = 2\left(\ln(1 + \|\boldsymbol{x}\|_2^2) + \frac{\|\boldsymbol{x}\|_2^2}{1 + \|\boldsymbol{x}\|_2^2}\right) x_j$$

Analizamos lo de dentro del paréntesis. Sabemos que $\ln(1 + \|\boldsymbol{x}\|_2^2) \geq 0$, y que en particular es 0 solo cuando $\boldsymbol{x} = 0$. Lo mismo con la fracción que lo acompaña. Finalmente, el término x_j también es 0 solo cuando dicha componente es 0. Como queremos que $\nabla f = 0$, esto ocurre solo en $(x, y, z) = (0, 0, 0)$, el cual es nuestro punto crítico.

Finalmente, queremos ver que tipo de punto es. Inspeccionamos los puntos cercanos al punto crítico en cuestión. Sea $(x, y, z) \in B_\varepsilon(0, 0, 0) \setminus \{(0, 0, 0)\}$, denotamos por $\eta = x^2 + y^2 + z^2$. Notamos que $0 < \eta \leq \varepsilon^2$. Así:

$$0 < f(x, y, z) = \eta \ln(1 + \eta) \leq \varepsilon^2 \ln(1 + \varepsilon^2)$$

por lo que cada punto en una bola de radio ε alrededor del punto crítico (menos el mismo punto crítico), el valor de la función f es superior a cuanto vale en el punto crítico. Es decir, el punto crítico en cuestión $(0, 0, 0)$ es un mínimo local.

Nota: Si calculó el Hessiano y llegó a decir que el criterio es inconcluyente porque evaluado en el punto crítico corresponde a la matriz $\mathbf{0}$, otorgar puntaje parcial.

4. En este problema estudiaremos el problema de regresión lineal vía mínimos cuadrados. Considere un conjunto de datos $\{\boldsymbol{x}_i\}_{i=1,\ldots,m} = \{(x_i, y_i)\}_{i=1,\ldots m} \subset \mathbb{R}^{m\times 2}$. Queremos encontrar la mejor recta que ajuste a dichos datos. Una forma de hacerlo ya lo vimos en la tarea 2, con el algoritmo de descenso de gradiente estocástico. Esto algoritmo, por otro lado, no lleva a una única solución, y las soluciones obtenidas pueden tener varianza elevada.

   Formulamos el problema como: encontrar constantes a, b tal que la recta $f(x) = ax + b$ minimice la distancia al cuadrado delos puntos a la recta. Es decir, resolver el problema de optimización:

$$\min_{(a,b)\in\mathbb{R}^2} \frac{1}{2} \sum_{i=1}^{m} (ax_i + b - y_i)^2 = \min_{(a,b)\in\mathbb{R}^2} L(a, b; \boldsymbol{x})$$

   Esta función $L(a, b; \boldsymbol{x})$ se puede escribir en términos de $\mathbf{X} \in \mathbb{R}^{m\times 2}$, $\boldsymbol{\theta} = (a, b) \in \mathbb{R}^2$ y $\boldsymbol{y} \in \mathbb{R}^m$, para $m > 2$, donde esta matrices y vectores corresponden a un sistema sobredeterminado dado por:

$$\mathbf{X}\boldsymbol{\theta} = \boldsymbol{y}$$

(a) (3 puntos) Escriba L en términos de $\mathbf{X}, \boldsymbol{y}$ y $\boldsymbol{\theta}$ (es decir, $L = L(\boldsymbol{\theta}; \mathbf{X}, \boldsymbol{y})$). Como recuerdo, la construcción de la matriz $\mathbf{X}$, y el vector $\boldsymbol{y}$ es análoga a lo que se hizo en la tarea 2.

Solution: Partimos escribiendo la función L original:

$$L(a, b; \boldsymbol{x}) = \frac{1}{2} \sum_{i=1}^{m} (ax_i + b - y_i)^2$$

Luego, definimos $\mathbf{X} \in \mathbb{R}^{m \times 2}$, $\boldsymbol{\theta} \in \mathbb{R}^2$ y $\boldsymbol{y} \in \mathbb{R}^m$ como sigue:

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}, \quad \boldsymbol{\theta} = (b, a), \quad \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

De esta forma, reescribimos la función L como:

$$L(\boldsymbol{\theta}; \mathbf{X}, \boldsymbol{y}) = \frac{1}{2} \left\| \mathbf{X}\boldsymbol{\theta} - \boldsymbol{y} \right\|_2^2$$

(b) (7 puntos) Encuentre los puntos críticos de L y verifique que estos son mínimos. Recuerde que como estamos minimizando sobre los parámetros, queremos calcular el gradiente respecto a $\boldsymbol{\theta}$ solamente.

Solution: Los puntos críticos de L se obtienen derivando L con respecto a $\boldsymbol{\theta}$:

$$\begin{aligned} \nabla_{\boldsymbol{\theta}} L &= \frac{1}{2} \nabla_{\boldsymbol{\theta}} \langle \mathbf{X}\boldsymbol{\theta} - \boldsymbol{y}, \mathbf{X}\boldsymbol{\theta} - \boldsymbol{y} \rangle \\ &= \frac{1}{2} \nabla_{\boldsymbol{\theta}} \left(\boldsymbol{\theta}^\top \mathbf{X}^\top \mathbf{X} \boldsymbol{\theta} - 2\boldsymbol{\theta}^\top \mathbf{X}^\top \boldsymbol{y} + \|\boldsymbol{y}\|_2^2 \right) \\ &= \mathbf{X}^\top \mathbf{X} \boldsymbol{\theta} - \mathbf{X}^\top \boldsymbol{y} \end{aligned}$$

por lo cual, los puntos críticos se encuentran en aquellos puntos tales que:

$$\boldsymbol{\theta}^* = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X} \boldsymbol{y}$$

Además, viendo la matriz Hessiana de L:

$$\nabla_{\boldsymbol{\theta}}^2 L = \mathbf{X}^\top \mathbf{X}$$

que es positiva definida, y por lo tanto el punto crítico es mínimo local. Demostrar este hecho, por ejemplo, viendo que si nos damos un vector $\boldsymbol{c} \in \mathbb{R}^2 \setminus \{0\}$:

$$\boldsymbol{c}^\top \mathbf{X}^\top \mathbf{X} \boldsymbol{c} = \boldsymbol{w}^\top \boldsymbol{w} > 0$$

(Formalmente faltan unos supuestos para esto realmente, pero basta suponer que todo se comporta bien).

(c) (5 puntos) Considere el set de datos entregado `HW3_data.csv`. Grafique la recta obtenida junto a dicho set de datos y calcule el valor de $L(\boldsymbol{\theta}; \mathbf{X}, \boldsymbol{y})$ para esta recta.

Solution: Resolver el sistema que determina el punto crítico y graficar según este.