Entrega: **20 de febrero, 2024**

Versión: 12 de febrero de 2024

Tarea 1

Las siguientes preguntas requieren tiempo para pensar, pero no requieren respuestas extensas. Deben resolverse utilizando hasta donde sea posible las propiedades de gradientes y trazas vistas en clase. Propiedades adicionales de las trazas y los gradientes los encuentra en el "Matrix Codebook" de Petersen y Pedersen (2012). Nótese que allí en vez de $\nabla_{\boldsymbol{X}}$, se usa la notación $\frac{\partial}{\partial \boldsymbol{X}}$, pero en esta tarea debe utilizar en sus soluciones la notación indicada en la sección I, es decir $\nabla_{\boldsymbol{X}}$.

El objetivo de la tarea es asegurar que se maneja el lenguaje matemático a utilizar en varias partes del curso, así como con GNU/Octave ayudar a ilustrar algunos conceptos fundamentales asociados.

I. Aclaraciones sobre notación

Tómese en cuenta que:

- Vector $\underline{\boldsymbol{x}}$ siempre es un vector columna y $\underline{\boldsymbol{x}}^\top$ es siempre un vector fila.
- Matriz $\boldsymbol{A}$ tiene columnas $\underline{\boldsymbol{a}}_{:,i}$ y filas $\underline{\boldsymbol{a}}_{i,:}^\top$; esto es, $\underline{\boldsymbol{a}}_{i,:}$ es la i-ésima fila como columna.
- Producto punto: $\underline{\boldsymbol{x}}^\top \underline{\boldsymbol{y}} = \underline{\boldsymbol{x}} \cdot \underline{\boldsymbol{y}}$ es escalar.
- Producto externo: $\underline{\boldsymbol{x}}\underline{\boldsymbol{y}}^\top$ es una matriz.
- Producto Schur: $\underline{\boldsymbol{x}} \odot \underline{\boldsymbol{y}}$ es un vector.
- $\underline{\boldsymbol{x}}\underline{\boldsymbol{y}}$: no está definido.

Y con respecto a las derivadas:

- El operador $\frac{d}{dx} f(x)$ aplica para $f : \mathbb{R} \to \mathbb{R}$, x una variable *escalar*.
- El operador $\frac{\partial}{\partial x_i} f(\underline{\boldsymbol{x}})$ aplica para $f : \mathbb{R}^n \to \mathbb{R}$, y x_i una variable *escalar*, componente de $\underline{\boldsymbol{x}}$.
- La notación $f'(x)$ aplica únicamente para $f : \mathbb{R} \to \mathbb{R}$, pues $f'(x) = \frac{d}{dx} f(x)$.
- El operador $\nabla_{\underline{\boldsymbol{x}}} f(\underline{\boldsymbol{x}})$ aplica para $f : \mathbb{R}^n \to \mathbb{R}$, con $\underline{\boldsymbol{x}} \in \mathbb{R}^n$.
- El operador $\nabla_{\boldsymbol{X}} f(\boldsymbol{X})$ aplica para $f : \mathbb{R}^{m \times n} \to \mathbb{R}$, con $\boldsymbol{X} \in \mathbb{R}^{m \times n}$.

II. Gradientes

Sea la matriz $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ simétrica, y sea el vector $\underline{\boldsymbol{x}} \in \mathbb{R}^n$.

1. Encuentre el gradiente $\nabla_{\underline{\boldsymbol{x}}} f(\underline{\boldsymbol{x}})$ para $f(\underline{\boldsymbol{x}}) = \frac{1}{2}\underline{\boldsymbol{x}}^\top \boldsymbol{A}\underline{\boldsymbol{x}} + \underline{\boldsymbol{b}}^\top \underline{\boldsymbol{x}}$ utilizando las propiedades de trazas y gradientes vistas en clase. (10 pts)

2. Realice una función en GNU/Octave que reciba una matriz $\boldsymbol{A}$ de tamaño 2×2 y un vector $\underline{\boldsymbol{b}}$ de dos dimensiones, y que grafique la superficie paraboloide tridimensional dada por

$$f(\underline{\boldsymbol{x}}) = \frac{1}{2}\underline{\boldsymbol{x}}^\top \boldsymbol{A}\underline{\boldsymbol{x}} + \underline{\boldsymbol{b}}^\top \underline{\boldsymbol{x}}$$

   en un rango de los componentes x_1 y x_2 de $\underline{\boldsymbol{x}} = [x_1, x_2]^\top$ también dado por el usuario.

   Lo más relevante de este punto es que el código esté *vectorizado*, es decir, que se evite utilizar ciclos `for`, y que todo se plantee en términos de sumas y productos de matrices y vectores (incluyendo productos de Schur cuando sea necesario).

   Algunas funciones que puede utilizar son, entre otras: `sum`, `sumsq`, `vecnorm`, `dot`, `meshgrid`, `surf`, `contour`. (10 pts)

3. Sabiendo que el mínimo se encuentra donde el gradiente de la función es cero, encuentre dónde está el mínimo de la función. (5 pts)

4. Use su función para mostrar tres casos de paraboloides: (10 pts)

   4.1. Matriz $\boldsymbol{A}$ igual a la matriz identidad escalada $c\boldsymbol{I}$.
   4.2. Matriz diagonal pero con los dos elementos de la diagonal distintos.
   4.3. Matriz simétrica no diagonal, que debe ser positiva definida.

   Elija $\underline{\boldsymbol{b}}$ distinto de cero y los rangos de $\underline{\boldsymbol{x}}$, de modo que el mínimo del paraboloide sea visible.

5. Utilice ahora la función `quiver` para mostrar además el gradiente de la función paraboloide del punto 4.3. (5 pts)

6. Demuestre que el gradiente de la función $f(\underline{\boldsymbol{\theta}})$ definida como (5 pts)

$$f(\underline{\boldsymbol{\theta}}) = \frac{1}{2}\|\boldsymbol{X}\underline{\boldsymbol{\theta}} - \underline{\boldsymbol{y}}\|^2$$

   respecto a $\underline{\boldsymbol{\theta}} \in \mathbb{R}^n$, está dado por

$$\nabla_{\underline{\boldsymbol{\theta}}} f(\underline{\boldsymbol{\theta}}) = \boldsymbol{X}^\top(\boldsymbol{X}\underline{\boldsymbol{\theta}} - \underline{\boldsymbol{y}})$$

   donde $\boldsymbol{X} \in \mathbb{R}^{m\times n}$. Debe utilizar las propiedades de trazas y gradientes vistas en clase.

7. Encuentre ahora el gradiente $\nabla_{\boldsymbol{X}} f(\boldsymbol{X})$ con (10 pts)

$$f(\boldsymbol{X}) = \frac{1}{2}\|\boldsymbol{X}\underline{\boldsymbol{\theta}} - \underline{\boldsymbol{y}}\|^2$$

   respecto a $\boldsymbol{X} \in \mathbb{R}^{m\times n}$, donde $\underline{\boldsymbol{\theta}} \in \mathbb{R}^n$. Debe utilizar las propiedades de trazas y gradientes vistas en clase.

   Ayuda: *Revise propiedades adicionales de las trazas y los gradientes en el "Matrix Codebook" de Petersen y Pedersen (2012). Nótese que allí en vez de $\nabla_{\boldsymbol{X}}$, se usa $\frac{\partial}{\partial \boldsymbol{X}}$. Debe utilizar en su solución la notación indicada en la sección I, es decir $\nabla_{\boldsymbol{X}}$.*

8. Encuentre ahora el gradiente $\nabla_{\underline{\boldsymbol{x}}} f(\underline{\boldsymbol{x}})$ con la composición $f(\underline{\boldsymbol{x}}) = g(h(\underline{\boldsymbol{x}}))$ con $g : \mathbb{R} \to \mathbb{R}$ y $h : \mathbb{R}^n \to \mathbb{R}$, ambas diferenciables. (10 pts)

9. Sea ahora $f(\underline{\boldsymbol{x}}) = g(\underline{\boldsymbol{a}}^\top \underline{\boldsymbol{x}})$ con $g : \mathbb{R} \to \mathbb{R}$ continuamente diferenciable y $\underline{\boldsymbol{a}} \in \mathbb{R}^n$ un vector. Encuentre $\nabla_{\underline{\boldsymbol{x}}} f(\underline{\boldsymbol{x}})$. (5 pts)

III. Matrices positivas definidas

1. Sea $\underline{\boldsymbol{z}} \in \mathbb{R}^n$ un vector n-dimensional. Muestre que $\boldsymbol{A} = \underline{\boldsymbol{z}}\underline{\boldsymbol{z}}^\top$ es positiva semidefinida. (10 pts)

2. Sea $\underline{\boldsymbol{z}} \in \mathbb{R}^n$ un vector n-dimensional no nulo, y $\boldsymbol{A} = \underline{\boldsymbol{z}}\underline{\boldsymbol{z}}^\top$. Encuentre cuál es el espacio nulo de $\boldsymbol{A}$ y su rango. (10 pts)

3. Compruebe sus hallazgos en GNU/Octave de la siguiente forma: Genere un vector $\underline{\boldsymbol{z}}$ aleatorio no nulo en tres dimensiones. Genere con él la matriz $\boldsymbol{A} = \underline{\boldsymbol{z}}\underline{\boldsymbol{z}}^\top$. Sintetice 100 vectores aleatorios $\underline{\boldsymbol{n}}_i$ en el espacio nulo. Compruebe que en cada caso $\boldsymbol{A}\underline{\boldsymbol{n}}_i = \underline{\mathbf{0}}$. (15 pts)

   Sugerencia: el procedimiento de Gram-Schimdt revisado en el curso de Señales y Sistemas le puede ser de utilidad.

IV. Ecuaciones normales

1. Usando los puntos del paraboloide en el punto 4.3, utilice las ecuaciones normales de regresión lineal para encontrar los parámetros de un plano que mejor ajuste dichos puntos. (10 pts)

2. Grafique el plano encontrado. (5 pts)

Esta tarea se realiza en parejas o individual. Deben enviar por medio del hilo correspondiente en el foro de *cafetería* del tecDigital la integración de los grupos, para activar la entrega en el tecDigital.

Solo lo que sea subido al tecDigital será calificado. Entregas por correo serán ignoradas, por lo que asegúrese de haber enviado su grupo (aún si es individual) a tiempo.

Los códigos deben manejarse utilizando esta invitación al Github-Classroom.

Toda tarea debe ser resultado del trabajo intelectual propio de la persona o personas que la entregan. Además de la literatura de referencia, solo puede utilizarse el material expresamente así indicado en la tarea, lo que incluye código brindado por el profesor o indicado en los enunciados a ser utilizado como base de la tarea. Expresamente quedan excluidos como material de referencia los trabajos entregados por otros estudiantes en el mismo semestre o semestres anteriores.

Nótese que esto no elimina la posibilidad de discutir estrategias de solución o ideas entre personas y grupos, lo cual es incluso recomendado, pero la generación concreta de cada solución, derivación o programa debe hacerse para cada entrega de forma independiente.

Para toda referencia de código o bibliografía externa deben respetarse los derechos de autor, indicando expresamente de dónde se tomó código, derivaciones, etc. Obsérvese que el código entregado por el profesor usualmente ya incluye encabezados con la autoría correspondiente. Si un estudiante agrega código a un archivo, debe agregar su nombre a los encabezados si la modificación es de más del 50 % del archivo, o indicar expresamente en el código, con comentarios claros, la autoría de las nuevas líneas de código, pues a la autora o al autor del archivo original no se le debe atribuir código que no es suyo.

Si se detecta código o deducciones teóricas iguales o muy cercanas a trabajos de otros estudiantes del mismo semestre o de semestres anteriores, se aplicará lo establecido por la reglamentación vigente, en particular el Artículo 75 del Reglamento de Régimen de Enseñanza y Aprendizaje.

Modificaciones de comentarios, cadenas alfanuméricas, nombres de variables, orden de estructuras independientes, y otras modificaciones menores de código se siguen considerando como clones de código, y las herramientas automatizadas de detección reportarán la similitud correspondiente.

Los estudiantes que provean a otros estudiantes del mismo o futuros semestres soluciones de sus tareas, también son sujetos a las sanciones especificadas en la reglamentación institucional. Por lo tanto, se advierte no poner a disposición soluciones de las tareas a otros estudiantes.