Guía 4

1. Calcular el gradiente $\nabla f = \frac{df}{d\boldsymbol{x}}$, para $f(\boldsymbol{x}) = \mathrm{tr}(\boldsymbol{x}\boldsymbol{x}^T + \sigma^2 I),\ \boldsymbol{x} \in \mathbb{R}^n$

2. Dada $f = Ax$, $A \in \mathbb{R}^{3\times 2}$, $x \in \mathbb{R}^2$. Calcular $\frac{df}{dA}$, ¿qué dimensiones tiene?

3. Dado el campo escalar $f(x, y) = e^{-x^2-y^2}(1 - x^2 - y^2 - 2\alpha xy)$:

   - Mostrar que para el vector nulo: $f(0) = 1$; $\nabla f(0) = 0$;
     $H(0) = \begin{pmatrix} -4 - 2\alpha \\ -2\alpha - 4 \end{pmatrix}$
   - Diagonalice la matriz Hessiana y halle el polinomio de Taylor de f de grado 2 alrededor del vector nulo.

4. Dada $f = (x^2 \cos(x))^2 + x^2 \cos(x)$, armar el grafo de cómputo y calcular la derivada de f respecto de x usando diferenciación automática.

5. Sea una red neuronal de dos capas, una de 3 neuronas y otra de 1 con los parámetros inicializados con los siguientes valores:

$$w^{(1)} = \begin{pmatrix} 0.1 & -0.5 \\ -0.3 & -0.9 \\ 0.8 & 0.02 \end{pmatrix}, b^{(1)} = \begin{pmatrix} 0.1 \\ 0.5 \\ 0.8 \end{pmatrix}, w^{(2)} = \begin{pmatrix} -0.4 & 0.2 & -0.5 \end{pmatrix}, b^{(2)} = 0.7$$

   y donde cada capa calcula su salida vía

$$y^{(i)} = \sigma(w^{(i)} \cdot x^{(i)} + b^{(i)})$$

   donde σ es la función sigmoidea.

   Dada la observación $x = \begin{pmatrix} 1.8 \\ -3.4 \end{pmatrix}$, $y = 5$ y la función de costo $J(\theta) = \frac{1}{2}(\hat{y}_\theta - y)^2$, calcular las derivadas de J respecto de cada parámetro.