Pontificia Universidad Católica de Chile
Instituto de Ingeniería Matemática y Computacional
Magister en Inteligencia Artificial - MIA

Tarea 2
Fundamentos Matemáticos para la Inteligencia Artificial
IMT3850 2023

Prof. Manuel A. Sánchez
Mayo 2023

Preguntas

1. (10 puntos) **Variables aleatorias**

   Suponga que x e y son variables aleatorias independientes con media 0 y varianza 1. Cual es la media μ_Z y la matriz de covarianza de $Z = (x, y, ax + by)$, $a, b \in \mathbb{R}$.

2. (20 puntos) **Algoritmo Random Quicksort.**

   En clases observamos que el número esperado de comparaciones realizadas por el algoritmo Quicksort Aleatorio es de $2n\ln(n) + \mathcal{O}(n)$, donde n es el largo de la lista S. El objetivo es testear este resultado. Para esto:

   a) Programe el algoritmo Quicksort con pivot aleatorio presentado en clases para ordenar de forma ascendente una lista S de números reales distintos. Corrobore que su algoritmo funciona mostrando la lista ordenada $S = [0, 5, 4, 1, 7, 6, 3, 2, 8, 9]$.
   b) Para largo de lista n fijo, obtenga lista aleatorias y aplique el algoritmo a cada una de ellas calculando el numero de comparaciones realizadas en el algoritmo. Calcule el promedio de estas para n fijo.
   c) Repite el procedimiento anterior para $n = [100, 200, 300, ..., 5000]$ y grafique n vs. el promedio de comparaciones para cada n. Ademas grafique las curvas correspondientes a $y = 2n\ln(n)$ y $y = 2n$.
   d) Explique porque los resultados del experimento corroboran los resultados teóricos.

3. (10 puntos) **Convexidad**

   Sea $f : \mathbb{R}^d \to \mathbb{R}$ una función es diferenciable. Pruebe que f es una función convexa si y sólo si se satisface

   $$f(y) \geq f(x) + \nabla f(x)^\top (y - x)$$

   para todo x, y en el dominio de f.

4. (20 puntos) **Descenso del gradiente**

   Considere el set de datos `datos_lineales.csv` que contiene pares de puntos $(x, y) \in \mathbb{R}^2$. Se busca resolver el problema:

   $$\min_{\theta \in \mathbb{R}^2} \frac{1}{2N} \|X\theta - y\|_2^2, \quad X \in \mathbb{R}^{N \times 2}, y \in \mathbb{R}^N$$

   donde θ son los coeficientes del polinomio lineal que mejor ajusta la matriz de datos X con los valores y.

   a) Programe el algoritmo de descenso de gradiente para resolver este problema. Considere como parámetros de este algoritmo: `X, y, NITMAX, gamma` donde `NITMAX` es el máximo número de iteraciones y `gamma` es la función tasa de aprendizaje.
   b) Programe el algoritmo de descenso de gradiente estocástico para resolver este problema. Considere como parámetros de este algoritmo: `X, y, NITMAX, gamma` donde `NITMAX` es el máximo número de iteraciones y `gamma` es la función tasa de aprendizaje.

c) Considere las siguientes tasas de aprendizaje:

$$\gamma(t) = \gamma_{\text{clases}}(t), \quad \gamma(t) = \log(t)$$

Donde γ_{clases} es la heurística vista en clases. Para cada una ejecute 100 veces el algoritmo estocástico y reporte el promedio de los resultados. Compare estos sus resultados con los del algoritmo determinista con la tasa correspondiente.

d) Repita el item anterior pero ahora con el set de datos `datos_cuadraticos.csv`. Notar que ahora $X \in \mathbb{R}^{N\times 3}$ y que $\theta \in \mathbb{R}^3$.

5. (10 puntos) **Bonus: Algoritmo de mediana aleatoria**

Programe el algoritmo mediana aleatoria visto en clases. Realice experimentos aleatorios y muestre evidencia de que el algoritmo calcula con éxito la mediana de una lista.