PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
INSTITUTO DE INGENIERÍA MATEMÁTICA Y COMPUTACIONAL
MAGISTER EN INTELIGENCIA ARTIFICIAL - MIA

TAREA 1
Fundamentos Matemáticos para la Inteligencia Artificial
IMT3850 2025

Prof. Manuel A. Sánchez
Marzo 2025

Preguntas

1. (5 puntos) **Normas.**

   Demuestre la siguiente identidad

   $$\mathrm{rms}_w^2(x) = \mathrm{avg}_w^2(x) + \mathrm{std}_w^2(x).$$

   donde $w_1, w_2, \ldots, w_n$ son pesos positives tales que la suma total de los pesos es

   $$W = \sum_{i=1}^{n} w_i > 0,$$

   y donde tenemos las definiciones

   $$\mathrm{avg}_w(x) = \frac{1}{W}\sum_{i=1}^{n} w_i x_i,$$
   $$\mathrm{std}_w^2(x) = \frac{1}{W}\sum_{i=1}^{n} w_i (x_i - \mathrm{avg}_w(x))^2,$$
   $$\mathrm{rms}_w(x) = \sqrt{\frac{1}{W}\sum_{i=1}^{n} w_i x_i^2}.$$

2. (15 puntos) **Aplicacion de algoritmo k-means.**

   Este problema trataremos el problema de clustering, es decir, determinar los k representates de un conjunto de datos y las etiquetas de estos que minimizan la distancia de cada cluster a su representante.

   a) Programe el algortimo de k-means. Construya una rutina

   `k_means_fit(X, Z0, NITERMAX)`

   donde `X` son los datos `Z0` son los representates iniciales, y `NITERMAX` es el número máximo de iteraciones.

   b) Use la base de datos `datakmeans.csv` para testear su algoritmo con $k = 5$ representantes. Grafique el conjunto de datos `X` por etiqueta y los representates de cada cluster. Grafique el comportamiento de la función respecto al número de iteraciones y discuta el número de iteraciones adecuado para el clustering.

   c) Use la base de datos de imágenes de dígitos de MNIST para testear su algoritmo con $k = 20$ representantes. Grafique cada representante y uno de los datos asignados a dicho cluster. Grafique el comportamiento de la función objetivo J^{clust} versus al número de iteraciones y discuta el numero de iteraciones adecuado para el clustering.

d) Discuta porque el algoritmo de $k - means$ converge, es decir, es seguro que de una iteración a la siguiente la función objetivo decrece?

3. (5 puntos) **Independencia lineal**

Considere una matriz de $A \in \mathbb{R}^{m\times n}$ con $m > n$ y un vector $b \in \mathbb{R}^m$. Buscamos el vector solución $x \in \mathbb{R}^n$ del sistema lineal $Ax = b$. Describa como puede asegurar de forma práctica si la solución del sistema existe o no.

Para el caso $n = 3$ y $m = 4$, considere el siguiente sistema de ecuaciones:

$$
\begin{aligned}
x + y + z &= 1,\\
2x - 3y + 4z &= 1,\\
3x + 5y - z &= 1,\\
2y - 4z &= -2.
\end{aligned}
$$

Aplique forma práctica propuesta anteriormente para determinar si el sistema tiene solución.

4. (15 puntos) **Clasificador binario: Perceptrón**

En este problema trataremos el problema de clasificación binaria, es decir, determinar si un dato pertenece a una de dos clases.

a) Programe el algoritmo del Perceptrón. Puede ser útil construir en este caso una rutina

```
Perceptron_fit(X, y, nitmax, eta)
```

que tome los datos `X`, sus etiquetas `y`, y que retorne los pesos `w` del clasificador. Otros hiperparámetros a considerar son `nitmax` y `eta` que corresponden respectivamente al número máximo de iteraciones del algoritmo y el *learning rate*.

b) Programe una rutina que dado un vector de pesos `w`, y un conjunto de datos `X` prediga en a que conjunto de datos pertenece cada dato.

c) Programe una rutina que dado un vector de pesos `w`, un conjunto de datos `X` y sus etiquetas `y`, entregue un score de que tan bien clasificados están los datos según las predicciones hechas por el clasificador.

d) Use las bases de datos `datos1.csv` y `datos2.csv`, grafique los datos identificando sus respectivas etiquetas y junto a ellos el resultado del hiperplano separador obtenido por su algoritmo de Perceptrón.

e) Cree 10 nuevos datos de la siguiente forma:

```
numpy.random.seed(18)
new_data = numpy.random.random(size = (10, 2))
new_data_labels = numpy.random.randint(2, size = 10)
```

Para el primer set de datos, prediga la clase a la que pertenecen cada uno de sus datos con la función de predicción creada y calcule el *score*. Grafique estos datos junto al conjunto inicial de datos y discuta la capacidad predictora de este clasificador.

5. (5 puntos) **Descomposición en valores singulares.**

Muestre que para $A \in \mathbb{R}^{m\times n}$ con valores singulares $\sigma_1, ..., \sigma_p$, $p = \min\{m, n\}$, su norma de Frobenius satisface la identidad

$$\|A\|_F = \sqrt{\sigma_1^2 + ... + \sigma_p^2}.$$

Verifique que la identidad anterior se cumple para

$$A = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix}$$

6. (15 puntos) **Análisis de componentes principales (PCA).**

   Considere nuevamente la base de datos de imágenes de dígitos de MNIST.

   a) Calcule la descomposicion en valores singulares de la matriz asociada a los datos utilizando la funcion `svd` de la librería `numpy.linalg`.

   b) Utilice la SVD para reducir la dimensión de los datos a la mitad, es decir realice un análisis de componentes principales.

   c) Grafique 10 imágenes y sus respectivas proyeccciones una vez realizado en el analisis de componentes principales.

   d) Para los datos proyectado calcule nuevamente su algoritmo de clustering y compare con los resultados obtenidos en la pregunta 2.