

TALLER DE PROCESAMIENTO DE SEÑALES

1er Cuatrimestre 2026 - Trabajo Práctico Nº 10 - Gaussian Variational Bayes

IMPORTANTE: En este ejercicio solamente se podrá importar `numpy`, `matplotlib.pyplot` y las funciones `gamma`, `digamma` y `softmax` de `scipy.special` (el resto deberá ser implementación propia).

Se desea modelar el tiempo (en minutos con respecto a la hora de entrada) en que diferentes grupos de estudiantes llegan a una determinada clase de la facultad. Estos comportamientos serán modelados como una mezcla de 6 comportamientos distintos dentro de las rutinas de las personas.

(a) *Creación del dataset:* Generar 100 muestras de una mezcla de gaussianas con pesos 0.1, 0.4, 0.2, 0.3, medias $-4, 0, 4, 5$ y varianzas 1, 1.96, 1.44, 1 respectivamente. *Interpretación:* En realidad existían 4 comportamientos marcados (aunque nosotros no lo sabíamos). Ellos pueden pensarse como:

- Un 10% de personas muy puntuales ($\mu = -4$ y varianza más pequeña).
- Un 40% de personas que tienden a llegar justo ($\mu = 0$).
- Un 20% de personas “relajadas” ($\mu = 4$).
- Un 30% de personas consistentemente impuntuales ($\mu = 5$ y varianza más pequeña).

(b) *K-means:*

1. Implementar un algoritmo de K-means para caracterizar la puntualidad de los estudiantes. El entrenamiento debe contener doble condición de parada (número de iteraciones y convergencia). El algoritmo debe ser definido dentro de una clase que posea al menos los métodos `init`, `fit` y `predict`.
2. Modelar una mezcla de gaussianas con los resultados del entrenamiento.
3. Graficar la densidad estimada y compararla con la original.

(c) *Expectation-Maximization:*

1. Implementar un algoritmo de EM para caracterizar la puntualidad de los estudiantes. Inicializar el algoritmo con el modelo de K-means. El entrenamiento debe contener doble condición de parada (número de iteraciones y convergencia). El algoritmo EM debe ser definido dentro de una clase que posea al menos los métodos `init`, `fit`, `predict_proba` y `predict`.
2. Graficar la densidad estimada, compararla con la de K-means y con la original.

(d) *Gaussian Variational Bayes:*

1. Implementar un Variational Bayes Gaussiano que permita computar el modelo. Suponer *a priori* $m = 0$, $\delta = \nu = \beta = 0.05$ y $\alpha = (1, 1, 1, 1, 1, 1)$, y utilizar el algoritmo EM para inicializar las probabilidades. El entrenamiento debe contener doble condición de parada (número de iteraciones y convergencia). El algoritmo debe ser definido dentro de una clase que posea al menos los métodos `init`, `fit`, `predict_proba` y `predict`.
2. Con la distribución *a posteriori* generar 3 muestras de parámetros y graficar la densidad de $X|\mu, \lambda, \pi$ para cada uno de esos conjuntos de parámetros. Compararla con la densidad verdadera, con la de K-means y con la del EM.
3. Incorporar al algoritmo anterior los métodos `predictive_predict_proba` y `predictive_predict`, utilizando la distribución predictiva en lugar de la distribución a posteriori.
4. Graficar la densidad *predictiva*. Compararla con la densidad verdadera, con la de K-means y con la del EM.