Tarea 2: Prior conjugado, MV y MAP

Profesor: Felipe Tobar
Auxiliares: José Díaz, Diego Garrido, Jou-Hui Ho, Luis Muñoz
Consultas: José Díaz, Jou-Hui Ho (U-cursos)
Período: 8/5/2020 — 17/5/2020

Formato entrega: Informe en formato PDF, con una extensión máxima de 3 planas para la P2 y P3 (puede usar un formato de doble columna). **Para la P1 incluya una foto de su desarrollo en formato PDF**. El desarrollo debe ser **ordenado y legible**. Adicionalmente debe entregar el jupyter notebook o el código utilizado. Si trabaja en parejas, realice solo un envío de la tarea.

P1. Prior conjugado (6.0 puntos)

Para el modelo gaussiano de una v.a. $X|\mu, \sigma^2 \sim \mathcal{N}(\mu, \sigma^2)$, encuentre el prior conjugado para $\theta = [\mu, \sigma^2]$.

P2. Máximo a posteriori (6.0 puntos)

En el procesamiento de señales de procesos aleatorios, generalmente se tiene que el valor de la señal depende de sus valores anteriores, donde nace la utilidad de los modelos basados en ecuaciones diferenciales. Esto es bajo el contexto de una señal continua. En el mundo discreto, se tiene su par análogo, donde un modelo se denomina **autoregresivo** cuando su variable de salida depende linealmente de sus valores anteriores. En particular, cuando solo depende del valor de una unidad de tiempo anterior, se habla del modelo **AR(1)**, dado por:

$$Y_t = c + \varphi Y_{t-1} + \epsilon_t$$

donde Y_t es la variable de salida en el tiempo t, ϵ_t es el ruido blanco del sistema en el tiempo t, modelado como una v.a. i.i.d, con $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$, y $|\varphi| < 1$.

En este tipo de modelos, se sabe además que:

$$\mathbb{E}(Y_t) = \frac{c}{1-\varphi}, \quad \mathbb{V}ar(Y_t) = \frac{\sigma^2}{1-\varphi^2}$$

(a) (0.5 puntos) Sea $\theta = [c, \varphi, \sigma^2]^T$ el vector de parámetros (desconocidos) del modelo, calcule la verosimilitud sobre la observación Y_1. Considere condición inicial nula conocida.

(b) (0.8 puntos) Dado que se conoce la observación en $t = 1$, explicite la distribución de Y_2. Con esto, calcule $L(\theta; Y_2|Y_1)$, es decir, la verosimilitud sobre Y_2, dado que se conoce el valor de Y_1.

(c) (0.7 puntos) Utilizando los resultados anteriores, calcule la verosimilitud sobre las primeras dos observaciones: (Y_1, Y_2). Con esto, generalice la expresión de la verosimilitud obtenida para t observaciones: $(Y_1, Y_2 \dots Y_t)$.
Nota: Para simplificar la notación, puede dejar expresados los términos que ya se conocen.

(d) (1.0 puntos) Genere una muestra de tamaño $T = 100$ del modelo descrito, utilizando $c = 1, \varphi = 0.5$ y $\sigma^2 = 1$. Implemente la estimación por MV de los parámetros del modelo.
Nota: Puede utilizar funciones de optimización predefinidas de Python, pero usted debe definir la función la minimizar (que en este caso es la verosimilitud). Recuerde setear la semilla de aleatoriedad para que sus resultados sean replicables.

(e) (1.0 puntos) Suponga que no se tiene ninguna información sobre los parámetros, salvo el rango en el que los parámetros pueden tomar valores: $c \in [-3, 3], |\varphi| < 1, \sigma^2 \in [0.5, 2]$. Obtenga el estimador MAP de θ.

(f) (1.0 puntos) Para cada parámetro, un experto le indica el valor probable cerca del cual puede estar. Proponga una nueva distribución a priori y repita el paso anterior. Justifique las distribuciones a priori escogidas.

(g) (1.0 puntos) Como en realidad los datos son sintéticos, usted conoce θ real. Compare las tres estimaciones e interprete los resultados. Sin embargo, en el mundo real, θ es desconocido, entonces ¿cómo podría evaluar la calidad de la estimación realizada? Impleméntela si es posible.

P3. Proyecto del curso

(i) Forme grupos de a lo más **3** personas (Si es de postgrado, máximo 1). Elijan una propuesta de proyecto y descríbalo.

(ii) Sobre el proyecto que piensa realizar, muestre y describa los datos que piensa utilizar.