
Recuperación de imágenes usando modelos auto-regresivos condicionales: CAR e IAR

Image recovery using conditional autoregressive models: CAR and IAR.

Danna Lesley Cruz Reyes^a
danna.cruz@urosario.edu.co

Resumen

Este artículo realiza la estimación Bayesiana de campos aleatorios gaussianos de Markov. En particular, se propone realizar un análisis de dependencia espacial por medio de un grafo que caracteriza las intensidades observadas de una imagen con un modelo ampliamente utilizado en estadística espacial y geoestadística conocido como modelo autorregresivo condicional (CAR por sus siglas en inglés). Este modelo es útil para obtener distribuciones conjuntas multivariadas de un vector aleatorio basado en especificaciones condicionales univariadas. Estas especificaciones condicionales se basan en las propiedades de Markov, de modo que la distribución condicional de un componente del vector aleatorio depende solo de un conjunto de vecinos, definido por el grafo. Los modelos autorregresivos condicionales son casos particulares de campos aleatorios de Markov y se utilizan como distribuciones *a priori*, que combinadas con la información contenida en los datos de la muestra (función de verosimilitud), inducen una distribución *a posteriori* en las que se basa la estimación. El modelo CAR tiene un caso particular llamado IAR, en el cual, la distribución *a priori* no es propia. En este artículo se aplica ambos modelos haciendo una comparación entre ellos. Todos los parámetros del modelo se estiman en un entorno completamente Bayesiano, utilizando el algoritmo Metropolis-Hastings. Los procedimientos completos de estimación posterior se ilustran y comparan utilizando varios ejemplos artificiales. Para estos experimentos, el modelo CAR y el modelo IAR se comporta muy favorablemente con imágenes homogéneas.

Palabras clave: Procesamiento probabilístico de imágenes, restauración de imágenes, modelado Bayesiano, modelo autorregresivo condicional.

Abstract

^aEscuela de Medicina y Ciencias de la Salud, Universidad del Rosario

This article performs Bayesian estimation of Gaussian Markov random fields. In particular, it is proposed to perform a spatial dependency analysis by means of a graph that characterizes the observed intensities of an image with a model widely used in spatial statistics and geostatistics known as the conditional autoregressive model (CAR). This model is useful for obtaining multivariate joint distributions from a random vector based on univariate conditional specifications. These conditional specifications are based on the Markov properties, so that the conditional distribution of a component of the random vector depends only on a set of neighbors, defined by the graph. Conditional autoregressive models are particular cases of random Markov fields and are used as *a priori* distributions, which, combined with the information contained in the sample data (likelihood function), induce a *a posteriori* distribution on which the estimate is based. The CAR model has a particular case called IAR, in which the *a priori* distribution is not proper, in this article both models are applied making a comparison between them. All model parameters are estimated in a completely Bayesian environment, using the Metropolis-Hastings algorithm. The complete estimation procedures are illustrated and compared using various artificial examples. For these experiments, the CAR model and the IAR model performed very favorably with homogeneous images.

Keywords: Probabilistic image processing, image restoration, Bayesian modeling, conditional autoregressive model.

1. Introducción

Los modelos estadísticos para analizar datos espaciales se dividen en dos clases generales: modelos geoestadísticos con soporte espacial continuo y modelos en una lattice, también llamados datos de área, donde los datos se producen en una cuadrícula (posiblemente irregular), con un conjunto enumerable de vértices o ubicaciones. Estos modelos autoregresivos se utilizan en muchos campos, incluyendo la cartografía de las tasas de infecciones, (Elliott & Wartenberg 2004), agricultura (Besag & Higdon 1999), econometría (LeSage & Thomas-Agnan 2015), ecología (Arslan & Akyürek 2018). En este trabajo, se presenta el modelo CAR como un ejemplo de los campos aleatorios gaussianos de Markov. Varios autores han analizado este modelo, mostrando las características de las covarianzas y las correlaciones dada una estructura espacial, por ejemplo, el modelo CAR produce variaciones no constantes en cada sitio, así como covarianzas desiguales entre regiones separadas por el mismo número de vecinos (ver Besag (1986), Assunção & Krainski (2009)). Wall (2004) estudió ampliamente la estructura de covarianza *a priori* que conlleva el modelo CAR. Encontró que la relación entre vecinos no parece tener explicación, ya que nodos con el mismo número de vecinos y en vecindades similares tienen covarianzas diferentes. Luego, Assunção & Krainski (2009) en su artículo esclarece todos estas rarezas y concluye que el modelo CAR es afectado por sus vecinos de mayor orden.

Por otro lado, los campos aleatorios de Markov (GMRF) se han utilizado ampliamente para el análisis de imágenes durante muchos años, (Xing et al. (2019), Liu & Cooper (2010), Besag (1986), Cross & Jain (1983), Qian & Titterton (1991), Chen & Huang (1993), Horiguchi et al. (1997)) . Uno de los modelos mas utilizados basados en los GMRF, es el modelo Ising, sin embargo, el modelo de Ising (Halim (2008), Mao & Jain (1992), Van Leemput et al. (1999)), funciona cuando la imagen es blanco y negro. Esto significa que el modelo de Ising no puede hacer diferencias de más de dos valores de gris. Por lo tanto, según Halim (2008) debería realizarse una modificación adicional para las imágenes con colores. Morris et al. (1996) afirma que a pesar de que los modelos de Ising han sido utilizado desde el primer trabajo en la segmentación de imágenes basada en MRF, los avances en las técnicas de Monte Carlo han resaltado las deficiencias de estos modelos. Por tanto, se presenta como alternativa, el modelo CAR.

El modelo CAR visualiza el dominio de la imagen como un grafo no dirigido con un vértice en cada pixel y una arista entre dos vértices si los pixeles correspondientes son vecinos. Esto crea vecinos bien definidos que se utilizan para definir la distribución conjunta o condicional. La distribución será la distribución normal multivariada.

El sistema de vecindad es un punto clave en los modelos autorregresivos condicionales o modelos CAR que se usan comúnmente en estadísticas espaciales. Para este caso, los grafos que apoyan la construcción del GMRF serán aquellos que expresen estas estructuras de vecindad. En este contexto, las aristas \mathcal{E} en el grafo $\mathcal{B} = (\mathcal{G}, \mathcal{E})$, representan las conexiones en la estructura de la imagen y, en consecuencia, definen los vecinos que se utiliza para modelar la dependencia espacial.

El término condicional, en el modelo CAR se usa porque cada elemento del proceso aleatorio se especifica condicionalmente en los valores de los nodos vecinos. En este artículo se presenta una descripción más detallada de la estructura implícita de este modelo, y en particular, se presenta la forma en cómo se puede aplicar para recuperar imágenes. La Sección 2 presenta el concepto de campo gaussiano de Markov y como un ejemplo de este, el modelo CAR. La Sección 3 define el modelo que será utilizado para la recuperación de imagen. La Sección 4 muestra algunos ejemplos y experimentos. Las conclusiones y la discusión final se encuentran en la Sección 5.

2. Campos aleatorios gaussianos de Markov

Los campos aleatorios son distribuciones multivariadas que, en general, se usan para describir la asociación espacial entre las variables θ . Un campo aleatorio de Markov extiende el concepto de cadena de Markov a un contexto espacial y supone que dicha distribución conjunta de θ esta definida como sigue: sea un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ con n vértices \mathcal{V} donde cada uno representa una de las componentes del vector $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ y \mathcal{E} el conjunto de aristas que conecta dos vértices θ_i e θ_j . Así,

$$f(\theta_i|\boldsymbol{\theta}_{-i}) = f(\theta_i|\boldsymbol{\theta}_{j\sim i}),$$

con $\boldsymbol{\theta}_{j\sim i}$ el vector formado por todos los componentes de $\boldsymbol{\theta}$ quienes son vecinos de i . En esta sección se discute brevemente los campos aleatorios gaussianos de Markov que a menudo se usan como una distribución *a priori* para efectos espaciales.

Un GMRF es un campo de Markov donde la distribución del vector aleatorio (de dimensión finita) y una distribución normal o gaussiana satisfacen los supuestos de independencia condicional. Se puede encontrar una discusión detallada de GMRF en Rue & Held (2005) y Assunção & Krainski (2009).

Todos los resultados válidos para la distribución normal también serán válidos para un GMRF. Por lo tanto, en la siguiente sección se presenta los resultados más relevantes de la distribución normal multivariante. Después de definir formalmente un GMRF con todas las propiedades heredadas de la distribución normal, se presenta la conexión entre el grafo \mathcal{G} y los parámetros de la distribución normal multivariada $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$. Se mostrará que toda la información en el grafo se condensa en la matriz de covarianza $\boldsymbol{\Sigma}$ por medio de la matriz de precisión $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}$, además, el vector promedio $\boldsymbol{\mu}$ no influirá en la estructura del vecindario del grafo.

2.1. La distribución normal multivariante.

Para facilitar la comprensión de los campos aleatorios gaussianos de Markov, se revisa la distribución normal multivariante y algunas de sus propiedades básicas.

Un vector aleatorio n -dimensional $\boldsymbol{\theta}_{n \times 1} = (\theta_1, \theta_2, \dots, \theta_n)^t$, $n < \infty$ tiene una distribución n -variada con vector de medias $\boldsymbol{\mu}_{n \times 1}$ y matriz de covarianza $\boldsymbol{\Sigma}_{n \times n}$ si su función de densidad de probabilidad toma la siguiente forma:

$$f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta} - \boldsymbol{\mu})\right\}, \quad \boldsymbol{\theta} \in \mathbb{R}^n. \quad (1)$$

Esta distribución será denotada por $\boldsymbol{\theta} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ donde $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$ son tales que $\mu_i = E(\theta_i)$, $\boldsymbol{\Sigma}_{ij} = Cov(\theta_i, \theta_j)$, $\boldsymbol{\Sigma}_{ii} = Var(\theta_i)$ y $Corr(\theta_i, \theta_j) = \boldsymbol{\Sigma}_{ij}(\boldsymbol{\Sigma}_{ii}\boldsymbol{\Sigma}_{jj})^{-1/2}$.

Para presentar algunas propiedades de la distribución dada en (1), se considera la siguiente partición: $\boldsymbol{\theta} = (\boldsymbol{\theta}_A, \boldsymbol{\theta}_B)^t$, $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$, y

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{pmatrix}, \quad \begin{bmatrix} \boldsymbol{\Sigma}_{AA} & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}_{AB} & \boldsymbol{\Sigma}_{BB} \end{bmatrix}$$

donde $\#(A) + \#(B) = n$. Suponiendo tal partición, algunas propiedades básicas de la distribución normal son:

- $\boldsymbol{\theta}_A \sim N(\boldsymbol{\mu}_A, \boldsymbol{\Sigma}_{AA})$ es la distribución marginal del vector $\boldsymbol{\theta}_A$ de orden $A \times 1$,
- $\boldsymbol{\Sigma}_{AB} = 0$ si y solo si $\boldsymbol{\theta}_A$ e $\boldsymbol{\theta}_B$ son independientes,
- la distribución condicional de $\boldsymbol{\theta}_A$ dado $\boldsymbol{\theta}_B$ es $N(\boldsymbol{\mu}_{A|B}, \boldsymbol{\Sigma}_{A|B})$ donde,

$$\boldsymbol{\mu}_{A|B} = \boldsymbol{\mu}_A + \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_{BB}^{-1}(\boldsymbol{\theta}_B - \boldsymbol{\mu}_B) \quad \text{y} \quad (2)$$

$$\boldsymbol{\Sigma}_{A|B} = \boldsymbol{\Sigma}_{AA} - \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_{BB}^{-1}\boldsymbol{\Sigma}_{BA}; \quad (3)$$

- Si $\boldsymbol{\theta} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y $\boldsymbol{\theta}_1 \sim N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$, entonces $\boldsymbol{\theta} + \boldsymbol{\theta}_1 \sim N(\boldsymbol{\mu} + \boldsymbol{\mu}_1, \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_1)$.

2.2. Definición básica y propiedades de GMRF

Para construir un GMRF se considera un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ con vértices n donde cada vértice representa uno de los componentes del vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)$ y las aristas conectan nodos que tienen algún tipo de asociación. Un GMRF supone que $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)^t \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y que las aristas del grafo conectan nodos i y j si y solo si $\theta_i \not\perp \theta_j | \boldsymbol{\theta}_{-ij}$, es decir, si θ_i es independiente de θ_j , dados los componentes de $\boldsymbol{\theta}$ excepto θ_i y θ_j .

TEOREMA 1. Si $\boldsymbol{\theta} \sim N(\boldsymbol{\mu}, \mathbf{Q})$, entonces para $i \neq j$,

$$\theta_i \perp \theta_j | \boldsymbol{\theta}_{-ij} \Leftrightarrow Q_{ij} = 0.$$

La demostración se puede encontrar en Rue & Held (2005). Por lo tanto, este resultado establece que los componentes no nulos de \mathbf{Q} determinan la relación de vecindad presente en \mathcal{G} . Esto implica que cualquier distribución normal con una matriz de covarianza definida positivamente también es un GMRF y viceversa. Por lo tanto, un GMRF se define formalmente de la siguiente manera:

DEFINICIÓN 1. Un vector aleatorio $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)^t \in \mathbb{R}^n$ es llamado GMRF correspondiente a un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ con media $\boldsymbol{\mu}$ y matriz de precisión $\mathbf{Q} > 0$, si y solamente si la función de densidad de probabilidad de $\boldsymbol{\theta}$ tiene la siguiente forma:

$$\pi(\boldsymbol{\theta}) = (2\pi)^{-n/2} |\mathbf{Q}|^{1/2} \exp\left(-\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\mu})^t \mathbf{Q}(\boldsymbol{\theta} - \boldsymbol{\mu})\right),$$

donde la matriz \mathbf{Q} cumple la siguiente condición:

$$Q_{ij} \neq 0 \Leftrightarrow \{i, j\} \in \mathcal{E}, \forall i \neq j.$$

Si \mathbf{Q} es una matriz densa, entonces \mathcal{G} está completamente conectado, es decir, el vértice está conectado a todos los demás vértices del grafo. En este artículo, se considera el caso donde \mathbf{Q} es esparsa, es decir, la mayoría de las componentes de la matriz son cero.

TEOREMA 2. *Sea un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ que representa $\boldsymbol{\theta}$ un GMRF, con media $\boldsymbol{\mu}$ y la matriz de precisión \mathbf{Q} simétrica y definitiva positiva. Luego, la distribución de cada componente θ_i de $\boldsymbol{\theta}$, dado el vector $\boldsymbol{\theta}_{-i}$ formado por todos los componentes de $\boldsymbol{\theta}$, excepto θ_i es una distribución normal tal que:*

$$\begin{aligned} \text{E}(\theta_i|\boldsymbol{\theta}_{-i}) &= \mu_i - \frac{1}{Q_{ii}} \sum_{j:j\sim i} Q_{ij}(\theta_j - \mu_j), \\ \text{Prec}(\theta_i|\boldsymbol{\theta}_{-i}) &= Q_{ii}, \\ \text{Corr}(\theta_i, \theta_j|\boldsymbol{\theta}_{-ij}) &= -\frac{Q_{ij}}{\sqrt{Q_{ii}Q_{jj}}}, \quad i \neq j, \end{aligned}$$

donde $i \sim j$ denota que el nodo j es vecino de nodo i .

2.3. Especificación de GMRF a través de condicionales completas

Una alternativa a la construcción de un GMRF es considerar las condicionales completas de las distribuciones $\pi(\theta_i|\boldsymbol{\theta}_{-i})$. Este enfoque pionero se adoptó en Besag (1974) para construir los modelos autorregresivos condicionales, conocidos como modelos CAR. Suponga que la distribución condicional completa de θ_i dada $\boldsymbol{\theta}_{-i}$ para todos $i = 1, \dots, n$ es una distribución normal con media y precisión dadas respectivamente por:

$$\text{E}(\theta_i|\boldsymbol{\theta}_{-i}) = \mu_i - \sum_{j:j\sim i} \beta_{ij}(\theta_j - \mu_j), \quad (4)$$

$$\text{Prec}(\theta_i|\boldsymbol{\theta}_{-i}) = \kappa_i, \quad (5)$$

con $\kappa_i > 0$, donde la suma en (4) está sobre todos los j de i vecinos. El siguiente teorema nos muestra que esta densidad conjunta de $\boldsymbol{\theta}$ existe y que es única.

TEOREMA 3. *Si θ_i dado $\boldsymbol{\theta}_{-i}$, $\forall i = 1, \dots, n$, tiene una distribución normal, con media y precisión dadas en (4) y (5), por lo que la distribución conjunta de $\boldsymbol{\theta}$ es un GMRF con vector de medias $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)$ y matriz de precisión \mathbf{Q} , entonces*

$$Q_{ij} = \begin{cases} \kappa_i \beta_{ij} & i \neq j \\ \kappa_i & i = j \end{cases} \quad (6)$$

en que $\kappa_i \beta_{ij} = \kappa_j \beta_{ji}$, $i \neq j$.

La demostración de este resultado se puede encontrar en Rue & Held (2005). La Expansión Brook se muestra a continuación.

TEOREMA 4. Si $f(\theta_1|\theta_{-1}), f(\theta_2|\theta_{-2}), \dots, f(\theta_n|\theta_{-n})$ son las n distribuciones condicionales completas de una densidad conjunta $f(\theta)$ y si \mathbf{x} es una configuración de referencia cualquiera tal que $f(\mathbf{x}) > 0$, entonces la densidad conjunta $f(\theta)$ se puede obtener por medio de la siguiente relación:

$$\begin{aligned} \frac{f(\theta)}{f(\mathbf{x})} &= \frac{f(\theta_1|\theta_2, \dots, \theta_n) f(\theta_2|x_1, \theta_3, \dots, \theta_n)}{f(x_1|\theta_2, \dots, \theta_n) f(x_2|x_1, \theta_3, \dots, \theta_n)} \cdots \frac{f(\theta_n|x_1, \dots, x_{n-1})}{f(x_n|x_1, \dots, x_{n-1})} \\ &= \prod_{i=1}^n \frac{f(\theta_i|\theta_1, \dots, \theta_{i-1}, x_{i+1}, \dots, x_n)}{f(x_i|\theta_1, \dots, \theta_{i-1}, x_{i+1}, \dots, x_n)}. \end{aligned}$$

La demostración se puede encontrar en Rue & Held (2005) .

2.4. Un ejemplo de Campos de Markov Gaussianos: el modelo CAR

Considere una región espacial que está dividida en subregiones n indexadas por enteros $1, 2, \dots, n$. Suponga que esta colección de subregiones está dotada de un sistema de vecindad $\{V_i : i : 1, \dots, n\}$, donde V_i denota la colección de subregiones que, en un sentido bien definido, son vecinos de la subregión i . En términos espaciales,

$$V_i = \{j : \text{subregiones } i \text{ y } j \text{ que comparten borde}\}, \quad \text{para } i \in \{1, 2, \dots, n\},$$

El modelo CAR supone que las distribuciones condicionales completas para efectos aleatorios correspondientes a las variables de respuesta para cada y_1, \dots, y_n , definidas como $\theta_1, \dots, \theta_n$ y dadas por $\theta = \theta_i|\theta_{-i}$, $i = 1, \dots, n$, siguen distribuciones normales con media y precisión dadas respectivamente en (4) y (5). Suponiendo que $\beta_{ij} = \rho/d_i$, si $i \sim j$, y $\beta_{ij} = 0$ caso contrario en que $\kappa_i = d_i/\sigma^2$, esto es,

$$\theta_i|\theta_{-i} \sim N(\mu_i + \rho\bar{\theta}_i, \frac{\sigma^2}{d_i}), \quad (7)$$

donde σ^2/d_i es la variante condicional de $\theta_i|\theta_{-i}$, ρ es una constante de proporcionalidad, d_i es el número de vecinos del nodo i en el grafo \mathcal{G} , la media de los vecinos del nodo i es $\bar{\theta}_i = \sum_{\mathcal{E}} (d_i^{\mathcal{G}})^{-1}(\theta_j)$ y $\mathcal{E} = \{(i, j) \in E(\mathcal{G}) : j \sim i\}$ es el conjunto de aristas que pertenecen al grafo \mathcal{G} .

Sea la matriz de adyacencia:

$$A = \begin{cases} 0, & \text{si } i = j \\ 1, & \text{si } i \sim j \\ 0, & \text{si } i \not\sim j \end{cases}$$

tome $M = \text{diag}\{d_1, d_2, \dots, d_n\}$. No es inmediato que (7) conduzca a una distribución conjunta completa de $\boldsymbol{\theta}$. Besag (1974) usa la expansión de Brook y muestra que cuando la matriz $(M - \rho A)^{-1}$ es definida positiva y simétrica, la distribución conjunta de $\boldsymbol{\theta}$ es:

$$\boldsymbol{\theta} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{\Sigma}), \quad (8)$$

donde $\boldsymbol{\Sigma} = (M - \rho A)^{-1}$. Para que la matriz de covarianza sea definida positiva, es necesario que $\rho < \frac{1}{\lambda_1}$ donde λ_1 es el valor propio más pequeño de la matriz $M^{-1/2} A M^{-1/2}$. La demostración se puede encontrar en Banerjee et al. (2004).

En resumen, el modelo CAR está definido por una estructura de correlación inducida por el grafo de la siguiente forma: la región de estudio V se divide en n unidades de área sobre el conjunto de regiones $\{V_i : i : 1, \dots, n\}$ que están vinculados a un conjunto correspondiente de respuestas $\mathbf{y} = (y_1, y_2, \dots, y_n)$ distribuidas condicionalmente como:

$$y_j | \theta_j, \sigma_y \sim N(\mu + \theta_j, \sigma_y), \quad \text{y} \quad (\theta_1, \dots, \theta_n)^t \sim N(\mathbf{0}, \sigma^2 \boldsymbol{\Sigma}), \quad \text{para cada } j \in \{1, \dots, n\}.$$

Así, el patrón espacial en la respuesta está modelado por el vector $\boldsymbol{\theta}$. Si el valor de $\rho = 1$, se conoce como modelo autorregresivo impropio, ya que si $\rho = 1$, la matriz $\boldsymbol{\Sigma}$ definida en (8) no posee inversa, lo que hace que la distribución normal sea impropia. El modelo ICAR fue propuesto por Besag & Kooperberg (1995). En este artículo se compara el comportamiento de los dos modelos, el CAR y el IAR por medio de la recuperación de una imagen. Se verifica que modelo se comporta mejor calculando los valores DIC y el ECM de cada uno de ellos.

3. Recuperación de una imagen

El modelo aquí propuesto se define de la siguiente manera: suponga que un espacio bidimensional se divide en n píxeles, etiquetados por los números enteros $1, 2, \dots, n$. Las ubicaciones de los píxeles formarán una red cuadrada regular. Suponga además que cada píxel es una variable aleatoria θ_i , $1 \leq i \leq n$ que puede tomar cualquier valor real, por tanto $\theta_i \in \mathbb{R}$. Los valores de las variables que representan píxeles se denominan intensidades.

Un ruido arbitrario se indicará con $\boldsymbol{\epsilon} = \{\epsilon_1, \epsilon_2, \dots, \epsilon_n\}$, así $\boldsymbol{\epsilon} \in \mathbb{R}^n$. En general, no es posible observar $\boldsymbol{\theta}$ directamente, sino que la imagen observada \mathbf{y} es una copia degradada de $\boldsymbol{\theta}$,

$$\mathbf{y} = \boldsymbol{\beta} + \boldsymbol{\theta} + \boldsymbol{\epsilon} \quad (9)$$

donde β es una constante y para cada $i \in \{1, \dots, n\}$, $\epsilon \sim N(0, \sigma_y^{-1})$ y ϵ_i e ϵ_j son independientes cuando $i \neq j$. Entonces, el vector $\mathbf{y} = (y_1, \dots, y_n)$ tiene como función de verosimilitud:

$$L(\boldsymbol{\theta}, \sigma, \sigma_y \rho, \mathbf{y}) = (2\pi\sigma_y^2)^{-n/2} \exp \left\{ -\frac{1}{2}(\mathbf{y} - \boldsymbol{\theta} - \beta)^t(\mathbf{y} - \boldsymbol{\theta} - \beta) \right\}$$

3.1. Formulación Bayesiana

Considerando el modelo definido en (9) para las intensidades $i = 1, \dots, n$, asignamos al vector $\boldsymbol{\theta}$ la distribución *a priori* correspondiente al modelo CAR o IAR, tal que,

$$(\theta_1, \dots, \theta_n)^t \sim N(\mathbf{0}, \sigma^2 \boldsymbol{\Sigma}), \quad \text{para cada } j \in \{1, \dots, N\},$$

donde $\boldsymbol{\Sigma}$ es definida como en (8), con esto, se presume que existe una correlación espacial entre los datos y que el efecto aleatorio ayuda a explicar este comportamiento. Para los demás parámetros, se suponen distribuciones *a priori* no informativas: $\rho \sim \text{Unif}(0, 1)$, $\sigma_y \sim \text{Gamma}(\alpha, \eta)$ $\sigma \sim \text{Gamma}(\alpha_0, \eta_0)$ y $\beta \sim N(\mu_0, \sigma_0^2)$.

Las distribuciones condicionales *a posteriori* para σ y σ_y tiene forma conocida. Específicamente,

$$\sigma_y^{-1} | \boldsymbol{\theta}, \sigma, \beta, \rho, \boldsymbol{\theta}, \boldsymbol{\Sigma}, \beta, \tau_{\boldsymbol{\theta}}, \gamma, \mathbf{y} \sim \text{Gamma}(\alpha + n/2, \eta + \frac{1}{2}(\mathbf{y} - (\beta + \boldsymbol{\theta}))^t(\mathbf{y} - (\beta + \boldsymbol{\theta})))$$

mientras que,

$$\sigma^{-1} | \boldsymbol{\theta}, \sigma_y, \beta, \rho, \boldsymbol{\theta}, \boldsymbol{\Sigma}, \beta, \tau_{\boldsymbol{\theta}}, \gamma, \mathbf{y} \sim \text{Gamma}(\alpha_0 + n/2, \eta_0 + \frac{1}{2}\boldsymbol{\theta}^t \boldsymbol{\Sigma} \boldsymbol{\theta}).$$

Lo mismo ocurre para β y $\boldsymbol{\theta}$; de forma que la distribución condicional completa para β es $N(\boldsymbol{\mu}_\beta, \boldsymbol{\sigma}_\beta^2)$ en donde,

$$\begin{aligned} \boldsymbol{\mu}_\beta &= \left(\frac{n}{\sigma_y^2} + \frac{1}{\sigma^2} \right)^{-1} \left(\sum_{i=1}^n \frac{y_i - \theta_i}{\sigma_y^2} + \frac{\mu_0}{\sigma_0^2} \right), \\ \boldsymbol{\sigma}_\beta^2 &= \left(\frac{n}{\sigma_y^2} + \frac{1}{\sigma^2} \right)^{-1}. \end{aligned}$$

La distribución condicional completa de $\boldsymbol{\theta} | \sigma, \sigma_y, \beta, \rho, \mathbf{y}$ es

$$\boldsymbol{\theta} | \sigma, \sigma_y, \beta, \rho, \mathbf{y} \sim N(\sigma_y^{-1} \boldsymbol{\Phi}^{-1}(\mathbf{y} - \beta), \boldsymbol{\Phi}^{-1}).$$

Con $\Phi = \sigma_y^{-1} \mathbf{I} + (\sigma)^{-1} \Sigma^{-1}$. La forma de la distribución condicional completa para ρ no coincide con alguna conocida. Por esta razón, la estimación de los parámetros se realiza mediante la obtención de muestras de la distribución *a posteriori* mediante métodos MCMC. El algoritmo completo para estimar los parámetros del modelo es un *Gibbs* con pasos de *Metropolis*. Para muestrear la distribución posterior de ρ se hace una reparametrización $U = \text{logit}(\rho)$. Este método de transformar el espacio de parámetros dentro del algoritmo Metropolis-Hastings es útil para problemas con espacios de parámetros restringidos, en este caso $\rho < \frac{1}{\lambda_1}$, (Givens & Hoeting 2012). Una vez que se ha generado la pseudomuestra a partir de la distribución *a posteriori*, se dispone de varios estimadores posibles, una opción es la media posterior, que puede estimarse mediante la media muestral de la pseudomuestra.

4. Aplicación

El ejemplo que se utiliza para recuperar una imagen utilizando un campo aleatorio gaussiano es la imagen dada en la Figura (1a). Para generar una imagen ruidosa, se le agrega a cada intensidad un ruido gaussiano independiente con $\sigma_y^2 = 25$, la imagen resultante es la que se muestra en la Figura (1b). En estos análisis, el ojo humano puede hacer un buen trabajo delimitando los límites y las intensidades de la imagen.

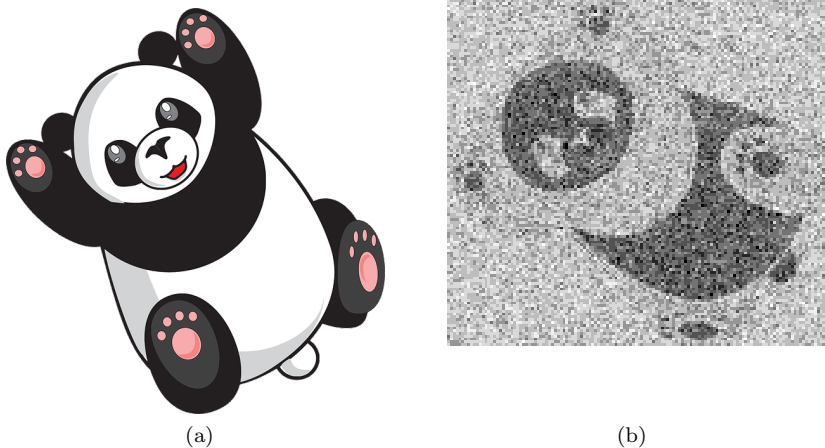


Figura 1: Ejemplo y datos (Imagen de OpenClipart-Vectors en Pixabay). La superficie de intensidad real se muestra en (a) y los datos simulados con ruido gaussiano correspondientes en (b).

	β	σ_y^2	DIC	ECM
CAR	0.285	22.732	-95122.38	$5.014e - 07$
IAR ($\rho = 1$)	0.285	22.732	-94544.02	$4.861e - 07$

Tabla 1: Estimativas posteriores de AIC y ECM.

La imagen tiene un tamaño de 134×125 píxeles. Esta imagen genera una red cuadrada regular de tamaño 16750×16750 , la cual, será la que caracterizará la matriz de adjacencia \mathbf{A} y la matriz de vecindades \mathbf{M} de la matriz de covarianza de la distribución normal definida en 8. Para el análisis, ejecutamos el MCMC durante 1000 iteraciones, con un periodo de *burn-in* de 500. Algo notable en estos modelos, es que son necesarias pocas iteraciones para obtener convergencia en la cadena.

La Tabla 1 muestra las medias *a posteriori* usando los modelos CAR e IAR para los parámetros β y σ_y^2 . Para comparar los modelos, se calculó el valor de DIC y el valor ECM (error cuadrático medio). La estimación del parámetro β y σ_y^2 fue de 0.285 y 22.732 aproximadamente la misma para ambos modelos. La estimación del parámetro ρ fue de 0.89, lo que podría indicar una correlación espacial fuerte ente píxeles vecinos. La estimación de σ_y^2 está cerca del verdadero valor del ruido 25, que fue incluido al realizar la simulación de los datos.

La Figura 2 muestra las cadenas e histogramas generados para los parámetros β , σ_y y para un sólo componente del vector θ por el método MCMC para el modelo CAR (no se puede mostrar las 16750 cadenas correspondientes a cada componente del vector θ). En cada uno de ellos hay un período *burn-in* de 500 iteraciones después de eso, se asume que la cadena de Markov está en equilibrio y las reconstrucciones subsecuentes son de hecho la distribución subsecuente necesaria. Las Figuras 2b y 2d muestran las densidades marginales estimadas para β y θ_1 , mostrando que son simétricas y muy estrechas. La densidad de la Figura 2f proviene del parámetro σ_y^2 , es casi simétrica y razonablemente estrecha.

La Figura 3 muestra la media posteriori del vector θ para cada modelo, es decir, la reconstrucción de la imagen. Cada uno de los componentes de este vector, representa la intensidad *a posteriori*, es decir, cada uno de los píxeles recuperados. En la Figura 3b muestra la reconstrucción siguiendo el modelo CAR. Note que los resultados son visualmente buenos sin suavizar excesivamente las discontinuidades. En la Figura 3c se muestra la reconstrucción utilizando el modelo IAR. La reconstrucción de intensidad es muy similar que la del modelo CAR. Los resultados son muy similares, lo que no nos permite concluir cual de los dos modelos es mejor.

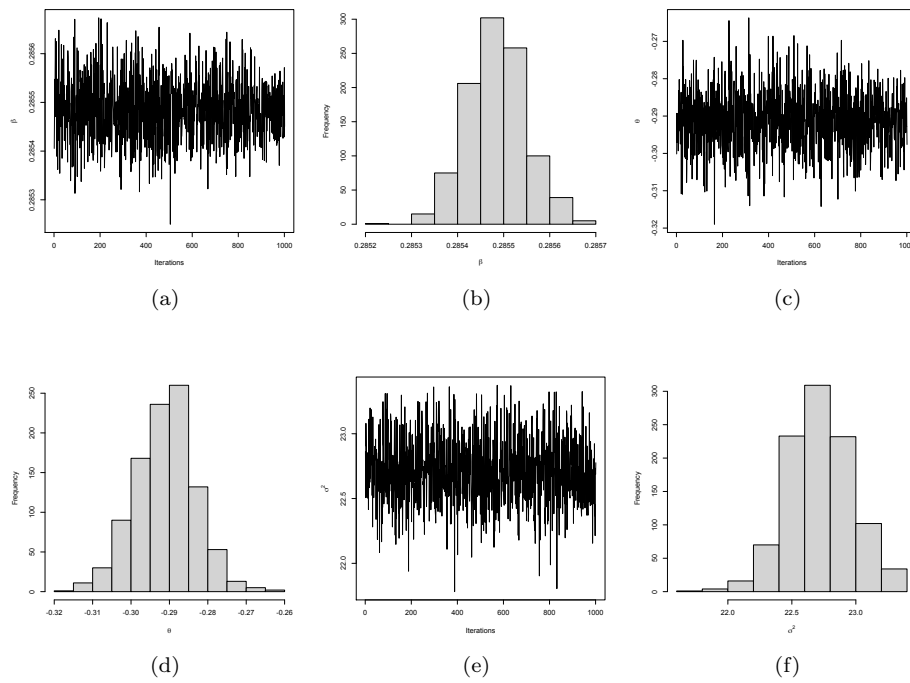


Figura 2: Ejemplos de estadísticas de seguimiento y densidades posteriores. Los histogramas de valores muestreados estiman densidades marginales de probabilidad posterior para β y σ_y^2 .

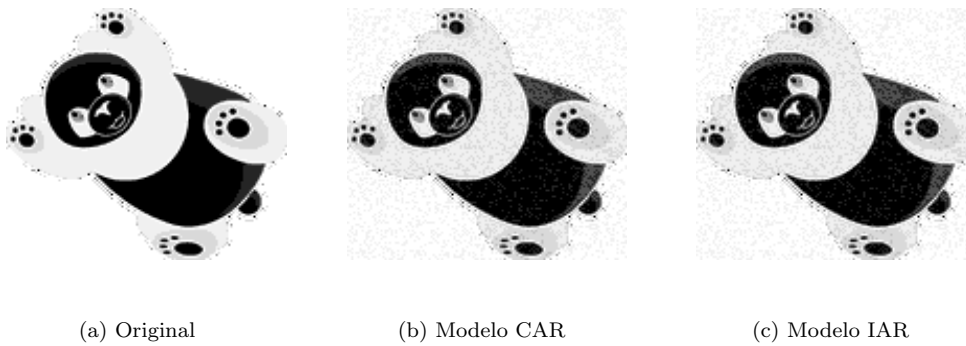


Figura 3: Reconstrucciones con el modelo CAR (b) y el modelo IAR (c) de la superficie original en la Figura (a).

5. Conclusiones

En este estudio, se propone y evalúa un método de eliminación de ruido para imágenes que siguen una distribución normal y un GMRF para las variables latentes. Aunque los resultados son generalmente buenos, es difícil interpretar el parámetro ρ ya que puede ocurrir que incluso valores muy altos de ρ inducen solo una modesta correlación espacial entre observaciones (ver Banerjee et al. (2004)).

A pesar de que el modelo CAR es ampliamente utilizado, tiene algunos aspectos no intuitivos. Para citar un caso, Wall (2004) estudió una estructura de covarianza *a priori* implícita en el modelo CAR; este estudio concluyó que la correlación espacial implícita entre diferentes ubicaciones utilizando un modelo CAR no parece seguir un esquema intuitivo, por ejemplo, Wall (2004) muestra que cuando ρ es negativo, las correlaciones entre áreas vecinas pueden ser positivas o negativas, sin importar el signo de ρ . Sin embargo, Assunção & Krainski (2009) explica cómo las consecuencias aparentemente contradictorias o poco prácticas, propuestas en el artículo de Wall (2004) se debe a que el modelo CAR utiliza la estructura completa del grafo de vecindad, no solo de vecindad inmediata. A través de los resultados del álgebra lineal, Assunção & Krainski (2009) descomponen la matriz de covarianza del modelo CAR *a priori* y *a posteriori* en una suma infinita de términos, esta descomposición aclara cómo toda la estructura de la vecindad del mapa afecta la covarianza de algunas áreas vecinas en términos de sus vecinos de mayor grado. Es decir, no se puede concluir simplemente que el valor de ρ indica la fuerza de correlación espacial entre vecinos; ρ es un parámetro más general que está influenciado por toda la estructura de vecindad del grafo. En el caso de la imagen, el valor obtenido fue de $\rho = 0.89$, pero sin un análisis más profundo no se puede interpretar detalladamente.

Otro problema encontrado con el uso del modelo CAR es el bajo valor de correlación marginal *a priori* que se genera, incluso cuando el parámetro ρ tiende a su límite superior 1. En el caso en que el valor ρ se aproxime lo suficiente a 1 y se utilice el modelo IAR, no se encuentra una notable diferencia entre los dos modelos. Besag & Kooperberg (1995) propusieron el modelo IAR, en parte basándose en lo siguiente: señalaron que para un modelo CAR, el valor de ρ tenía que ser 0,999972 para tener una correlación marginal cercana a 0.75. Algo similar muestra Ver Hoef et al. (2017), en el cual el valor del modelo CAR aplicado a sus datos fue $\rho = 0.7$ que produce una correlación marginal máxima de 0.07, algo que claramente no es esperado, ya que se espera que los vecinos de primer orden tenga una fuerte relación. Ver Hoef et al. (2017) también muestra que existe confusión al elegir entre el modelo IAR y el modelo CAR aunque ambos se describen a menudo juntos en la literatura, hay poca orientación en elegir entre ellos.

En la aplicación, tampoco se encontró diferencias notables, quizá la única ventaja del modelo IAR es que tiene un parámetro menos para estimar y podría mejorar el tiempo computacional, aunque este no fue el caso, ya que el tiempo fue casi el mismo (aproximadamente 8 minutos para ambos modelos). En conclusión, ambos

modelos pueden ser utilizados para recuperar imágenes ruidosas con buenos resultados computacionales y visuales, sin embargo, no se evidencia ninguna ventaja utilizando el modelo IAR contra el modelo CAR.

Recibido:

Aceptado:

Referencias

- Arslan, O. & Akyürek, (2018), ‘Spatial modelling of air pollution from pm10 and so2 concentrations during winter season in marmara region (2013-2014)’, *International Journal of Environment and Geoinformatics* pp. 1 – 16.
- Assunção, R. & Krainski, E. (2009), ‘Neighborhood dependence in bayesian spatial models’, *Biometrical Journal* **51**(5), 851–869.
*<https://onlinelibrary.wiley.com/doi/abs/10.1002/bimj.200900056>
- Banerjee, S., P. Carlin, B. & E. Gelfand, A. (2004), ‘Hierarchical modeling and analysis of spatial data’, *Chapman and Hall/CRC Monographs on Statistical and Applied Probability*; **101**.
- Besag, J. (1974), ‘Spatial interaction and the statistical analysis of lattice systems’, *Journal of the Royal Statistical Society. Series B (Methodological)* **36**(2), 192–236.
*<http://www.jstor.org/stable/2984812>
- Besag, J. (1986), ‘On the statistical analysis of dirty pictures’, *Journal of the Royal Statistical Society. Series B (Methodological)* **48**, 259–302.
- Besag, J. & Higdon, D. (1999), ‘Bayesian analysis of agricultural field experiments’, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61**(4), 691–746.
*<https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00201>
- Besag, J. & Kooperberg, C. (1995), ‘On conditional and intrinsic autoregressions’, *Biometrika* **82**(4), 733–746.
*<https://doi.org/10.1093/biomet/82.4.733>
- Chen, C.-C. & Huang, C.-L. (1993), ‘Markov random fields for texture classification’, *Pattern Recognition Letters* **14**(11), 907 – 914.
*<http://www.sciencedirect.com/science/article/pii/0167865593901557>
- Cross, G. R. & Jain, A. K. (1983), ‘Markov random field texture models’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-5**, 25–39.
- Elliott, P. & Wartenberg, D. E. (2004), ‘Spatial epidemiology: Current approaches and future challenges’.

- Givens, G. H. & Hoeting, J. A. (2012), *Computational statistics*, 2 edn, John Wiley & Sons, Hoboken, NJ, USA.
- Halim, S. (2008), 'Modified ising model for generating binary images', *Jurnal Informatika* **8**.
- Horiguchi, T., Honda, Y. & Miya, M. (1997), 'Restoration of digital images of the alphabet by using ising models', *Physics Letters A* **227**(5), 319 – 324.
*<http://www.sciencedirect.com/science/article/pii/S0375960197000807>
- LeSage, J. P. & Thomas-Agnan, C. (2015), 'Interpreting spatial econometric origin-destination flow models', *Journal of Regional Science* **55**(2), 188–208.
*<https://onlinelibrary.wiley.com/doi/abs/10.1111/jors.12114>
- Liu, S. & Cooper, D. B. (2010), 'Ray markov random fields for image-based 3d modeling: Model and efficient inference', pp. 1530–1537.
- Mao, J. & Jain, A. K. (1992), 'Texture classification and segmentation using multiresolution simultaneous autoregressive models', *Pattern Recognition* **25**(2), 173–188.
- Morris, R. D., Descombes, X. & Zerubia, J. (1996), 'The ising/potts model is not well suited to segmentation tasks', pp. 263–266.
- Qian, W. & Titterton, D. M. (1991), 'Multidimensional markov chain models for image textures', *Journal of the Royal Statistical Society: Series B (Methodological)* **53**(3), 661–674.
*<https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1991.tb01855.x>
- Rue, H. & Held, L. (2005), 'Gaussian Markov random fields: Theory and applications', **104**.
*<http://dx.doi.org/10.1201/9780203492024>
- Van Leemput, K., Maes, F., Vandermeulen, D. & Suetens, P. (1999), 'Automated model-based tissue classification of mr images of the brain', *IEEE Transactions on Medical Imaging* **18**(10), 897–908.
- Ver Hoef, J., Peterson, E., Hooten, M., Hanks, E. & Fortin, M.-J. (2017), 'Spatial autoregressive models for statistical inference from ecological data', *Ecological Monographs* .
- Wall, M. M. (2004), 'A close look at the spatial structure implied by the CAR and SAR models', *Journal of Statistical Planning and Inference* **121**(2), 311–324.
*<http://www.sciencedirect.com/science/article/pii/S0378375803001113>
- Xing, W., Deng, N., Xin, B., Chen, Y. & Zhang, Z. (2019), 'Investigation of a novel automatic micro image-based method for the recognition of animal fibers based on wavelet and markov random field', *Micron* **119**, 88 – 97.
*<http://www.sciencedirect.com/science/article/pii/S0968432818304797>