

Sobre la solución numérica de problemas inversos

OSCAR ALBERTO RODRÍGUEZ MELENDEZ
MATEMÁTICO



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE MATEMÁTICAS
BOGOTÁ, D.C.
OCTUBRE DE 2018

Sobre la solución numérica de problemas inversos

OSCAR ALBERTO RODRÍGUEZ MELENDEZ
MATEMÁTICO

DISERTACIÓN PRESENTADA PARA OPTAR AL TÍTULO DE
MAGÍSTER EN CIENCIAS - MATEMÁTICAS APLICADAS

DIRECTOR
JUAN GALVIS , PH.D.
DOCTOR EN MATEMÁTICAS

CODIRECTOR
MARCOS CAPISTRÁN, PH.D.
DOCTOR EN CIENCIAS MATEMÁTICAS



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE MATEMÁTICAS
BOGOTÁ, D.C.
OCTUBRE DE 2018

Título en español

Sobre la solución numérica de problemas inversos.

Resumen: En este trabajo se revisan algunos métodos numéricos y estadísticos para la solución de problemas inversos. El enfoque bayesiano presentado es cada vez más utilizado en las aplicaciones y proporciona una estimación de la función de densidad de probabilidad de los parámetros de interés, contrario a los métodos deterministas que calculan solo un valor estimado para cada uno de los parámetros.

Para ilustrar el potencial de los métodos bayesianos para la solución de problemas inversos, se presenta la solución de recuperación de imágenes y recuperación de parámetros de una ecuación diferencial ordinaria, estocástica y parcial, éste último está asociado al problema de tomografía de impedancia eléctrica. En cada uno de éstos ejemplos se consideran métodos numéricos adecuados para la solución del problema directo.

Abstract: In this work, we review some important statistical and numerical methods for the solution of inverse problems. The Bayesian approach presented here is useful in applications and it computes an estimated probability density function for the parameters of interest. This is the main difference with deterministic methods that provide only one value for the parameters with no confidence interval.

In order to illustrate the applicability of the presented Bayesian approach for the solution of inverse problems, we present several models such as image recovery, parameters recovery of deterministic and stochastic ordinary differential equations and partial differential equations. The later associated to the problem of electrical impedance tomography. In each example we consider the adequate numerical method for the solution of the direct problem.

Palabras clave: Métodos bayesianos, MCMC, Elementos finitos, recuperación de imágenes, Problemas inversos, Impedancia de tomografía eléctrica.

Keywords: Bayesian methods, MCMC, Finite elements, Image recovery, Inverse problems, Electrical impedance tomography.

Nota de aceptación

Trabajo de tesis

Aprobado

“Mención Meritoria o Laureada”

Jurado

Jurado

Director
Juan Carlos Galvis Arrieta

Codirector
Marcos Aurelio Capistrán Ocampo

Bogotá, D.C., Mayo de 2018

Dedicado a

A Dios, a mis padres, a Paola y a los profesores que me guiaron durante el proceso de aprendizaje.

Agradecimientos

A mi familia, por todo su apoyo incondicional.

A mi tutor Juan Carlos Galvis, por asesorar mi desempeño, por la dedicación y por el apoyo brindado en el transcurso de mi vida académica.

A mi tutor Marcos Aurelio Capistrán, por la dedicación y paciencia que me tuvo para el desarrollo de éste trabajo.

Índice general

Índice general	I
Índice de tablas	III
Índice de figuras	IV
Introducción	VII
1. Preliminares	1
1.1. Introducción a las ecuaciones diferenciales parciales	1
1.2. Problema directo e inverso	3
1.3. Enfoque determinista, regularización de Tikhonov	3
1.3.1. Principio de discrepancia de Morozov	4
2. Enfoque bayesiano	7
2.1. Algoritmos tipo Markov Chain Monte Carlo (MCMC)	7
2.1.1. Método de Monte Carlo	7
2.1.2. Muestreo por importancia	8
2.1.3. Algoritmo de Metropolis-Hastings	9
2.1.4. Algoritmo t-walk	10
2.1.5. Familias de muestras invariantes afines	11
2.2. Estimadores	11
2.3. Distribuciones normales	12
2.4. Densidad gaussiana	15
2.5. Introducción a la estimación de parámetros de una ecuación diferencial	19
2.5.1. Generación de datos simulados	20
2.5.2. Problema inverso	21

2.6. Campo aleatorio de Markov gaussiana	24
3. Introducción a los elementos finitos	27
3.1. El método de Galerkin	28
3.2. Problema de la temperatura de una barra	29
3.2.1. Formulación de Galerkin	30
3.2.2. Formulación matricial	31
3.3. Introducción a los elementos finitos (1D)	32
3.4. Introducción a los elementos finitos (2D)	35
4. Problema de impedancia tomográfica	37
4.1. Problema directo de impedancia tomográfica en formulación fuerte	40
4.1.1. Formulación matricial del problema directo	42
4.2. Problema inverso de impedancia tomográfica con planteamiento bayesiano .	45
4.2.1. Construcción de la distribución apriori de γ	45
4.2.2. Distribución posterior de γ	46
5. Resultados y aplicaciones de la metodología bayesiana	47
5.1. Aplicación a la recuperación de imágenes	47
5.2. Aplicación a la estimación de parámetros de ecuaciones diferenciales es- tocásticas (EDE's)	53
5.2.1. Estimación de β para el modelo CEV	54
5.3. Problema inverso de impedancia tomográfica	58
A. Algoritmos numéricos	67
A.1. Método Runge - Kutta de orden 4	67
A.2. Método de Milstein	67
Conclusiones, recomendaciones y trabajo futuro	69
Bibliografía	71

Índice de tablas

1.	Diferentes métodos considerados en el trabajo. El modelo estudiado con más énfasis es el de EIT.	VIII
2.	Distribuciones a priori utilizadas en los diferentes casos empleados en la tesis, en la distribución a priori para la recuperación de imágenes (ver [16]) se considera x el vector de píxeles de la imagen y α la varianza de ésta. En el modelo de población los parámetros r y k son la tasa de cambio y capacidad de carga respectivamente. En el modelo de valoración de activos β es el parámetro de palanca y en el problema de EIT γ es la conductividad, L la matriz de covarianza (conocida) y δ el hiperparámetro del modelo CAMG (ver [2]). En todas las anteriores C es la constante de proporcionalidad.	X
2.1.	Datos de la solución de la ecuación logística en el intervalo $[0,1]$ con tasa de cambio $r = 6$ y capacidad de carga $k = 15$ usando la solución analítica.	20
5.1.	Estimaciones de β para los estimadores MAP (máximo posterior), CM (media condicional) y moda respectivamente.	55
5.2.	Estimaciones del dólar para el 11 de Septiembre de 2017 con el modelo CEV y diferentes estimaciones de β	57
5.3.	Resultados obtenidos para la conductividad γ con 20000 iteraciones en una malla de 50×50	61
5.4.	Resultados obtenidos para la conductividad γ con 50000 iteraciones en una malla de 50×50 sobre el dominio 5.25 con los parámetros presentados en el mismo orden.	64

Índice de figuras

1.1. Imagen de 50 pixeles con ruido generado con un operador A de tipo convolución.	5
1.2. Imagen de 50 pixeles recuperada con regularización de Tikhonov para $\delta = 0.01$	5
1.3. Imagen de 50 pixeles con borrado y ruido aditivo gaussiano.	6
1.4. Imagen de 50 pixeles recuperada con Tikhonov de una imagen borrada y con ruido aditivo gaussiano.	6
2.1. Solución de la ecuación logística en el intervalo $[0,1]$ con tasa de cambio $r = 6$ y capacidad de carga $k = 15$ usando la solución analítica para obtener resultados precisos.	21
2.2. Datos de la ecuación logística en el intervalo $[0,1]$ con un ruido poisson de parámetro $x(t_i, \theta)$ para $i = 1, \dots, 101$ y $\theta = (r, k)$ representa los parámetros de tasa de cambio y capacidad de carga.	22
2.3. Comportamiento de las muestras para 10000 iteraciones	23
2.4. Histograma de las muestras para 10000 iteraciones	24
2.5. Gráfico del logaritmo de las probabilidades de aceptación de las muestras 10000 iteraciones	24
2.6. Diez trayectorias de una normal multivariada en 1D, tomando solo las medias y varianzas asociadas de cada trayectoria se muestra el camino que genera a través de las iteraciones centrándose en 0 (punto más probable).	25
2.7. Una muestra en 2D de una grilla normal multivariada de 100x100, es decir cada pixel posee una distribución normal.	26
3.1. Funciones que pasan por dos puntos.	27
3.2. Triangulación del dominio $[0,1] \times [0,1]$ en triángulos.	36
3.3. Triangulación del dominio $[0,1] \times [0,1]$ en cuadrados.	36
4.1. Esquema del problema de impedancia tomográfica. (Tomado y modificado de “ https://www.aip2015.fips.fi/intro-part5.png ” from Katya Krupchyk)	40

5.1. Imagen de 5×5 pixeles.	48
5.2. Imagen de la Figura 5.1 con un borrado de la forma 5.1 para $\sigma = 0.0821$. . .	48
5.3. Comparación entre la imagen real, imagen borrada, resultado obtenido a partir de la regularización de Tikhonov y el máximo posterior de la función de densidad (punto más probable) con función de densidad dada por 5.5 para $\sigma = 0.0821$ con cien mil (100000) iteraciones del modelo MCMC.	49
5.4. Logaritmo de la probabilidad de aceptación del MCMC para la posterior (5.5) con cien mil (100000) iteraciones.	49
5.5. Imágenes de la Figura 5.1 generadas con un ruido gaussiano presentado en (5.1) con media cero y varianza 0.0821.	51
5.6. Comparación de la recuperación de la imagen de la Figura 5.1 por medio del promedio de la regularización de Tikhonov a cada imagen con el máximo posterior de las muestras obtenidas para la función de densidad posterior (5.7).	51
5.7. Logaritmo de la probabilidad de aceptación de cada muestra obtenida por el modelo MCMC para la función de densidad posterior (5.7) con 2 millones (2000000) de iteraciones.	52
5.8. Comparación de la recuperación de la imagen de la Figura 5.1 utilizando la moda y la media condicional de las muestras obtenidas por el modelo MCMC aplicado a la función de densidad posterior (5.7).	53
5.9. Histograma de la función posterior para β dada por (5.13) para 5000 iteraciones.	55
5.10. Logaritmo natural de la probabilidad de aceptación por el modelo MCMC de la función posterior para β dada por (5.13) con 5000 iteraciones.	56
5.11. Ajuste y predicción de del dólar en Colombia usando el modelo CEV para β dado por el máximo posterior.	56
5.12. Ajuste y predicción de del dólar en Colombia usando el modelo CEV para β dado por la media condicional.	57
5.13. Ajuste y predicción de del dólar en Colombia usando el modelo CEV para β dado por la moda.	57
5.14. Conductividad con inclusión en el centro.	58
5.15. Dominio del problema EIT con cien particiones en el eje X y cien particiones en el eje Y para un total de cien mil pixeles, tomando una discretización cuadrada como en 3.3 con 10 electrodos en los bordes ocupando 20 pixeles cada electrodo.	59
5.16. Voltajes obtenidos de la solución del problema directo mediante la aplicación del método de los elementos finitos.	60
5.17. Voltajes con ruido.	61
5.18. Corrientes con ruido.	61
5.19. Datos iniciales con ruido gaussiano de media cero (0) y varianza 0.00000022.	61

5.20. Versus entre los estimadores de la media condicional, moda, máximo posterior y el valor real de la inclusión de la conductividad γ	62
5.21. Logatirno de la probabilidad de aceptación de cada muestra obtenida por el modelo MCMC para la 20000 iteraciones.	62
5.22. Función de densidad marginal para el exterior de la inclusión.	63
5.23. Función de densidad marginal para la inclusión en el centro del dominio. . .	63
5.24. Funciones de densidad marginales obtenidas para los parámetros estimados por el método bayesiano.	63
5.25. División del subdominio en cinco “subdominios verticales” donde se recuperarán 5 parámetros.	63
5.26. Máximo posterior de los cinco parámetros estimados.	64
5.27. Logaritmo natural de las probabilidades de aceptación de las muestras para los parámetros de la Figura 5.26.	64
5.28. Media condicional de la conductividad obtenida al estimar 9 parámetros sobre el dominio.	65
5.29. Logaritmo natural de las probabilidades de aceptación de las muestras para 9 parámetros.	65
5.30. Máximo posterior de la conductividad obtenida al estimar 25 parámetros sobre el dominio.	66
5.31. Logaritmo natural de las probabilidades de aceptación de las muestras para 25 parámetros.	66

Introducción

En este trabajo se estudian los métodos bayesianos para la solución de problemas inversos, vea [16]. En particular para ilustrar el potencial de éstos métodos se consideran los siguientes problemas:

- i) Determinar los coeficientes de un modelo de crecimiento poblacional, ver (2.18).
- ii) Determinar la conductividad eléctrica de un medio, asociado al problema de tomografía de impedancia eléctrica (electrical impedance tomography) o EIT por su nombre en inglés, ver (4.11).
- iii) Recuperar una imagen con borrado conocido usando estadística bayesiana, usando un modelo de problema de convolución (5.1).
- iv) Recuperar un parámetro de una ecuación diferencial estocástica (5.9) para ajustar los datos a un modelo con intervalo de confianza.

En especial, se presenta en detalle el problema de EIT que es el problema más complejo de los mencionados. La EIT es un procedimiento que permite estimar la conductividad eléctrica de un cuerpo (o más generalmente la distribución de la impedancia) si se conoce algunos datos de corriente y voltaje en la frontera del mismo, ver [16]. Existen diversas aplicaciones para la solución numérica del problema de la impedancia. Cabe mencionar que en el campo de la medicina, la tomografía es de gran interés en la actualidad; por ejemplo, es posible usar imágenes y estimaciones generadas como medio para obtener una imagen del interior del cerebro (esto es posible debido a la diferencia en conductividad eléctrica de los principales tejidos de la cabeza). Como aplicación directa se tiene la posibilidad de detectar anomalías y enfermedades tales como tumores cancerígenos; de la misma manera, se puede implementar para construir imágenes de otras partes del cuerpo u órganos.

El problema de la EIT se puede clasificar como un problema inverso. En general, un problema inverso es un problema de determinación de parámetros en un problema directo, dado un conjunto de datos. El problema directo puede ser interpretado como una aplicación que, dados algunos parámetros, calcula la solución de una ecuación (parametrizada por los datos). Por ejemplo, el problema directo asociado al problema inverso de EIT es la solución de la ecuación de difusión de la corriente eléctrica en el tejido del órgano a estudiar. Es claro que, para poder analizar y resolver un problema inverso se debe tener en cuenta el análisis y algoritmos relacionados al problema directo, ver [17]. Para comprender la relación entre estos problemas vea la Tabla 1.

Tópico / Modelo	Imágenes	Modelo de Población	Valoración de activos	EIT
Problema directo (PD)	Borrado de la imagen	Evolución temporal de una población	Evolución de un activo	Difusión del voltaje en un medio
Modelo matemático	Operador de convolución	Problema de valor inicial determinista	Problema de valor inicial estocástico	Problema de frontera para una ecuación diferencial parcial
Datos	Imágenes borradas	Datos de población en el tiempo	Valor del activo en el tiempo	Corriente y voltaje en los electrodos
Problema inverso	Recuperación de la imagen original	Recuperación de la tasa de crecimiento y la capacidad de carga	Recuperar el parámetro de palanca	Recuperar la impedancia y conductividad
Parámetros a estimar	Valor de cada pixel de la imagen	Un valor real para la tasa de crecimiento y un valor real para la capacidad de carga	Un valor real para el parámetro de palanca	Valor de la impedancia en los electrodos y valor de la conductividad en cada pixel de una malla en el dominio
Método numérico del PD	Representación matricial de la convolución	Método Runge-Kutta	Método de Milstein	Método de elementos finitos bilineal por partes

TABLA 1. Diferentes métodos considerados en el trabajo. El modelo estudiado con más énfasis es el de EIT.

Existen varias metodologías para analizar y resolver problemas inversos, a grosso modo se pueden clasificar en metodologías deterministas y estadísticas. Entre los métodos deterministas se encuentran métodos de proyección y métodos de regularización (el más popular siendo el de regularización de Tikhonov) [16, pág 16]. Por otro lado, está el enfoque estadístico, el cual ha venido ganando popularidad en años recientes [17]. Por ejemplo, se puede aplicar un método bayesiano en el que se cuenta con información con respecto a un conjunto de datos recolectados, se asume que dicha información contiene errores incorporados. Estos errores pueden provenir de la medición, condiciones climáticas, entre otros, según sea la naturaleza de los datos, al no conocerse la medida exacta del error, se puede decir que dichos errores son aleatorios. De dichos errores aleatorios se tiene alguna información en términos de una función de densidad (la función de densidad depende de la naturaleza de los datos y es una pieza clave en la construcción del modelo que permite obtener una solución estimada), hallando así la distribución posterior de tales parámetros que puede ser muestreada mediante algoritmos tipo Markov Chain Monte Carlo (MCMC);

ver [20]. La idea es que en lugar de estimar un solo valor para los parámetros del problema inverso, se estima la funciones de densidad de los posibles valores de estos parámetros, obteniendo así una densidad que representa el comportamiento de dicho parámetro. Se menciona aquí, que en la práctica se utilizan los métodos de tipo de Monte Carlo con cadenas de Markov (MCMC) para aproximar su densidad. En caso de requerirse un único valor para este parámetro se puede estimar mediante el máximo posterior de la distribución a posteriori, la media de dicha distribución o cualquier otro estimador puntual adecuado para el parámetro de interés.

La aplicabilidad y conveniencia de la solución numérica de problemas inversos depende en gran medida de la calidad de los algoritmos para resolver el problema directo asociado. Por ejemplo, para el caso de EIT, se puede usar el método de elementos finitos para lidiar con el problema directo. Es conocido que la solución por elementos finitos es de alta complejidad computacional (más aún para mallas muy finas y conductividades complicadas).

Los algoritmos de muestreo usados en el enfoque bayesiano requieren la solución numérica del problema directo para cada nueva muestra o propuesta a evaluar de los valores de los parámetros de interés. En éstos casos es fundamental para la aplicabilidad de los algoritmos, que la solución del problema directo sea eficiente y que el número de propuestas a evaluar sea moderado.

En ésta tesis se aborda este problema desde el punto de vista de mantener moderado el número de propuestas necesarias, para ésto se debe seleccionar adecuadamente el valor inicial y la distribución apriori de los parámetros de interés. Ésta es una de las principales contribuciones de éste trabajo, y esto se debe a que se observó que pueden ser usadas las siguientes distribuciones prior para los casos planteados:

- i) Recuperación de la imagen: Para este caso la apriori escogida fue una normal con varianza desconocida, pues se considero que un borrado usual es el ruido gaussiano presente en la naturaleza, debido al comportamiento del movimiento de ciertas partículas como una caminata aleatoria, la cual bajo ciertas condiciones puede llegar a generar una distribución normal.
- ii) Coeficiente del modelo de población: Para este caso las apriori escogidas fueron gammas pues los parámetros son positivos y no están acotados en los reales.
- iii) Parámetro de palanca del modelo CEV: Para este caso la apriori escogida fue una gamma pues el parámetros es positivo y no está acotado en los reales.
- iv) Parámetro de conductividad de la EIT: Para este caso la apriori escogida fue un campo aleatorio de Markov gaussiana, pues ésta es una prior adaptativa a la malla escogida para la solución del problema directo por elementos finitos.

En la Tabla 2 se pueden observar las distribuciones apriori que fueron utilizadas en los problemas.

Modelo	Imágenes	Modelo de Población	Valoración de activos	EIT
Apriori	$C \exp\left(-\frac{1}{2\alpha}x^T x - \alpha\right)$	$C \exp(-r - k)$	$C \exp(-\beta)$	$C \exp\left(-\frac{\delta}{2}(\gamma^T L \gamma + 1)\right)$

TABLA 2. Distribuciones apriori utilizadas en los diferentes casos empleados en la tesis, en la distribución apriori para la recuperación de imágenes (ver [16]) se considera x el vector de píxeles de la imagen y α la varianza de ésta. En el modelo de población los parámetros r y k son la tasa de cambio y capacidad de carga respectivamente. En el modelo de valoración de activos β es el parámetro de palanca y en el problema de EIT γ es la conductividad, L la matriz de covarianza (conocida) y δ el hiperparámetro del modelo CAMG (ver [2]). En todas las anteriores C es la constante de proporcionalidad.

Cabe aclarar que en este trabajo no se hace enfoque en la eficiencia de la solución del problema directo, éste problema puede ser considerado en problemas futuros.

El resto del trabajo ésta organizado como sigue: En el capítulo 1 se presentan los preliminares necesarios para entender lo que es una ecuación diferencial parcial y una introducción al porque es conveniente usar estadística bayesiana para los problemas inversos. En el capítulo 2 se muestra el enfoque bayesiano del trabajo, presentando algunas demostraciones sobre las distribuciones gaussianas que son principalmente las utilizadas en las aplicaciones aquí presentadas, al igual que se da una introducción a una apriori gaussiana de construcción especial para problemas inversos. En el capítulo 3 se presenta una introducción al método de los elementos finitos que es el método con el que se resuelven la ecuación diferencial parcial asociada a la EIT. En el capítulo 4 se presenta el problema de impedancia tomográfica, mostrando la deducción de la ecuación matemática que viene dada de un problema físico y gobierna el problema directo de la EIT. En el capítulo 5 se presentan las aplicaciones de la metodología bayesiana a los problemas inversos en el problema de convolución, de la ecuación diferencial estocástica y la EIT. Por último se presentan algunas conclusiones del trabajo y un posible trabajo futuro.

CAPÍTULO 1

Preliminares

En este capítulo se realiza una corta revisión de conceptos básicos relacionados al tema de estudio. En particular se revisa las definiciones básicas de ecuaciones diferenciales parciales y de los problemas inversos. En la sección 1.1 se presentan algunas definiciones básicas de lo que es una ecuación diferencial parcial. En la sección 1.2 se presenta la diferencia entre lo que es un problema inverso y el problema directo, con el ejemplo del problema de convolución que es solucionado utilizando la regularización de Tikhonov.

1.1. Introducción a las ecuaciones diferenciales parciales

Para el material aquí presentado vea [15]. Una ecuación en derivadas parciales o ecuación diferencial parcial (EDP) es una ecuación que envuelve dos o más variables independientes $x_1, x_2, \dots, x_n, \dots$ y derivadas parciales de una función desconocida $u = u(x_1, x_2, \dots, x_n, \dots)$, es decir una EDP en n variables independientes es de la forma:

$$F\left(x_1, \dots, x_n, u, \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n}, \frac{\partial^2 u}{\partial x_1^2}, \dots, \frac{\partial^2 u}{\partial x_1 \partial x_n}, \dots, \frac{\partial^k u}{\partial x_n^k}\right) = 0. \quad (1.1)$$

En particular una EDP lineal de segundo orden,

$$\sum_{i,j=1}^n a_{ij}(x) D_i D_j u + \sum_{j=1}^n b_j(x) D_j u + c(x)u + d(x) = 0,$$

donde alguno de los coeficientes a_{ij} no es idénticamente nulo (para segundo orden). Aquí $D_j u = \frac{\partial u}{\partial x_j}$ y $D_i D_j u = \frac{\partial^2 u}{\partial x_i \partial x_j}$. Si en la expresión (1.1) se tiene que $d(x) = 0$ se dice que la EDP es homogénea, en caso contrario se dice que es no homogénea.

Ejemplo 1.1.1. *La ecuación de onda*

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2},$$

es una ecuación lineal, homogénea de segundo orden. Donde la variable $t > 0$ representa el tiempo, $x \in \mathbb{R}$ es la variable espacial y $c > 0$ es una constante (velocidad de propagación de la onda).

Considere una EDP de primer o segundo orden con n variables independientes con notación vectorial $x = (x_1, \dots, x_n)$. Considere la ecuación (1.1), ésta puede ser reescrita en la forma,

$$Lu = f,$$

donde $f(x) = -d(x)$, entonces,

$$(Lu)(x) = \sum_{i,j=1}^n a_{ij}(x) D_i D_j u(x) + \sum_{j=1}^n b_j(x) D_j u(x) + c(x)u(x). \quad (1.2)$$

A cada función u (suficientemente diferenciable) le corresponde una única función Lu , de ésta forma se define un operador o transformación L . De forma más precisa, sea Ω un subconjunto abierto de \mathbb{R}^n y suponga que las funciones a_{ij} , b_{ij} y c con $1 \leq i, j \leq n$, son continuas en Ω y toman valores reales, entonces se puede definir,

$$\begin{aligned} L : C^k(\Omega) &\rightarrow C(\Omega) \\ u &\rightarrow Lu, \end{aligned} \quad (1.3)$$

donde Lu es dado por la fórmula (1.2) y $C^k(\Omega)$ (con $k = 1, 2$ el orden de la ecuación) es un conjunto de las funciones $u : \Omega \rightarrow \mathbb{R}$ que son k veces continuamente diferenciables.

En el caso de las EDP's el espacio de las variables independientes es multidimensional, por ejemplo definidas en un abierto $\Omega \subseteq \mathbb{R}^n$, es natural sustituir los extremos del intervalo por el borde $\partial\Omega$ de la región Ω . Cuando se impone condiciones sobre el valor de la solución y sus derivadas en el borde de la región (condiciones de contorno) se tiene un problema de contorno.

Ejemplo 1.1.2. Se presenta la siguiente EDP,

$$\alpha u(x) + \beta \frac{\partial u}{\partial n}(x) = f(x), \quad x \in \partial\Omega.$$

Donde α y β son constantes dadas, f es una función definida en $\partial\Omega$ y $\frac{\partial u}{\partial n}$ es la derivada de u en la dirección normal a $\partial\Omega$. En el caso en que $\alpha = 0$, la condición es conocida como **condición de Neumann**, en el caso $\beta = 0$ es conocida como **condición de Dirichlet**.

Las condiciones iniciales se presentan imponiendo un valor de la solución y sus derivadas normales a lo largo de una curva o superficie inicial, el problema correspondiente es un **problema de Cauchy** o de **valor inicial**. Suele ocurrir que los valores de la solución y de sus derivadas en relación al tiempo hasta el orden $k - 1$ (suponiendo que la EDP es de orden k en t) son descritos en el instante $t = 0$ como función de x, y, z (condición inicial), al mismo tiempo en que son impuestas condiciones de contorno, para todo $t \geq 0$, en relación a las variables espaciales, tales problemas son llamados **problemas mixtos**, un ejemplo de esto es el problema de onda en un intervalo finito o como se verá posteriormente el problema de impedancia eléctrica.

Ejemplo 1.1.3. *El problema de la ecuación de onda en un intervalo finito,*

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \text{ en } (0, l) \times (0, +\infty), \\ u(0, t) = 0 = u(l, t), \quad t \geq 0, \\ u(x, 0) = f(x), \quad x \in [0, l], \\ u_t(x, 0) = g(x), \quad x \in [0, l], \end{cases} \begin{array}{l} \text{condición inicial,} \\ \text{condición de contorno,} \\ \text{condición de contorno,} \end{array}$$

puede ser considerado como un problema mixto con condiciones iniciales $u(x, 0) = f(x)$, $u_t(x, 0) = g(x)$, $x \in [0, l]$ y condiciones de contorno $u(0, t) = 0 = u(l, t)$, $t \geq 0$.

Se llama **formulación débil** o **formulación variacional** a una forma alternativa en que las EDP's se escriben en forma integral, dando lugar a ecuaciones tratables mediante distintos métodos sobre espacios vectoriales o funcionales. La ventaja de la formulación variacional es que, por ejemplo, la formulación variacional de una EDP de segundo orden usa solo primeras derivadas.

1.2. Problema directo e inverso

El objetivo de recopilar datos de cualquier sistema es poder obtener información significativa de dicho sistema, pero en muchas ocasiones hay cantidades que son imposibles de medir directamente. Se llama problema directo, al problema en el que se predicen los datos de observación mediante un modelo concreto con parámetros ya dados. Se llama problema inverso al proceso de, dado un conjunto de datos se asume que estos datos provienen de un problema directo (existe una ecuación que describe el comportamiento), construir un modelo para obtener información acerca de los parámetros que implican estos datos. Dicho de otra forma, un problema inverso se puede considerar como un problema donde es medido un efecto y se desea conocer la causa.

En las ciencias aplicadas se utilizan comúnmente estas metodologías, pero no siempre son cobijadas bajo el nombre de problema inverso, así que las características distintivas de un problema inverso vienen dadas de problemas físicos con análisis matemático y alguna metodología para su solución (comúnmente computación numérica avanzada). Por ejemplo el problema de regresión lineal usa la metodología de problema inverso, en el que dado un conjunto de datos se busca encontrar los parámetros de una ecuación lineal que describa el comportamiento de dichos datos, sin embargo éste problema bien conocido no es considerado un problema inverso.

1.3. Enfoque determinista, regularización de Tikhonov

Suponga que se tiene una ecuación como la siguiente, $Y = AX$, donde A es un operador compacto. No siempre es posible obtener la solución $X = A^{-1}Y$, especialmente cuando la matriz A es muy grande. Calcular la inversa no es un problema sencillo computacionalmente (por su costo computacional). Para poder solucionar problemas como estos es necesario implementar otros modelos, que aunque no obtengan una solución exacta proporcionan una buena aproximación con un error admisible. Entre algunos de los métodos de solución a problemas de este tipo se encuentra los procesos de regularización, especialmente uno muy conocido fue introducido por Tikhonov en 1977 el cual es revisado a continuación usando como base [16].

La regularización se refiere al proceso de introducir información adicional a los datos para impedir el sobreajuste de los mismos, en el caso específico de Tikhonov el proceso de regularización hace referencia a dado el problema $y = Ax$ minimizar el siguiente funcional,

$$F_\delta(x) = \|Ax - y\|^2 + \delta\|x\|^2, \quad (1.4)$$

donde el parámetro $\delta > 0$ es llamado el parámetro de regularización. Cuando A es un operador lineal, la solución analítica al anterior problema viene dada por,

$$x_\delta = (A^*A + \delta I)^{-1} A^*y, \quad (1.5)$$

donde A^* representa la adjunta conjugada de A .

Uno de los aspectos más importantes de la regularización de Tikhonov es la selección del parámetro δ , pues una buena escogencia del δ nos reduce el error en la solución propuesta. Para esto se introducirá una noción del principio de discrepancia.

1.3.1. Principio de discrepancia de Morozov

El principio de discrepancia de Morozov es un método que se usa para determinar un parámetro de regularización, para ello se define,

$$f(\delta) = \|Ax_\delta - y\|.$$

La estimación del error $\epsilon > 0$ de la solución a un problema de aproximación viene dado por,

$$\|Ax - y\| < \epsilon.$$

El principio de discrepancia de Morozov nos dice que δ puede ser escogido para que cumpla la siguiente condición

$$f(\delta) = \|Ax_\delta - y\| = \epsilon,$$

cuando ésta sea posible, es decir, la solución de regularización x_δ no debe tratar de satisfacer los datos con mayor precisión que hasta el nivel de ruido que contienen los datos (evitar problemas de sobreajuste).

Un ejemplo de aplicación en el caso determinista es dada la imagen con borrado de la Figura 1.1. Aplicando la regularización de Tikhonov con una matriz A de suavizamiento de la imagen, se puede recuperar la imagen original para algún δ . Aplicando el principio de Morozov se puede escoger el mejor δ hasta obtener la imagen de la Figura 1.2.

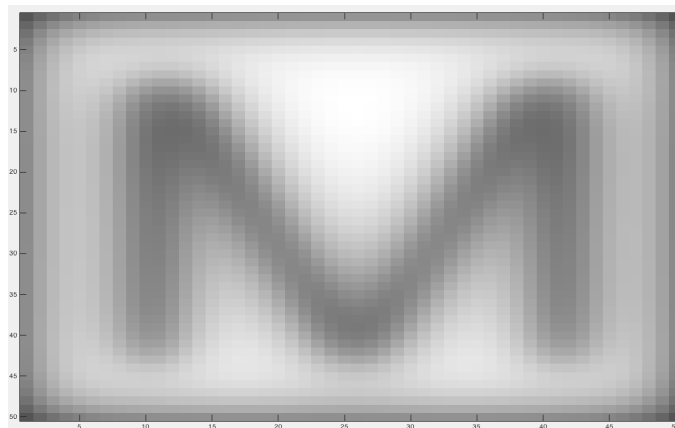


FIGURA 1.1. Imagen de 50 pixeles con ruido generado con un operador A de tipo convolución.



FIGURA 1.2. Imagen de 50 pixeles recuperada con regularización de Tikhonov para $\delta = 0.01$.

Como se puede observar en la Figura 1.2 la recuperación es bastante buena, pero esto no necesariamente es así si la imagen presenta un ruido aditivo o borrado más ruido aditivo como el de la Figura 1.3, donde la recuperación por medio de Tikhonov se presenta en la Figura 1.4.

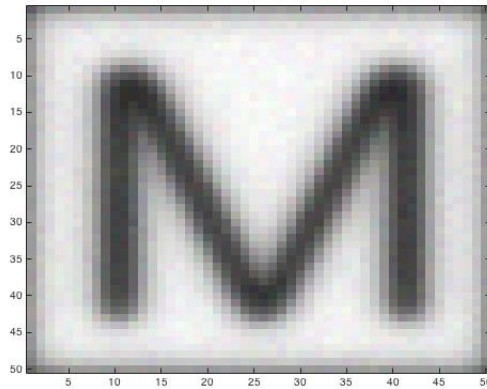


FIGURA 1.3. Imagen de 50 pixeles con borrado y ruido aditivo gaussiano.

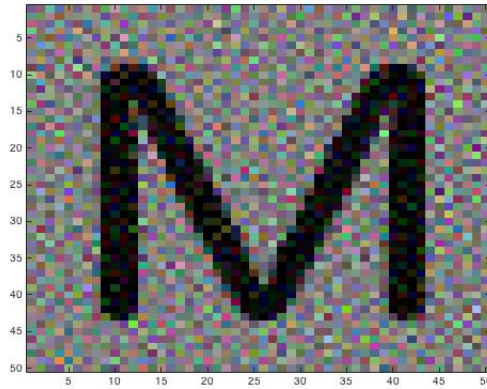


FIGURA 1.4. Imagen de 50 pixeles recuperada con Tikhonov de una imagen borrada y con ruido aditivo gaussiano.

Como se puede observar en la Figura 1.4 al aplicar el método de Tikhonov sin ninguna modificación, se recupera una imagen sin borrado pero no elimina el ruido que trae consigo la imagen. El problema está en el primer término del funcional ya que este término no considera el ruido aditivo, esto motiva un modelamiento más preciso de este tipo de problemas

Para ello existen diferentes metodologías, entre ellas está la recuperación por medio de un enfoque bayesiano.

En el siguiente capítulo se presentará un enfoque bayesiano, algunos cálculos sobre distribuciones normales y la relación del enfoque bayesiano con la regularización de Tikhonov.

Enfoque bayesiano

En este capítulo se presentará la metodología bayesiana (lo necesario para el problema inverso trabajado) que será de gran utilidad en los problemas inversos, y que en este caso será de gran ayuda para el problema de impedancia tomográfica. En la sección 2.1 se presentan algunos algoritmos que serán de utilidad para entender como son el algoritmo de Metropolis Hastings, el t-walk y el GWMCMC, en la sección 2.2 se presentan algunos estimadores estadísticos puntuales, en la sección 2.3 se podrán observar algunas propiedades de distribuciones normales, en la sección 2.4 algunos teoremas sobre densidades gaussianas, en la sección 2.5 se presenta un ejemplo de la metodología para recuperar los parámetros de una ecuación diferencial ordinaria y por último se presenta el campo aleatorio de Markov gaussiano que será de utilidad al momento de construir la distribución apriori para la conductividad en el problema inverso de EIT.

2.1. Algoritmos tipo Markov Chain Monte Carlo (MCMC)

Los algoritmos MCMC por sus siglas en inglés (Markov Chain Monte Carlo) o en español Monte Carlo basados en cadenas de Markov que generan muestras independientes de una distribución aplicando métodos de Monte Carlo y así poder estimar la función de probabilidad de dicha distribución. En estos métodos se simulan valores sucesivos de una densidad propuesta apriori, que no tiene que ser necesariamente parecida a la densidad a posterior, pues entre las muestras generadas no todas son aceptadas, para ésto se usa un método de aceptación. Las muestras aceptadas son utilizadas para estimar la función de densidad requerida. Éstos métodos no son exclusivos de la inferencia bayesiana, son muy utilizados para funciones de densidad como para distribuciones posterior que no son fáciles de muestrear. Cada valor generado depende solo del valor anterior, de ahí que se enmarcan en la noción de cadenas Markov.

2.1.1. Método de Monte Carlo

Suponga que θ tiene una distribución posterior con densidad $g(\theta|y)$ y se está interesado en conocer la función $h(\theta)$ (obtener muestras de $h(\theta)$ para ver el comportamiento de la

variable aleatoria). La media de $h(\theta)$ está dada por,

$$E(h(\theta)|y) = \int h(\theta)g(\theta|y)d\theta. \quad (2.1)$$

Ahora se obtiene m muestras $\theta^1, \dots, \theta^m$ de $g(\theta|y)$, luego la estimación de la media viene dada por,

$$\hat{h} = \sum_{j=1}^m \frac{h(\theta^j)}{m}.$$

Un ejemplo simple de la aplicación del método es para la solución de integrales como se observa a continuación:

Ejemplo 2.1.1. *Integrar la función $f(x) = x^2$ en el intervalo $[0, 1]$ (es decir, calcular $\int_0^1 x^2 dx$). Considere $\theta = (\theta^1, \theta^2, \dots, \theta^{100})$, es decir cien muestras con distribución uniforme en el intervalo $[0, 1]$, luego*

$$\hat{h} = \frac{1}{100} \sum_{j=1}^{100} (\theta^j)^2.$$

Realizando la simulación se obtiene como resultado 0.337656 con el siguiente código en R;
`c<-runif(100,min=0,max=1); fun<-function(x){x^2}; h<-sum(fun(c))/100.`

2.1.2. Muestreo por importancia

Considere (2.1), se calculará una aproximación implementando Monte Carlo con algunas modificaciones donde $f(y|\theta) = \frac{g(\theta,y)}{g(\theta)}$ es la función de verosimilitud de los datos, $g(y)$ la función de distribución de los datos y $g(\theta)$ la distribución a priori para el parámetro θ .

Note primero que vale

$$\begin{aligned} E(h(\theta)|y) &= \int h(\theta)g(\theta|y)d\theta \\ &= \int h(\theta) \frac{g(\theta,y)}{g(\theta)} d\theta \\ &= \int \frac{h(\theta)f(y|\theta)g(\theta)}{g(y)} d\theta \\ &= \int \frac{h(\theta)f(y|\theta)g(\theta)}{\int g(y,\theta)d\theta} d\theta \\ &= \int \frac{h(\theta)f(y|\theta)g(\theta)}{\int f(y|\theta)g(\theta)d\theta} d\theta. \end{aligned}$$

Ahora se multiplica por una función $p(\theta)$ tal que,

- La función $p(\theta)$ sea fácil de muestrear (distribución conocida).
- La densidad de $p(\theta)$ se asemeja a la densidad posterior de interés g en términos de ubicación y propagación.
- Existe una constante c para todo θ tal que $g(\theta|y) \leq cp(\theta)$.

Así, se tiene

$$\begin{aligned} E(h(\theta)|y) &= \int \frac{h(\theta) \frac{g(\theta)f(y|\theta)}{p(\theta)} p(\theta)}{\int \frac{g(\theta)f(y|\theta)}{p(\theta)} p(\theta) d\theta} d\theta \\ &= \int \frac{h(\theta)w(\theta)p(\theta)}{\int w(\theta)p(\theta)d\theta} d\theta. \end{aligned}$$

Con $w(\theta) = \frac{g(\theta)f(y|\theta)}{p(\theta)}$ se calculan m muestras de la función p , se aplica método de aceptación rechazo como se sigue:

- Se simulan θ independientes de p , donde θ es el vector de parámetros. Posteriormente se obtiene U una variable aleatoria uniforme en el intervalo unidad.
- Si $U \leq \frac{g(\theta|y)}{cp(\theta)}$ se acepta θ , en otro caso se rechaza.
- Continúe los anteriores pasos hasta que halle la muestra m deseada.

Si $\theta^1, \dots, \theta^m$ son muestras aleatorias a partir de p , entonces el muestreo por importancia genera

$$\hat{h}_{MI} = \frac{\sum_{j=1}^m h(\theta^j)w(\theta^j)}{\sum_{j=1}^m w(\theta^j)}.$$

Para más resultados y ejemplos ver [20].

2.1.3. Algoritmo de Metropolis-Hastings

El algoritmo de Metropolis-Hastings es un modelo MCMC es uno de los algoritmos más conocidos en la estadística bayesiana, pues es el primer algoritmo para obtener muestras de una función de densidad propuesta no necesariamente simétrica, ésta es una versión generalizada del algoritmo creado por Nicholas Metropolis ver [13]. Sea $\pi_{pos}(x)$ la distribución a posterior de X , para obtener información sobre la distribución se siguen los siguientes pasos;

1. Se escoge un valor $x_k \in \mathbb{R}^n$ para $k = 1$.
2. Se obtiene una muestra $y \in \mathbb{R}^n$ mediante una distribución propuesta $q(x_k|y)$ bajo las condiciones del muestreo por importancia y se calcula el radio de aceptación definido por,

$$\alpha(x_k, y) = \min\left(1, \frac{\pi_{pos}(y)q(y|x_k)}{\pi_{pos}(x_k)q(x_k|y)}\right).$$

3. Se toma $u \in [0, 1]$ con probabilidad de densidad uniforme.
4. Si $\alpha(x_k, y) \geq u$, se escoge $x_{k+1} = y$, en otro caso $x_{k+1} = x_k$.
5. Volver al paso 2 y continuar hasta obtener el número de muestras deseadas.

2.1.4. Algoritmo t-walk

Se presenta una breve explicación del algoritmo *t-walk* tomado de [5], el cual es una modificación del Metrópolis-Hastings. Se busca obtener muestras de una distribución $\pi(x)$, considere $f(x, x') = \pi(x)\pi(x')$ con $(x, x') \in C \subset \mathbb{R}^n \times \mathbb{R}^n$, mientras la propuesta (apriori) será de forma general, $q\{(y, y') = (x, x')\}$, bajo las siguientes restricciones propuestas:

$$(y, y') = \begin{cases} (x, h(x', x)), & \text{si } I = 1, \\ (h(x, x'), x'), & \text{si } I = 0. \end{cases} \quad (2.2)$$

Donde $I \sim \text{Bernoulli}(0.5)$, h es una variable aleatoria que sirve de propuesta y se busca solo x o x' en cada paso. Notando $h(x, x')$ por $g(\cdot|x, x')$, en estas condiciones se tiene que la expresión para el radio de convergencia está dada por:

$$\frac{\pi(y')g(x'|y', x)}{\pi(x')g(y'|x', x)}. \quad (2.3)$$

Recuerde que para una caminata aleatoria escalonada h puede verse como:

$$h_W(x, x')_j = \begin{cases} x_j + (x_j - x'_j)\alpha_j, & I_j = 1 \\ x_j, & I_j = 0, \end{cases} \quad (2.4)$$

para $j = 1, 2, \dots, n$, donde $\alpha_j \in \mathbb{R}$ son independientes e igualmente distribuidas, e I_j tiene una distribución Bernoulli con parámetro 0.5. En la propuesta del *t-walk* se define una propuesta para h como se sigue:

$$h_t(x, x')_j = \begin{cases} x'_j + (x'_j - x_j)\beta, & I_j = 1, \\ x_j, & I_j = 0, \end{cases} \quad (2.5)$$

para $j = 1, 2, \dots, n$, donde $\beta \in \mathbb{R}^+$ con función de densidad conocida, e $I_j \sim \text{Bernoulli}(0.5)$. Se define la función β como,

$$\beta = \begin{cases} u^{\frac{1}{1+a_t}}, & \text{con probabilidad } \frac{a_t-1}{2a_t}, \\ u^{\frac{1}{1-a_t}}, & \text{con probabilidad } \frac{a_t+1}{2a_t}, \end{cases} \quad (2.6)$$

con $u \sim U(0, 1)$ y $a_t > 1$.

Sea $n_I = \sum_{j=1}^n I_j$ los cambios transversales realizados en n iteraciones, el radio de aceptación para el *t-walk* está dado por $\frac{\pi(y')}{\pi(x')} \beta^{n_I-2}$ para el caso 1 o $\frac{\pi(y)}{\pi(x)} \beta^{n_I-2}$ en el caso 2 en (2.5). Lo que conlleva al siguiente algoritmo MCMC:

Algoritmo t-walk-

Sea $\pi_{pos}(x)$ la distribución a posterior de X , para obtener información sobre la distribución se siguen los siguientes pasos (<https://www.cimat.mx/jac/twalk/>);

1. Se escogen dos valores x_k y $x'_k \in \mathbb{R}^n$ para $k = 1$.
2. Se obtiene una muestra y o $y' \in \mathbb{R}^n$ mediante la distribución propuesta $g_t(x_k|y, x')$ o $g_t(x'_k|y', x)$ respectivamente, escogiéndolos de la siguiente forma:
 - 2.1. Se obtiene un número aleatorio de distribución binomial con parámetro 0.5, dado el resultado se procede a calcular h_t .

- 2.2. Se obtiene un número aleatorio de distribución binomial con parámetro 0.5, si se obtiene 1 se obtiene la muestra a partir de $g_t(x|y, x')$ en cuyo caso se calcula el radio de aceptación dado por $r_1 = \frac{\pi_{pos}(y')}{\pi_{pos}(x')} \beta^{k_I-2}$, en el otro caso se toma la muestra de la función $g_t(x'|y', x)$ con radio de aceptación $r_0 = \frac{\pi_{pos}(y')}{\pi_{pos}(x')} \beta^{k_I-2}$.
3. Se toma $u \in [0, 1]$ con probabilidad de densidad uniforme.
4. Si $\min(1, r_1) \geq u$, se escoge $x_{k+1} = y$, en otro caso $x_{k+1} = x_k$. o si $\min(1, r_0) \geq u$, se escoge $x'_{k+1} = y'$, en otro caso $x'_{k+1} = x'_k$.
5. Volver al paso 2 y continuar hasta obtener el número de muestras deseadas.

2.1.5. Familias de muestras invariantes afines

El método MCMC presentado por Goodman y Weare notado como (GWMCMC, ver [12]), que se basa en ideas de varios caminos, presentadas t-walk. Éste modelo propone una familia de muestras de conjuntos invariantes afines, para una muestra de una densidad $\pi(x)$ con $x \in \mathbb{R}^n$, se considera obtener muestras de,

$$\prod(\vec{y}) = \pi_{A_1, b_1}(y_1) \pi_{A_2, b_2}(y_2) \cdots \pi_{A_{2n}, b_{2n}}(y_{2n}).$$

Donde $\pi_{A_i, b_i}(y_i)$ es la función de densidad de la transformación invariante afín $y_i = A_i x_i + b_i$, al igual que el t-walk se presenta una interacción entre las posiciones a modificar en cada iteración.

A diferencia del t-walk, éste modelo propone obtener muestras de $2n$ caminos para estimar n parámetros, es decir se necesitan $2(n-1)$ caminos demás para la estimación. Sin embargo la convergencia o simplicidad del modelo se debe a realizar transformaciones afines a las distribuciones propuestas para cada parámetro.

2.2. Estimadores

El enfoque bayesiano será utilizado para la obtener solución al problema inverso, pero dicha solución vendrá dada por una función de densidad de probabilidad posterior π_{pos} . Sin embargo una función de densidad de probabilidad como su nombre lo indica es una función que permitirá varios resultados puntuales a un problema dado. Por ésta razón es importante contar con estimadores que permitan extraer la suficiente información para dar solución puntual al problema.

Para ello se trabajaran con cuatro estimadores, el primero de ellos es conocido como el estimador de *máxima probabilidad* o *máximo posterior* denotado como MAP por las palabras en inglés *maximum a posteriori*, dicho estimador viene dado por:

$$\hat{\theta}_{\text{MAP}}(x) = \arg \max_{x \in \mathbb{R}} \pi(x|y). \quad (2.7)$$

Es decir, $\hat{\theta}_{\text{MAP}}$ es el punto que maximiza la función de densidad posterior, note que éste punto no es necesariamente único o no siempre existe.

Otro estimador que será utilizado es el de la *moda*, que es el punto con más repeticiones en la muestra. Éste estimador será utilizado para ver numéricamente que el MAP aparece en la muestra obtenida con mayor frecuencia que los otros números o si los números con mayor frecuencia en la muestra están cerca de la estimación MAP.

El tercer estimador que será utilizado es el estimador de la *media condicional* notado como CM por las frase en inglés *conditional mean* ver [6, pág 33], éste estimador estada dado por:

$$\hat{\theta}_{\text{MC}}(X|Y) = E(X|Y) = \int_{-\infty}^{\infty} x\pi(x|y). \quad (2.8)$$

Otro tipo de estimador que será de gran utilidad para conocer la variabilidad que presenta la muestra es la varianza condicional notado como CV por la frase *conditional variance* ver [6, pág 36], es un estimador de dispersión que permitirá conocer la proximidad de los resultados en la muestra obtenida. Éste estimador está dado por:

$$\text{var}(X|Y) = E\left(\left(X - E(X|Y)\right)^2 \mid Y\right). \quad (2.9)$$

La varianza (2.9) puede ser obtenida mediante (2.8) siempre que $\hat{\theta}_{\text{MC}}(X|Y) < \infty$.

2.3. Distribuciones normales

Se considerará algunas variables aleatorias con función de densidad normal y la relación que existe entre estas y el método de Tikhonov, ésta sección se basa en [cap. 3, [16]], donde los teoremas y ejemplos aquí presentados son tomados del mismo, salvo en algunos casos con modificaciones.

Definición 2.3.1. *Sea,*

$$\Gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \in \mathbb{R}^{n \times n},$$

es una matriz simétrica y definida positiva, donde $\Gamma_{1 \times 1} \in \mathbb{R}^{k \times k}$, $\Gamma_{2 \times 2} \in \mathbb{R}^{(n-k) \times (n-k)}$ con $n < k$, y $\Gamma_{21} = \Gamma'_{12}$.

Se define el complemento de Schur $\tilde{\Gamma}_{jj}$ de Γ_{jj} , $j = 1, 2$, por las fórmulas:

$$\tilde{\Gamma}_{22} = \Gamma_{11} - \Gamma_{12}\Gamma_{22}^{-1}\Gamma_{21} \quad y \quad \tilde{\Gamma}_{11} = \Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12}.$$

De lo anterior se puede observar que Γ_{jj} para $j = 1, 2$, son definidas positivas pues Γ es definida positiva y sus menores también lo son.

Lema 2.3.1. *Sea Γ la matriz que satisface la definición anterior, entonces los complementos de Schur son invertibles y la inversa de Γ está dada por:*

$$\Gamma^{-1} = \begin{pmatrix} \tilde{\Gamma}_{22}^{-1} & -\tilde{\Gamma}_{22}^{-1}\Gamma_{12}\Gamma_{22}^{-1} \\ -\tilde{\Gamma}_{11}^{-1}\Gamma_{21}\Gamma_{11}^{-1} & \tilde{\Gamma}_{11}^{-1} \end{pmatrix}.$$

Demostración: Considere,

$$|\Gamma| = \begin{vmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{vmatrix} \neq 0,$$

multiplicando la primera fila por $-\Gamma_{21}\Gamma_{11}^{-1}$ y sumándolo a la segunda fila, que por propiedades de los determinantes este se mantiene, se obtiene:

$$|\Gamma| = \begin{vmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{11} & \Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12} \end{vmatrix} = \begin{vmatrix} \Gamma_{11} & \Gamma_{12} \\ 0 & \tilde{\Gamma}_{11} \end{vmatrix} = |\Gamma_{11}||\tilde{\Gamma}_{11}| \neq 0.$$

Lo que implica que $|\Gamma_{11}| \neq 0$ y $|\tilde{\Gamma}_{11}| \neq 0$, así $\tilde{\Gamma}_{11}$ es invertible. De forma similar para $\tilde{\Gamma}_{22}$ multiplicando la segunda fila por $-\Gamma_{12}\Gamma_{22}^{-1}$ y sumándolo a la primera fila. Note que también se obtiene que Γ_{11} y Γ_{22} son invertibles.

Ahora considere el sistema $y = \Gamma x$, tal que,

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Desarrollando el anterior sistema se obtienen las siguientes ecuaciones,

$$\begin{aligned} (1) \quad \Gamma_{11}x_1 + \Gamma_{12}x_2 &= y_1 \Rightarrow x_1 = \Gamma_{11}^{-1}(y_1 - \Gamma_{12}x_2). \\ (2) \quad \Gamma_{21}x_1 + \Gamma_{22}x_2 &= y_2 \Rightarrow x_2 = \Gamma_{22}^{-1}(y_2 - \Gamma_{21}x_1). \end{aligned}$$

Sustituyendo x_2 en (1) se obtiene,

$$\begin{aligned} \Gamma_{11}x_1 + \Gamma_{12}\Gamma_{22}^{-1}(y_2 - \Gamma_{21}x_1) &= y_1 \\ \Gamma_{11}x_1 + \Gamma_{12}\Gamma_{22}^{-1}y_2 - \Gamma_{12}\Gamma_{22}^{-1}\Gamma_{21}x_1 &= y_1 \\ \underbrace{(\Gamma_{11} - \Gamma_{12}\Gamma_{22}^{-1}\Gamma_{21})}_{\tilde{\Gamma}_{22}}x_1 &= y_1 - \Gamma_{12}\Gamma_{22}^{-1}y_2 \\ \tilde{\Gamma}_{22}x_1 &= y_1 - \Gamma_{12}\Gamma_{22}^{-1}y_2 \\ x_1 &= \tilde{\Gamma}_{22}^{-1}y_1 - \tilde{\Gamma}_{22}^{-1}\Gamma_{12}\Gamma_{22}^{-1}y_2. \end{aligned}$$

De forma análoga sustituyendo en (2),

$$\begin{aligned} \Gamma_{21}\Gamma_{11}^{-1}(y_1 - \Gamma_{12}x_2) + \Gamma_{22}x_2 &= y_2 \\ \underbrace{(\Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12})}_{\tilde{\Gamma}_{11}}x_2 &= y_2 - \Gamma_{21}\Gamma_{11}^{-1}y_1 \\ x_2 &= \tilde{\Gamma}_{11}^{-1}y_2 - \tilde{\Gamma}_{11}^{-1}\Gamma_{21}\Gamma_{11}^{-1}y_1. \end{aligned}$$

Expresando las ecuaciones obtenidas para x_1 y x_2 en forma matricial se obtiene,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{\begin{pmatrix} \tilde{\Gamma}_{22}^{-1} & -\tilde{\Gamma}_{22}^{-1}\Gamma_{12}\Gamma_{22}^{-1} \\ -\tilde{\Gamma}_{11}^{-1}\Gamma_{21}\Gamma_{11}^{-1} & \tilde{\Gamma}_{11}^{-1} \end{pmatrix}}_{\Gamma^{-1}} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \quad \square$$

Nota Dado que $\Gamma^T = \Gamma$, se tiene que $(\Gamma^T)^{-1} = \Gamma^{-1}$, así $(\Gamma^{-1})^T = \Gamma^{-1}$. Es decir la inversa es simétrica y se obtiene la igualdad

$$\left(\tilde{\Gamma}_{22}^{-1}\Gamma_{12}\Gamma_{22}^{-1}\right)^T = \tilde{\Gamma}_{11}^{-1}\Gamma_{21}\Gamma_{11}^{-1} = \Gamma_{22}^{-1}\Gamma_{21}\tilde{\Gamma}_{22}^{-1}.$$

Teorema 2.3.1. Sea $X : \Omega \rightarrow \mathbb{R}^n$ y $Y : \Omega \rightarrow \mathbb{R}^k$ dos variables aleatorias gaussianas, con función de densidad de probabilidad,

$$\pi(x, y) \propto \exp \left(-\frac{1}{2} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}^T \underbrace{\begin{bmatrix} \tilde{\Gamma}_{22}^{-1} & -\tilde{\Gamma}_{22}^{-1}\Gamma_{12}\Gamma_{22}^{-1} \\ -\tilde{\Gamma}_{11}^{-1}\Gamma_{21}\Gamma_{11}^{-1} & \tilde{\Gamma}_{11}^{-1} \end{bmatrix}}_{\Gamma^{-1}} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \right).$$

Entonces la función de densidad de probabilidad de X condicionada a $Y = y$, $\pi(x|y) : \mathbb{R}^n \rightarrow \mathbb{R}^+$ es de la forma,

$$\pi(x|y) \propto \exp \left(-\frac{1}{2} (x - \bar{x})^T \tilde{\Gamma}_{22}^{-1} (x - \bar{x}) \right).$$

Donde $\bar{x} = x_0 + \Gamma_{12}\Gamma_{22}^{-1}(y - y_0)$.

Demostración: Considere la función de densidad de probabilidad $\pi(x, y)$ que desarrollando algunos términos se obtiene,

$$\begin{aligned} \pi(x, y) &\propto \exp \left(-\frac{1}{2} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}^T \underbrace{\begin{bmatrix} \tilde{\Gamma}_{22}^{-1} & -\tilde{\Gamma}_{22}^{-1}\Gamma_{12}\Gamma_{22}^{-1} \\ -\tilde{\Gamma}_{11}^{-1}\Gamma_{21}\Gamma_{11}^{-1} & \tilde{\Gamma}_{11}^{-1} \end{bmatrix}}_{\Gamma^{-1}} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \right) \\ &= \exp \left(-\frac{1}{2} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}^T \begin{bmatrix} \tilde{\Gamma}_{22}^{-1}(x - x_0) - \tilde{\Gamma}_{22}^{-1}\Gamma_{12}\Gamma_{22}^{-1}(y - y_0) \\ \tilde{\Gamma}_{11}^{-1}(y - y_0) - \tilde{\Gamma}_{11}^{-1}\Gamma_{21}\Gamma_{11}^{-1}(x - x_0) \end{bmatrix} \right). \end{aligned}$$

Por lo visto anteriormente, con la simetría de Γ^{-1} , se obtiene,

$$\begin{aligned} \pi(x, y) &\propto \exp \left(-\frac{1}{2} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}^T \begin{bmatrix} \tilde{\Gamma}_{22}^{-1}(x - x_0) - \tilde{\Gamma}_{22}^{-1}\Gamma_{12}\Gamma_{22}^{-1}(y - y_0) \\ \tilde{\Gamma}_{11}^{-1}(y - y_0) - \left(\tilde{\Gamma}_{22}^{-1}\Gamma_{21}\Gamma_{11}^{-1} \right)^T (x - x_0) \end{bmatrix} \right) \\ &= \exp \left(-\frac{1}{2} \left((x - x_0)^T \tilde{\Gamma}_{22}^{-1} (x - x_0) - (x - x_0)^T \tilde{\Gamma}_{22}^{-1} \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) + (y - y_0)^T \tilde{\Gamma}_{11}^{-1} (y - y_0) \right. \right. \\ &\quad \left. \left. - (y - y_0)^T \left(\tilde{\Gamma}_{22}^{-1} \Gamma_{12} \Gamma_{22}^{-1} \right)^T (x - x_0) \right) \right) \\ &= \exp \left(-\frac{1}{2} \left((x - x_0)^T \tilde{\Gamma}_{22}^{-1} (x - x_0) - (x - x_0)^T \tilde{\Gamma}_{22}^{-1} \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) + (y - y_0)^T \tilde{\Gamma}_{11}^{-1} (y - y_0) \right. \right. \\ &\quad \left. \left. - (y - y_0)^T \Gamma_{22}^{-1} \Gamma_{21} \tilde{\Gamma}_{22}^{-1} (x - x_0) \right) \right) \end{aligned}$$

Tomando solo el exponente, sumando y restando $(y - y_0)^T \Gamma_{22}^{-1} \Gamma_{21} \tilde{\Gamma}_{22}^{-1} \Gamma_{12} \Gamma_{22}^{-1} (y - y_0)$ para completar cuadrados, se obtiene:

$$\begin{aligned}
& (y - y_0)^T \tilde{\Gamma}_{11}^{-1} (y - y_0) - (y - y_0)^T \Gamma_{22}^{-1} \Gamma_{21} \tilde{\Gamma}_{22}^{-1} (x - x_0) - (x - x_0)^T \tilde{\Gamma}_{22}^{-1} \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) \\
& + (x - x_0)^T \tilde{\Gamma}_{22}^{-1} (x - x_0) + (y - y_0)^T \Gamma_{22}^{-1} \Gamma_{21} \tilde{\Gamma}_{22}^{-1} \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) - (y - y_0)^T \Gamma_{22}^{-1} \Gamma_{21} \tilde{\Gamma}_{22}^{-1} \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) \\
& = \left((x - x_0)^T \tilde{\Gamma}_{22}^{-1} - (y - y_0)^T \Gamma_{22}^{-1} \Gamma_{12} \tilde{\Gamma}_{22}^{-1} \right) \left((x - x_0) - \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) \right) + (y - y_0)^T \tilde{\Gamma}_{11}^{-1} (y - y_0) \\
& - (y - y_0)^T \Gamma_{22}^{-1} \Gamma_{21} \tilde{\Gamma}_{22}^{-1} \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) \\
& = \left((x - x_0)^T - (y - y_0)^T \Gamma_{22}^{-1} \Gamma_{12} \right) \tilde{\Gamma}_{22}^{-1} \left((x - x_0) - \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) \right) + c \\
& = \left((x - x_0) - \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) \right)^T \tilde{\Gamma}_{22}^{-1} \left((x - x_0) - \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) \right) + c.
\end{aligned}$$

Donde $c = (y - y_0)^T \left(\tilde{\Gamma}_{11}^{-1} - \Gamma_{22}^{-1} \Gamma_{21} \tilde{\Gamma}_{22}^{-1} \Gamma_{12} \Gamma_{22}^{-1} \right) (y - y_0)$. Así

$$\pi(x, y) \propto \exp \left(-\frac{1}{2} \left(x - x_0 - \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) \right)^T \tilde{\Gamma}_{22}^{-1} \left(x - x_0 - \Gamma_{12} \Gamma_{22}^{-1} (y - y_0) \right) \right). \quad \square$$

Teorema 2.3.2. Sea $Y = AX + E$ tal que $X : \Omega \rightarrow \mathbb{R}^n$ y $E : \Omega \rightarrow \mathbb{R}^m$ son variables aleatorias gaussianas mutuamente independientes,

$$X \sim N(x_0, \Gamma_{apr}), \quad E \sim N(e_0, \Gamma_{ruido}),$$

$\Gamma_{apr} \in \mathbb{R}^{n \times n}$ y $\Gamma_{ruido} \in \mathbb{R}^{k \times k}$ son definidas positivas. Además $A \in \mathbb{R}^{k \times n}$ es una matriz conocida. Entonces la función de densidad de probabilidad posterior de X condicionado a $Y = y$ es

$$\pi(x|y) \propto \exp \left(-\frac{1}{2} (x - \bar{x})^T \Gamma_{pos}^{-1} (x - \bar{x}) \right),$$

donde

$$\bar{x} = x_0 + \Gamma_{apr} A^T (A \Gamma_{apr} A^T + \Gamma_{ruido})^{-1} (y - Ax_0 - e_0) \quad (2.10)$$

y

$$\Gamma_{pos} = \Gamma_{apr} - \Gamma_{apr} A^T (A \Gamma_{apr} A^T + \Gamma_{ruido})^{-1} A \Gamma_{apr} \quad (2.11)$$

2.4. Densidad gaussiana

En ésta sección se presentarán algunos ejemplos especiales de densidad gaussiana para el sistema $Y = AX + E$, ésta sección se basa en [cap. 3, [16]], donde los teoremas y ejemplos aquí presentados son tomados del mismo, salvo en algunos casos con modificaciones.

Ejemplo 2.4.1. Considere el caso $Y = AX + E$, con A una matriz conocida, $X \sim N(0, \gamma^2 I)$ y $E \sim N(0, \sigma^2 I)$, entonces se tendrá que para la densidad posterior la media viene dada por:

$$\bar{x} = \gamma^2 A^T (\gamma^2 A A^T + \sigma^2)^{-1} y = A^T (A A^T + \alpha I)^{-1} y,$$

donde $\alpha = \sigma^2 / \gamma^2$.

Observe que,

$$\pi_{pos}(x|y) = \frac{\pi_{X,Y}(x, y)}{\pi_Y(y)} \propto \pi_{X,E}(x, y - Ax) = \pi_{apr}(x) \cdot \pi_{ruido}(y - Ax).$$

La última igualdad es debido a que X es independiente de E , así,

$$\begin{aligned}
\pi_{pos}(x|y) &\propto \exp\left(-\frac{1}{2}\left(\frac{x^T x}{\gamma^2} + \frac{e^T e}{\sigma^2}\right)\right) \\
&= \exp\left(-\frac{1}{2\gamma^2\sigma^2}(x^T x\sigma^2 + (y - Ax)^T(y - Ax)\gamma^2)\right) \\
&= \exp\left(-\frac{1}{2\sigma^2\gamma^2}(x^T x\sigma^2 + y^T y\gamma^2 - x^T A^T y\gamma^2 - y^T Ax\gamma^2 + x^T A^T Ax\gamma^2)\right) \\
&\propto \exp\left(-\frac{1}{2\sigma^2\gamma^2}(x^T x\sigma^2 - x^T A^T y\gamma^2 - y^T Ax\gamma^2 + x^T A^T Ax\gamma^2)\right)
\end{aligned}$$

Se busca llegar en el exponente a la forma:

$$(x - \bar{x})^T \Gamma_{pos}^{-1} (x - \bar{x}) = x^T \Gamma_{pos}^{-1} x - x^T \Gamma_{pos}^{-1} \bar{x} - \bar{x}^T \Gamma_{pos}^{-1} x + \bar{x}^T \Gamma_{pos}^{-1} \bar{x}.$$

Comparando términos se obtiene,

$$\begin{aligned}
x^T \Gamma_{pos}^{-1} x &= x^T x\sigma^2 + x^T A^T Ax\gamma^2 = x^T (\sigma^2 I + \gamma^2 A^T A)x = x^T \underbrace{\gamma^2 (\alpha I + A^T A)}_{\Gamma_{pos1}^{-1}} x \\
x^T \Gamma_{pos}^{-1} \bar{x} &= x^T A^T y\gamma^2 = x^T \Gamma_{pos}^{-1} \Gamma_{pos} A^T y\gamma^2 = x^T \underbrace{\gamma^2 (\alpha I + A^T A)}_{\Gamma_{pos1}^{-1}} \underbrace{\gamma^{-2} (\alpha I + A^T A)^{-1} A^T y}_{\Gamma_{pos1}} \gamma^2 \\
&= x^T \underbrace{\gamma^2 (\alpha I + A^T A)}_{\Gamma_{pos1}^{-1}} \underbrace{(\alpha I + A^T A)^{-1} A^T y}_{\bar{x}} \\
\bar{x}^T \Gamma_{pos}^{-1} x &= y^T A\gamma^2 x = y^T A(\alpha I + A^T A)^{-1} (\alpha I + A^T A)\gamma^2 x = ((\alpha I + A^T A)^{-1} A^T y)^T \gamma^2 (\alpha I + A^T A)x \\
&= \underbrace{((\alpha I + A^T A)^{-1} A^T y)^T}_{\bar{x}^T} \underbrace{\gamma^2 (\alpha I + A^T A)}_{\Gamma_{pos1}^{-1}} x \\
\bar{x}^T \Gamma_{pos}^{-1} \bar{x} &= y^T A(\alpha I + A^T A)^{-1} \gamma^2 (\alpha I + A^T A) (\alpha I + A^T A)^{-1} A^T y \\
&= \gamma^2 y^T A(\alpha I + A^T A)^{-1} A^T y.
\end{aligned}$$

Luego,

$$\begin{aligned}
\pi_{pos}(x|y) &\propto \exp\left(-\frac{\gamma^{-2}\sigma^{-2}}{2}\left((x - \bar{x})^T \Gamma_{pos1}^{-1} (x - \bar{x}) - \gamma^2 y^T A(\alpha I + A^T A)^{-1} A^T y\right)\right) \\
&\propto \exp\left(-\frac{1}{2}(x - \bar{x})^T \Gamma_{pos}^{-1} (x - \bar{x})\right).
\end{aligned}$$

Donde:

$$\bar{x} = (\alpha I + A^T A)^{-1} A^T y \quad (2.12)$$

$$\Gamma_{pos} = \gamma^2 \sigma^2 \Gamma_{pos1} = \sigma^2 \gamma^2 (\alpha I + A^T A)^{-1} \gamma^{-2} = (\gamma^{-2} I + \sigma^{-2} A^T A)^{-1}. \quad (2.13)$$

Ahora el siguiente ejemplo presenta la relación entre algunas densidades gaussianas y el principio de Morozov.

Ejemplo 2.4.2. Sea X una variable aleatoria distribuida normalmente con media $x_0 \in \mathbb{R}^n$ y matriz de covarianza $\Gamma \in \mathbb{R}^{n \times n}$, es decir $X \sim N(x_0, \Gamma)$.

Suponga que se tiene datos con un ruido aditivo con distribución conocida, es decir,

$$Y = AX + B, \quad (2.14)$$

donde B representa el ruido aditivo y tiene una distribución $\pi(b)$, luego $Y - AX$ tiene la misma distribución de B , dado que X son los datos a recuperar a partir de los datos dados Y , y la distribución obtenida es $\pi(Y|X) = \pi(b)$. Haciendo uso del Teorema de Bayes se puede decir que X tiene una distribución proporcional conocida, es decir,

$$\pi(X|Y) \propto \pi(Y|X)\pi(X).$$

En términos del lenguaje bayesiano lo anterior se traduce a,

$$\pi_{\text{pos}}(X) \propto \pi(Y|X)\pi_{\text{apr}}(X),$$

donde π_{pos} y π_{apr} definen la distribución posterior y la distribución a priori respectivamente. Por ejemplo considere (2.14) con ruido aditivo gaussiano, es decir $B \sim N(0, \Gamma_{\text{ruido}})$. Ahora se asume un modelo a priori gaussiano para X ,

$$X \sim N(x_0, \alpha^{-1}\Gamma_{\text{apr}}),$$

con $\alpha > 0$. Luego la distribución a priori es conocida y de la forma

$$\pi(x|\alpha) = \frac{\alpha^{n/2}}{\sqrt{(2\pi)^n |\Gamma_{\text{apr}}|}}. \quad (2.15)$$

Asumiendo una distribución a priori para α dada por

$$\pi_{\text{apr}}(\alpha) = \frac{\alpha}{\alpha_0^2} \exp\left(-\frac{1}{2} \left(\frac{\alpha}{\alpha_0}\right)^2\right), \quad \alpha > 0, \text{ y una constante } \alpha_0 > 0.$$

Realizando cálculos se halla la distribución posterior para x , obteniendo la siguiente función de densidad,

$$\pi(x|\alpha, y) \sim \exp\left(-\frac{1}{2}\alpha\|x\|^2 - \frac{1}{2\sigma^2}\|y - Ax\|^2\right), \quad (2.16)$$

con $\sigma^2 I = \Gamma_{\text{ruido}}$, y $\Gamma_{\text{apr}} = I$ e I nota la matriz identidad. Aplicando (1.5) para solucionar (2.14) se debe hacer mediante métodos MCMC dado que las variables tienen distribuciones asociadas y el parámetro de regularización δ debe cumplir la siguiente condición por el principio de Morozov,

$$\|Ax_\delta - y\|^2 = \text{Tr}(\Gamma_n) = n\sigma^2,$$

y el parámetro α por el principio de Morozov debe ser,

$$\alpha_{\text{Morozov}} = \frac{\delta}{\sigma^2}.$$

Buscando generalizar el ejemplo anterior, se presenta la relación que existe entre la regularización de Tikhonov y densidades gaussianas con varianza conocida.

Ejemplo 2.4.3. *Suponga que se busca solución a la ecuación $Y = AX + E$ por los métodos clásicos de regularización. Suponga además que $X \in \mathbb{R}^n$ representa valores discretizados de alguna función $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, con una distribución a priori conocida y dos veces diferenciable sobre D . Para expresar ésta información como una restricción, se presenta el funcional de Tikhonov,*

$$F_\alpha(x) = \|Ax - y\|^2 + \alpha\|x\|^2.$$

Del ejemplo anterior se obtiene que para $X \sim N(0, \gamma^2 I)$ y $E \sim N(0, \sigma^2 I)$ la distribución posterior está dada por:

$$\pi_{pos}(x|y) \propto \exp\left(-\frac{1}{2}\left(\frac{x^T x}{\gamma^2} + \frac{1}{\sigma^2}(y - Ax)^T(y - Ax)\right)\right),$$

donde se puede observar que,

$$\frac{1}{\gamma^2}x^T x = \frac{1}{\gamma^2}\|x\|^2 \quad y \quad \frac{1}{\sigma^2}(y - Ax)^T(y - Ax) = \frac{1}{\sigma^2}\|y - Ax\|^2.$$

Entonces,

$$\pi_{pos}(x|y) \propto \exp\left(-\frac{1}{2\gamma^2}\|x\|^2 - \frac{1}{2\sigma^2}\|y - Ax\|^2\right). \quad (2.17)$$

Se define el potencial posterior como $V(x|y) = \frac{1}{2\gamma^2}\|x\|^2 + \frac{1}{2\sigma^2}\|y - Ax\|^2 = \frac{1}{2\sigma^2}F_\alpha(x)$, entonces (2.17) queda expresado de la forma,

$$\pi_{pos}(x|y) \propto \exp(-V(x|y)) = \exp\left(-\frac{1}{2\sigma^2}F_\alpha(x)\right).$$

En la gráfica de una densidad gaussiana se puede observar que el parámetro que maximiza la función es la media, por el exponente negativo el maximizar la función es equivalente a minimizar el exponente. Anteriormente se observó que la solución al funcional de Tikhonov (el término que minimiza el funcional) está dado por (1.5). Con esto se obtiene que $\bar{x} = x_\alpha$, con $\alpha = \frac{\sigma^2}{\gamma^2}$.

En el anterior ejemplo se evidencia como se relaciona el funcional con Tikhonov con una distribución normal, donde la media que es precisamente la que maximiza la función de distribución normal, coincide con la solución de Tikhonov que es la que maximiza dicho funcional. Mostrado así como es posible obtener la misma solución trabajando con un modelo determinista o aleatorio.

Con los ejemplos que se presentan en ésta sección se puede observar algunas relaciones entre las densidades gaussianas y los métodos de regularización, lo que presenta una forma de pasar de un modelo aleatorio a un modelo determinista. De la misma forma en que es posible en algunas ocasiones el pasar de un modelo aleatorio a un modelo determinista, se puede realizar el procedimiento inverso siguiendo ciertos supuestos sobre los parámetros y los datos, una forma de hacer esto se presenta a continuación, donde se utiliza una metodología bayesiana para estimar parámetros de modelos deterministas a partir de supuestos sobre estos asumiendo que son variables aleatorias con cierta distribución dada.

2.5. Introducción a la estimación de parámetros de una ecuación diferencial

En ésta sección presentamos un ejemplo de aplicación de los métodos estadísticos para la solución de problemas inversos. Se presenta una introducción al problema inverso de estimar algunos parámetros de una ecuación diferencial ordinaria que modela el crecimiento de una población. El problema directo consiste en resolver un problema de valor inicial asociado a la ecuación diferencial ordinaria considerada. La ecuación diferencial usada es ecuación de Verhulst (o modelo de crecimiento logístico). El problema de valor inicial es,

$$\begin{cases} \frac{dx(t)}{dt} &= rx(t) \left(1 - \frac{x(t)}{k}\right) \\ x(0) &= x_0, \end{cases} \quad (2.18)$$

los parámetros de este problema de valor inicial son x_0 , condición inicial, r , tasa de crecimiento y k la capacidad de carga del ecosistema considerado. Ésta ecuación modela una población en donde los individuos interactúan bajo la hipótesis de recursos limitados. Por ejemplo, en el caso de una especie animal, el ecosistema que los alberga tiene alimento limitado. Ésta ecuación también puede simular la propagación de rumores, extensión de innovación tecnológica, entre otras aplicaciones.

El problema de valor inicial (2.18) tiene solución analítica dada por,

$$x(t) = \frac{x_0 k}{x_0 + (k - x_0)e^{-rt}}. \quad (2.19)$$

Note que en el problema directo, dados los valores de los parámetros mencionados (condición inicial x_0 , tasa de crecimiento r y capacidad de carga k) se pueden calcular soluciones $x(t)$ que describen el comportamiento que se busca entender o predecir de la población estudiada.

Por otro lado, en el problema inverso (usando métodos estadísticos) debemos estimar la distribución de los parámetros tasa de crecimiento r y capacidad de carga k de la población observada. En lo que sigue de ésta sección usaremos la notación $\theta = (r, k)$ para referirnos a los parámetros objetos del ejercicio de estimación. Aquí solamente presentaremos el caso en donde, para obtener la distribución de $\theta = (r, k)$ se asume conocido el valor de x_0 y se dispone de datos observados de la evolución de la población.

Para calcular la distribución de θ se puede usar el enfoque de estadística bayesiana para la solución de problemas inversos. Recuerde que en este enfoque se asume que los datos observados tienen un algún ruido o son obtenidos de una distribución de probabilidad. Se tiene también que los parámetros requeridos también cuentan con una distribución la cual debemos muestrear. En el enfoque bayesiano asignamos al parámetro θ una distribución apriori que permita obtener una distribución posterior (o final).

Para este ejemplo se usa una implementación del algoritmo GWCMC en el software Matlab¹. El algoritmo GWCMC es una versión del algoritmo de familias de muestras invariantes afines enunciado en la sección 2.1.5 y presentado por [12]. Los parámetros de entrada de éste algoritmo son tres;

¹La implementación particular usada se puede encontrar en la página oficial de Matlab “<https://es.mathworks.com/matlabcentral/fileexchange/49820-ensemble-mcmc-sampler>”.

- 1) Una matriz de valores iniciales de tamaño $m \times 2m$, donde m es el número de parámetros que se desean recuperar. Para el ejemplo presentado en ésta sección son dos parámetros $\theta = (r, k)$ (tasa de cambio y capacidad de carga).
- 2) El logaritmo de la función de densidad de probabilidad de la distribución posterior.
- 3) El número de iteraciones que se desean realizar. Mientras más grande sea éste número, se tiene una mayor probabilidad de obtener un número de muestra de mayor tamaño.

A continuación presentamos un ejemplo práctico de cómo muestrear la distribución estimada de los parámetros r y k cuando estos hacen parte del modelo (2.18). En ese ejemplo se se usarán datos sintéticos generados a partir de la solución analítica del problema de valor inicial (2.18) modificados aditivamente con un ruido pequeño de tipo Poisson.

2.5.1. Generación de datos simulados

Considere la ecuación diferencial (2.18) con parámetros de tasa de cambio $r = 6$, condición inicial $x_0 = 1$ y capacidad de carga $k = 15$. Se obtienen n datos $x(t_i, \theta)$ de la solución (2.19) evaluados en tiempos t_i para $i = 1, \dots, n$, con parámetros de tasa de cambio y capacidad de carga representados por $\theta = (6, 15)$. Podemos entonces simular el comportamiento de la ecuación a lo largo del tiempo con estos valores de parámetros. Ésta simulación se pueden obtener mediante algún método numérico conocido como *método de Euler* o alguno de los *métodos Runge-Kutta*, entre otros. Alternativamente podemos usar la solución explícita (2.19). Obtenemos los datos presentados en la Tabla (2.1) y la Figura 2.1.

Tiempo	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
Dato	1.00	1.05	1.11	1.18	1.24	1.31	1.39	1.47	1.55	1.63	1.72
Tiempo	0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18	0.19	0.20	0.21
Dato	1.82	1.91	2.02	2.12	2.24	2.35	2.48	2.60	2.73	2.87	3.01
Tiempo	0.22	0.23	0.24	0.25	0.26	0.27	0.28	0.29	0.30	0.31	0.32
Dato	3.16	3.31	3.47	3.63	3.80	3.97	4.15	4.33	4.52	4.71	4.91
Tiempo	0.33	0.34	0.35	0.36	0.37	0.38	0.39	0.40	0.41	0.42	0.43
Dato	5.11	5.31	5.52	5.73	5.95	6.16	6.38	6.60	6.83	7.05	7.27
Tiempo	0.44	0.45	0.46	0.47	0.48	0.49	0.50	0.51	0.52	0.53	0.54
Dato	7.50	7.72	7.95	8.17	8.39	8.62	8.83	9.05	9.26	9.48	9.68
Tiempo	0.55	0.56	0.57	0.58	0.59	0.60	0.61	0.62	0.63	0.64	0.65
Dato	9.89	10.09	10.28	10.48	10.66	10.84	11.02	11.20	11.36	11.53	11.68
Tiempo	0.66	0.67	0.68	0.69	0.70	0.71	0.72	0.73	0.74	0.75	0.76
Dato	11.84	11.98	12.12	12.26	12.39	12.52	12.64	12.76	12.87	12.98	13.08
Tiempo	0.77	0.78	0.79	0.80	0.81	0.82	0.83	0.84	0.85	0.86	0.87
Dato	13.18	13.27	13.36	13.45	13.53	13.60	13.68	13.75	13.82	13.88	13.94
Tiempo	0.88	0.89	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98
Dato	14.00	14.05	14.10	14.15	14.20	14.24	14.28	14.32	14.36	14.40	14.43
Tiempo	0.99	1	—	—	—	—	—	—	—	—	—
Dato	14.46	14.49	—	—	—	—	—	—	—	—	—

TABLA 2.1. Datos de la solución de la ecuación logística en el intervalo $[0,1]$ con tasa de cambio $r = 6$ y capacidad de carga $k = 15$ usando la solución analítica.

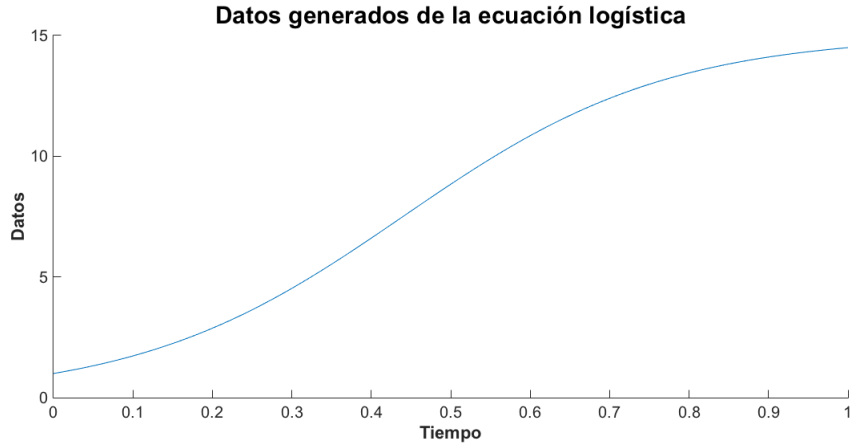


FIGURA 2.1. Solución de la ecuación logística en el intervalo $[0,1]$ con tasa de cambio $r = 6$ y capacidad de carga $k = 15$ usando la solución analítica para obtener resultados precisos.

2.5.2. Problema inverso

En el problema inverso se tendrán datos z_i con $i = 1, \dots, n$. Estos datos provienen de algún comportamiento poblacional que se presume puede ser representado mediante (2.18) con x_0 conocida.

Así, el fin es poder encontrar los valores que representan la tasa de cambio r y la capacidad de carga k en la ecuación logística a partir de los datos dados inicialmente. Para esto se asume como ejemplo práctico que los datos provienen de distribuciones Poisson² con parámetro $\lambda(t_i) = x(t_i, \theta)$ con $t_i \in [0, 1]$ y $\theta = (r, k)$.

Más precisamente, se asumen datos independientes $z_i \sim \text{Poisson}(x(t_i, \theta))$ para $i = 1, 2, \dots, 101$, con parámetro dado por los datos reales de la solución de la ecuación logística (2.19), el valor $x(t_i, \theta)$ no es conocido, sin embargo los datos z_i si lo son. En la Figura (2.2) se puede observar como son estos datos (z_i) en un periodo de tiempo $[0,1]$.

²Se escoge ésta distribución dada su relación con la probabilidad de ocurrencia de un determinado número de eventos en un tiempo determinado.

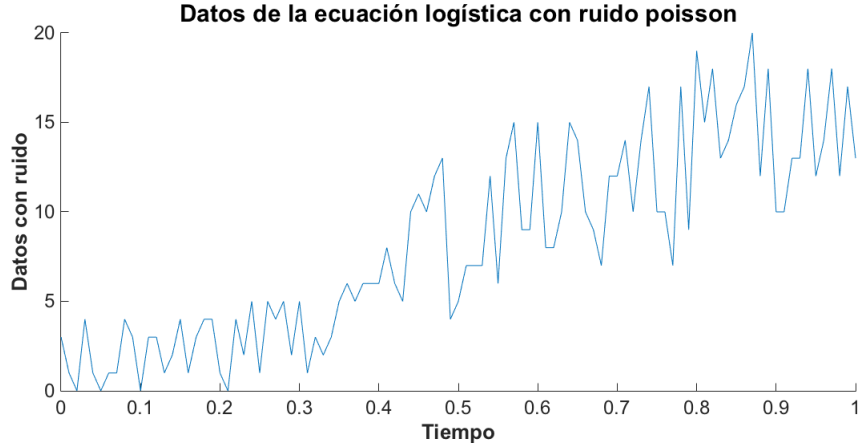


FIGURA 2.2. Datos de la ecuación logística en el intervalo $[0,1]$ con un ruido Poisson de parámetro $x(t_i, \theta)$ para $i = 1, \dots, 101$ y $\theta = (r, k)$ representa los parámetros de tasa de cambio y capacidad de carga.

Para encontrar los parámetros de tasa de cambio r y capacidad de carga k de la ecuación logística, se emplea una metodología bayesiana, que consiste en asumir distribuciones apriori para la tasa de cambio r y la capacidad de carga k dadas por $r \sim \text{Gamma}(1, 1)$ y $k \sim \text{Gamma}(1, 1)$. Esto es,

$$\pi_{apr}(r) = e^{-r} \quad \text{y} \quad \pi_{apr}(k) = e^{-k}. \quad (2.20)$$

Dado que se asumen los datos z_i independientes con distribución Poisson, se tiene como función de densidad conjunta o verosimilitud para $Z = (z_1, z_2, \dots, z_{101})$,

$$\pi(Z|\theta) = \prod_{i=1}^{101} \left((x(t_i, \theta))^{z_i} \cdot \frac{e^{-x(t_i, \theta)}}{z_i!} \right). \quad (2.21)$$

Luego, por el teorema de Bayes se obtiene el modelo jerárquico,

$$\begin{aligned} \pi_{\text{pos}}(\theta|Z) &= \prod_{i=1}^{101} \left((x(t_i, \theta))^{z_i} \cdot \frac{e^{-x(t_i, \theta)}}{z_i!} \right) \cdot e^{-r} \cdot e^{-k} \\ &= \prod_{i=1}^{101} \left((x(t_i, \theta))^{z_i} \cdot \frac{e^{-x(t_i, \theta)}}{z_i!} \right) \cdot e^{-(r+k)}. \end{aligned} \quad (2.22)$$

La función logposterior de la posterior (2.22) estada dada por,

$$\ln(\pi_{\text{pos}}(\theta|Z)) = \sum_{i=1}^{101} \left(z_i \ln(x(t_i, \theta)) - x(t_i, \theta) - \ln(z_i!) \right) - (r + k). \quad (2.23)$$

Una vez se obtiene la logposterior se procede a obtener muestras de dicha distribución con el algoritmo GWCMC mencionado anteriormente obteniendo la distribución para $x(t_i, \theta)$ de dos posibles formas diferentes con el fin de comprar las soluciones obtenidas del uso de un modelo dependiente de la solución explícita y de un modelo dependiente de la solución numérica;

- 1) En la primera forma se considera que $x(t_i, \theta)$ es la solución analítica del problema de valor inicial (2.19) asociado a la ecuación logística. Se realiza el muestreo evaluando los puntos t_i y los valores $\theta = (r, k)$ proporcionados por el algoritmo bayesiano en (2.19).
- 2) En la segunda forma se considera el problema de valor inicial (2.18), con $x(t_i, \theta)$ la solución de la ecuación en cada punto t_i del tiempo. Obteniendo ésta al resolver dicha ecuación con el método Runge-Kutta de orden 4 (RK4) que es de rápida convergencia, realizando el muestreo evaluando los puntos t_i y los valores $\theta = (r, k)$ proporcionados por el algoritmo bayesiano en (2.18). Éste método es más demorado que la forma 1, puesto que por cada posible punto que proporcione el algoritmo se debe realizar la solución de la ecuación logística (2.18) por el método numérico, haciendo que se tarde más tiempo en obtener las muestras.

Una vez implementadas ambas metodologías se obtienen los resultados presentados en las Figuras 2.3, 2.4 y 2.5. En la Figura 2.4 se puede observar que los resultados del muestreo, usando ya sea la solución analítica o la solución numérica, tienen un comportamiento similar. Esto debido a las buenas propiedades de aproximación del método numérico para el problema de valor inicial. También se puede observar una buena convergencia en el MCMC pues para 10000 iteraciones las muestras aceptadas se aglomeran al rededor de los valores exactos de los parámetros como se puede ver en la Figura 2.3.

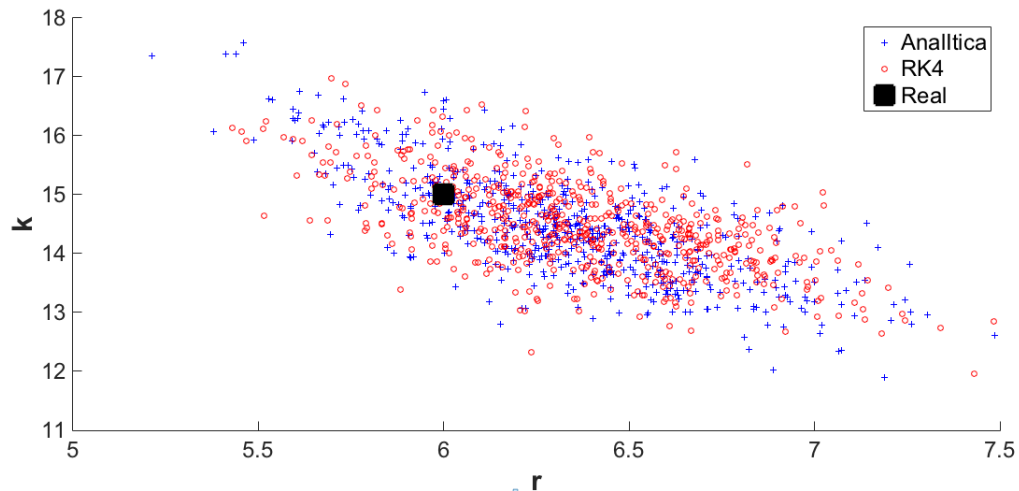


FIGURA 2.3. Comportamiento de las muestras para 10000 iteraciones

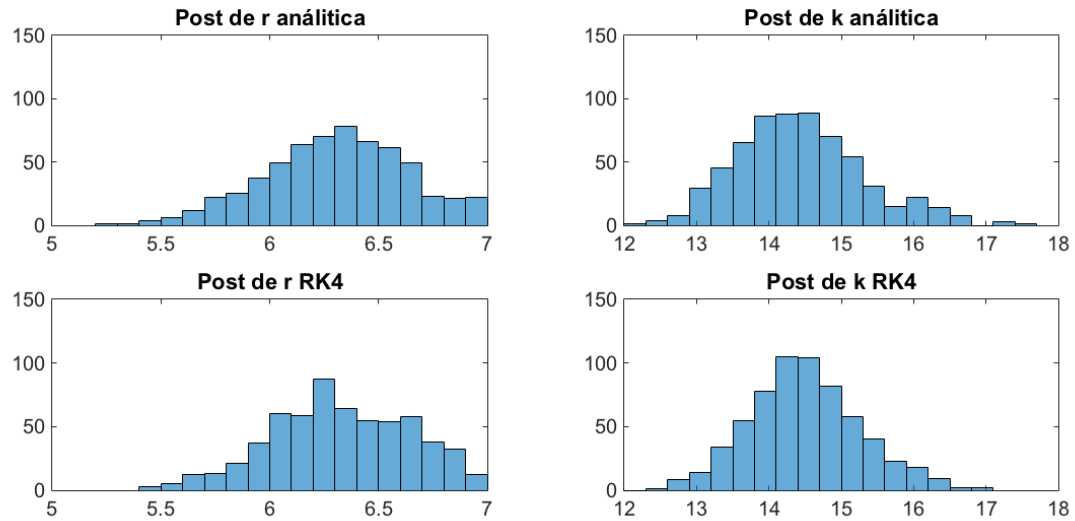


FIGURA 2.4. Histograma de las muestras para 10000 iteraciones

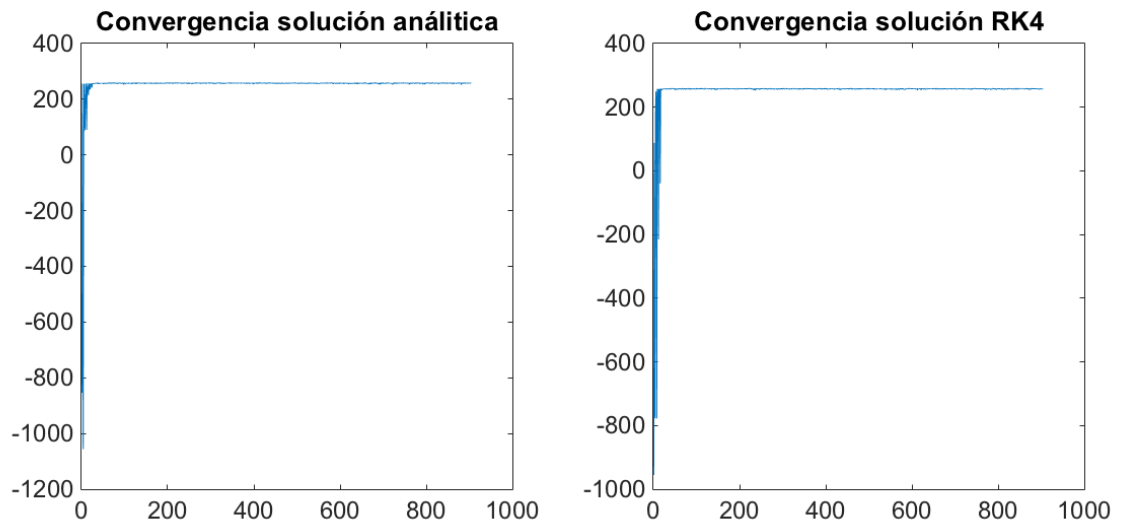


FIGURA 2.5. Gráfico del logaritmo de las probabilidades de aceptación de las muestras 10000 iteraciones

Con los anteriores resultados se puede concluir que con un buen método numérico (de buena convergencia) se pueden obtener resultados casi idénticos que si se estuviera trabajando con la solución analítica de las ecuaciones. Lo cual facilitará el encontrar los parámetros de ecuaciones que no cuentan con solución explícita y que se evidencian en fenómenos de la vida cotidiana.

2.6. Campo aleatorio de Markov gaussiano

Se introduce el *campo aleatorio de Markov gaussiano* (CAMG) basado en [2]. CAMG es un campo de vectores aleatorios con distribución gaussiana y además con la propiedad

de Markov, la matriz de covarianzas de ésta distribución depende de la discretización del dominio y las condiciones de frontera de la ecuación diferencial que se ésta trabajando.

Sea $X = (x_1, \dots, x_n)$ y $X_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ el vector X sin el i -ésimo término, se llama condicional completa a $x_i|X_{-i}$ y se le asigna una distribución gaussiana de la forma

$$x_i|X_{-i} \sim \text{Nor} \left(\sum_{j \neq i} \beta_{ij} x_j, k_i^{-1} \right), \quad (2.24)$$

donde β_{ij} y k_i son parámetros que satisfacen $\beta_{ij}k_i = \beta_{ji}k_j$ para todo i y todo j . En [2, pág 399] se puede observar que

$$p(X) = (2\pi)^{-n/2} |Q|^{1/2} \exp \left(-\frac{1}{2} X^T Q X \right), \quad (2.25)$$

donde

$$[Q]_{ij} = \begin{cases} k_i, & \text{si } i = j, \\ -k_i \beta_{ij}, & \text{si } i \neq j. \end{cases} \quad (2.26)$$

En la Figura 2.6 se observa un ejemplo de (2.25) de una normal multivariada con 10 trayectorias discretizando el intervalo $[0,1]$ en 60 nodos, para la construcción de Q , y en la Figura 2.7 se presenta una muestra para la discretización de $[0,1] \times [0,1]$, el el que se considera obtener muestras de tamaño 50×50 , donde cada pixel de la grilla es un parámetro de la muestra.

En la construcción de la apriori para el problema inverso será de gran importancia (2.25), pues la apriori que se presentará sera una CAMG con ciertas propiedades para la matriz de covarianza Q adaptándose a los requerimientos del método para resolver EDP's con elementos finitos.

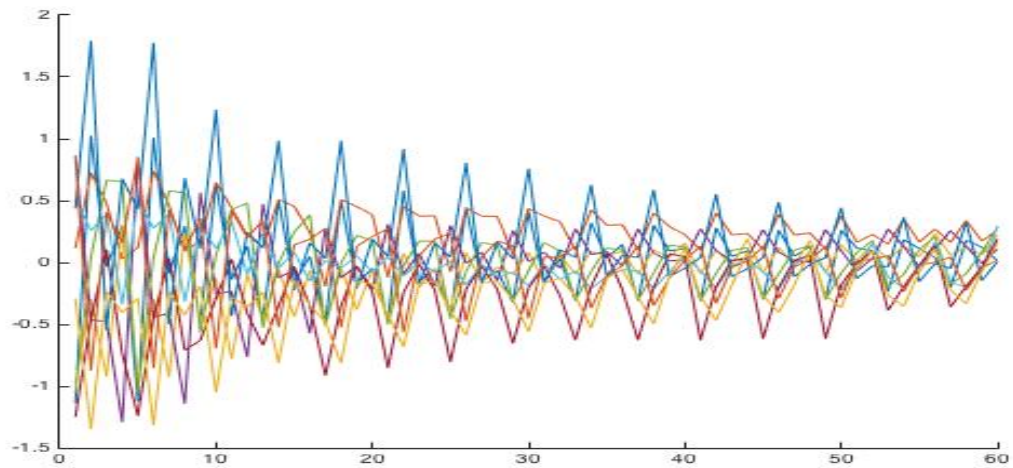


FIGURA 2.6. Diez trayectorias de una normal multivariada en 1D, tomando solo las medias y varianzas asociadas de cada trayectoria se muestra el camino que genera a través de las iteraciones centrándose en 0 (punto más probable).

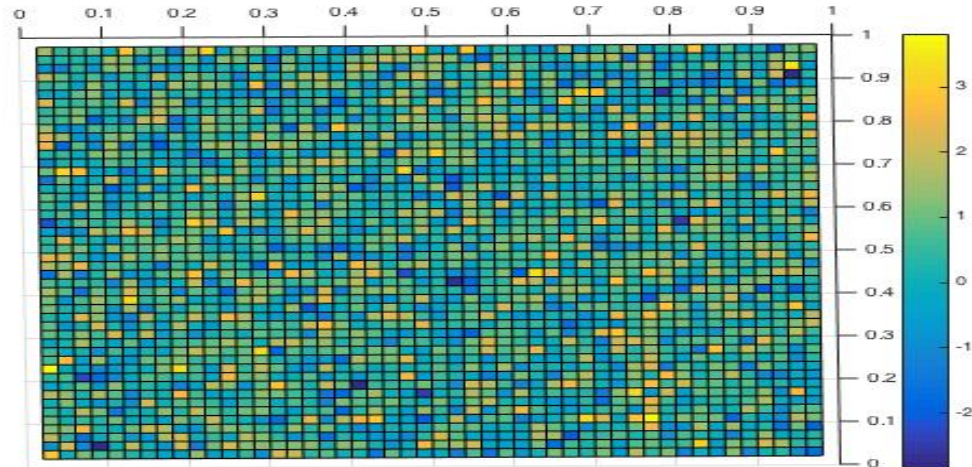


FIGURA 2.7. Una muestra en 2D de una grilla normal multivariada de 100x100, es decir cada pixel posee una distribución normal.

Ésta distribución propuesta en [2] será de gran ayuda como una propuesta apriori en el problema de impedancia tomográfica, la cual tiene como parámetros a recuperar, valores de la grilla en un dominio como en la Figura 2.7.

Introducción a los elementos finitos

En este capítulo se presentará una breve introducción al método de los elementos finitos con base en [7], éste usa el cálculo de variaciones y busca minimizar un funcional de la forma,

$$I(y) = \int_a^b F \left(x, y, \frac{dy}{dx} \right) dx. \quad (3.1)$$

Un funcional puede ser visto como una función de funciones, entonces cuando se hace referencia a minimizar el funcional se busca $y(x)$ que minimice $I(y)$. Como se puede observar en el siguiente ejemplo:

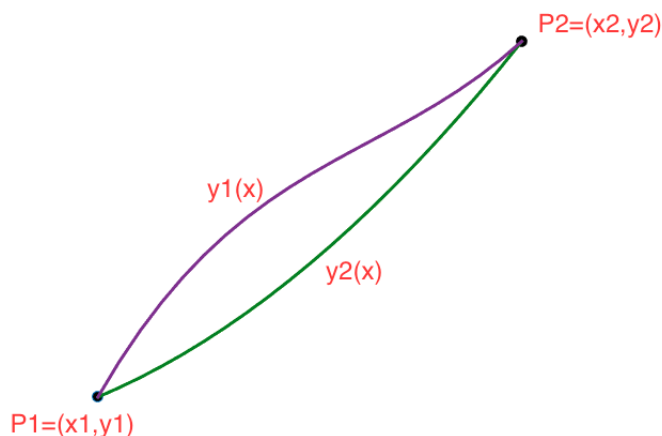


FIGURA 3.1. Funciones que pasan por dos puntos.

Ejemplo 3.0.1. Con respecto a la Figura 3.1 se busca minimizar la distancia entre los puntos P_1 y P_2 . Para ello se debe encontrar una función que minimice esas distancias de entre un conjunto de funciones respectivas. Así el funcional será,

$$I(y) = \int_{x_1}^{x_2} \sqrt{(dx)^2 + (dy)^2} = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx} \right)^2} dx.$$

Como en el cálculo usual se debe igualar la derivada a cero, además para una trayectoria óptima se debe satisfacer la ecuación de Euler-Lagrange que viene dada por:

$$\frac{d}{dx} \left[\frac{\partial}{\partial y'} F(x, y, y') \right] - \frac{\partial}{\partial y} F(x, y, y') = 0. \quad (3.2)$$

Así, para $F(x, y, y') = \sqrt{1 + (y')^2}$, se tiene:

$$\frac{\partial F}{\partial y} = 0, \quad y \quad \frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}}.$$

Aplicando la ecuación de Euler-Lagrange,

$$\frac{d}{dx} \left[\frac{\partial F}{\partial y'} \right] = \frac{d}{dx} \left[\frac{y'}{\sqrt{1 + (y')^2}} \right] = 0 = \frac{\partial F}{\partial y}.$$

De lo anterior se concluye que $\frac{y'}{1+(y')^2} = c$, donde c es una constante. Así,

$$\begin{aligned} y' &= c\sqrt{1 + (y')^2} \\ (y')^2 &= c^2 (1 + (y')^2)^2 \\ y' &= \sqrt{\frac{c^2}{1 - c^2}} = m \quad \text{com } b \text{ una constante.} \end{aligned}$$

De lo anterior se obtiene que $\frac{dx}{dy} = m$, e integrando en ambas partes $y = mx + b$. Dado que y debe pasar por los puntos $P_1 = (x_1, y_1)$ y $P_2 = (x_2, y_2)$ se obtendrá que la función $y(x)$ que minimiza el funcional es:

$$y(x) = \frac{y_2 - y_1}{x_2 - x_1} x + \frac{x_2 y_1 - x_1 y_2}{x_2 - x_1}.$$

El anterior ejemplo es un caso simple e introductorio que permite entender a grandes rasgos el concepto de funcional.

3.1. El método de Galerkin

El método de Galerkin es usado ampliamente en la teoría de elementos finitos, donde se define un residuo $R(x)$ como la diferencia entre el funcional con la ecuación diferencial, por ejemplo, si $y'' + Q(x)y = F(x)$ el residuo asociado $R(x)$ se define como la diferencia $R(x) = y'' + Qy - F$. El residuo $R(x)$ se multiplica por funciones de ponderación v_i o funciones de prueba, donde las v_i son linealmente independientes.

Una vez seleccionadas las v_i , los coeficientes desconocidos se calculan al igualar a cero la integral definida sobre un intervalo $[a, b]$ de residuo ponderado:

$$\int_a^b R(x)v_i(x) dx = 0, \quad i = 0, 1, \dots, n. \quad (3.3)$$

Para observar mejor el método de Galerkin se presentará el siguiente ejemplo:

Ejemplo 3.1.1. (Tomado de [7]) Resuelva

$$y'' + y = 3x^2, \quad y(0) = 0, \quad y(2) = 3.5.$$

Considere el espacio de los polinomios de grado 3 ($\mathbb{P}_3(x)$) y una aproximación $u(x)$ de $y(x)$ tal que,

$$u(x) = \frac{7}{4}x + c_1(x)(x-2) + c_2(x^2)(x-2),$$

de modo que $v_1(x) = x$, $v_2(x) = (x)(x-2)$ y $v_3 = (x^2)(x-2) \in \mathbb{P}_3$ y son linealmente independientes. Se tiene el residuo $R(x) = y'' + y - 3x^2$, que al sustituir y'' por u'' e y por u se obtiene como resultado:

$$R(x) = c_3(x^3 - 2x^2 + 6x - 4) + c_2(x^2 - 2x + 2) - \left(3x^2 - \frac{7}{4}x\right).$$

Así se obtiene el siguiente planteamiento:

- $\int_0^2 v_2 R(x) = 0 \rightarrow \int_0^2 (x^2 - 2x)R(x) = -\frac{8}{4}c_3 - \frac{8}{5}c_2 + \frac{37}{15} = 0.$
- $\int_0^2 v_3 R(x) = 0 \rightarrow \int_0^2 (x^3 - 2x^2)R(x) = -\frac{64}{21}c_3 - \frac{8}{5}c_2 + \frac{18}{5} = 0.$

Lo anterior nos da origen a un sistema en forma matricial que se presenta a continuación:

$$\begin{bmatrix} \frac{8}{5} & \frac{8}{5} \\ \frac{64}{21} & \frac{8}{5} \end{bmatrix} \begin{bmatrix} c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} \frac{37}{15} \\ \frac{18}{5} \end{bmatrix}.$$

Solucionando el anterior sistema se obtiene que $c_2 = \frac{593}{1312}$ y $c_3 = \frac{496}{591}$, sustituyendo en $u(x)$ se obtiene como aproximación a la función $y(x)$ la función:

$$u(x) = \frac{496}{591}x^3 - \frac{2025}{1651}x^2 + \frac{555}{656}x.$$

3.2. Problema de la temperatura de una barra

Sea $k(x)$ una función que representa la conductividad de la barra y $u(x)$ una función que representa la temperatura en $x \in [a, b]$, se desea calcular u . Así la formulación del problema se presenta como sigue:

1 Forma fuerte:

$$\begin{cases} \text{Encontrar } u : [a, b] \rightarrow \mathbb{R} \text{ tal que} \\ -(k(x)u'(x))' = f(x) \quad \text{para todo } x \in (a, b) \\ u(a) = 0, u(b) = 0. \end{cases} \quad (3.4)$$

El anterior problema cuenta con las condiciones de frontera de Dirichlet $u(a) = u(b) = 0$.

2 Forma débil (o formulación variacional)

- Se multiplica por una función de prueba $v \in C_0^\infty(a, b)$,
- Integración por partes

Se integra (3.4) en el intervalo (a, b) ,

$$-\int_a^b (k(x)u'(x))' v(x) dx = \int_a^b f(x)v(x) dx.$$

Aplicando la integración por partes en la parte izquierda de la ecuación anterior se obtiene la siguiente igualdad:

$$\int_a^b k(x)u'(x)v'(x) dx - [k(a)u'(a)v(a) - k(b)u'(b)v(b)] = \int_a^b f(x)v(x) dx.$$

Aplicando la condición de contorno dada en (3.4) se obtiene la siguiente formulación:

$$\begin{cases} \text{Encontrar } u \text{ tal que:} \\ \int_a^b k(x)u'(x)v'(x) dx = \int_a^b f(x)v(x) dx \text{ para todo } v \in C_0^\infty(a, b) \\ u(a) = 0, u(b) = 0. \end{cases} \quad (3.5)$$

En la anterior formulación u es tal que todas las integrales tienen sentido. Ésta no es propiamente la formulación débil del problema. Sin embargo este ejemplo permite mostrar como se inicia la formulación del problema débil (presentada más adelante) de una EDP.

Note que si u resuelve (3.4) entonces también resuelve (3.5), pero si u resuelve (3.5) no necesariamente resuelve (3.4) (puede que no exista la segunda derivada).

Observaciones sobre linealidad de (3.5)

1. $\int_a^b k u' v_1' dx = \int_a^b f v_1 dx.$
2. $\int_a^b k u' v_2' dx = \int_a^b f v_2 dx.$

Multiplicando por escalares α_1, α_2 y sumando se obtiene:

$$\int_a^b k u' (\alpha_1 v_1 + \alpha_2 v_2)' dx = \int_a^b f (\alpha_1 v_1 + \alpha_2 v_2) dx.$$

Si valen las condiciones de linealidad (1) y (2) entonces la anterior igualdad vale para cualquier combinación lineal de v_1 y v_2 .

3.2.1. Formulación de Galerkin

Idea: En (3.5) busque u y use v en espacios de dimensión finita.

Escoja una base $\{\varphi_1, \varphi_2, \dots, \varphi_N\}$ de N funciones linealmente independientes y genere el espacio V , donde,

$$V = \{\varphi_i\}_{i=1}^N = \{\alpha_1 \varphi_1 + \alpha_2 \varphi_2 + \dots + \alpha_N \varphi_N \mid \alpha_1, \alpha_2, \dots, \alpha_N \in \mathbb{R}\}.$$

Así obtenemos la formulación de Galerkin,

$$\begin{cases} \text{Encontrar } u \text{ tal que:} \\ \int_a^b k(x)u'(x)v'(x) dx = \int_a^b f(x)v(x) dx \text{ para todo } v \in V \\ u(a) = 0, u(b) = 0. \end{cases} \quad (1) \quad (3.6)$$

3.2.2. Formulación matricial

Note que:

- 1 Como $u \in V$, $u = x_1\varphi_1 + x_2\varphi_2 + \cdots + x_N\varphi_N$
- 2 Basta usar $v = \varphi_i$, $i = 1, \dots, N$ como función de prueba.

Usando las condiciones anteriores (1) y (2) en la ecuación (3.9) se tiene:

$$\int_1^b k \underbrace{(x_1\varphi_1 + x_2\varphi_2 + \cdots + x_N\varphi_N)'}_{u'} \underbrace{\varphi_i'}_{v'} = \int_a^b f \underbrace{\varphi_i}_v, \quad i = 1, 2, \dots, N. \quad (3.7)$$

Aplicando propiedades de linealidad de la integral en (3.7) se obtiene la ecuación,

$$x_1 \underbrace{\int_a^b k\varphi_1'\varphi_i'}_{a_{i1}} + x_2 \underbrace{\int_a^b k\varphi_2'\varphi_i'}_{a_{i2}} + \cdots + x_N \underbrace{\int_a^b k\varphi_N'\varphi_i'}_{a_{iN}} = \underbrace{\int_a^b k\varphi_1'\varphi_i}_{b_i} \quad i = 1, 2, \dots, N. \quad (3.8)$$

La ecuación (3.8) puede escribirse como,

$$\sum_{j=1}^N a_{ij}x_j = b_i, \quad i = 1, \dots, N.$$

Si $A = [a_{ij}]$ y $b = [b_1, b_2, \dots, b_N]^T$ entonces se tiene la formulación $Ax = b$, así el problema en formulación matricial queda como sigue,

$$\begin{cases} \text{Encontrar } x_1, \dots, x_N \text{ tal que:} \\ u = \sum_{i=1}^N x_i\varphi_i \text{ y } Ax = b \\ a_{ij} = \int_a^b k\varphi_i'\varphi_j', b_i = \int_a^b f\varphi_i, u(a) = u(b) = 0 \end{cases} \quad (3.9)$$

Ejemplo 3.2.1. Considere la siguiente EDP en formulación fuerte.

$$\begin{cases} u''(x) = -1, & x \in (0, 1), \\ u(0) = 1, & u(1) = 0. \end{cases} \quad (3.10)$$

Aplicando integrales y multiplicando por una función de prueba el problema (3.10) se obtiene una reformulación del problema como sigue,

$$\begin{cases} \text{Encontrar } u : [0, 1] \rightarrow \mathbb{R} \text{ tal que} \\ \int_0^1 u'v' = -\int_0^1 v, \text{ para todo } v \in C_0^\infty(0, 1) \\ u(0) = u(1) = 0. \end{cases}$$

El anterior problema tiene como solución exacta $u(x) = \frac{1}{2}x(1-x)$. Si se toma $N = 1$, $\varphi_1(x) = \sin(\pi x)$ y $V = \{\varphi_i\} = \{\alpha \sin(x) \mid \alpha \in \mathbb{R}\}$. Tomando la matriz $A = [a_{11}]$ y el vector $b = [b_1]$, se obtiene las siguientes igualdades:

$$a_{11} = \int_0^1 \varphi_1'\varphi_1' = \int_0^1 \pi^2 \cos^2(\pi x) = \pi^2 \int_0^1 \cos^2(\pi x) = \pi^2 \left[\frac{2\pi x + \sin(2\pi x)}{4\pi} \right] \Big|_0^1 = \pi^2 \left[\frac{2\pi}{4\pi} \right] = \frac{\pi^2}{2},$$

y

$$b_1 = - \int_0^1 1 \cdot \sin(\pi x) = x \frac{\cos(\pi x)}{\pi} \Big|_0^1 = \frac{1}{\pi} + \frac{1}{\pi} = \frac{2}{\pi}.$$

Por el método de Galerkin se busca $u \in V$ tal que $\int_0^1 u'v' = \int_0^1 1 \cdot v$ para todo $v \in V$, lo anterior implica la condición $u(0) = 0$ y $u(1) = 0$. En la forma matricial se busca encontrar x_1 , tal que, $Ax = b$, o sea, $a_{11}x_1 = b_1$, donde se reemplazan los términos encontrados anteriormente, $\frac{\pi^2}{2}x_1 = \frac{2}{\pi}$, despejando finalmente se obtiene $x_1 = \frac{4}{\pi^3}$. Luego $u(x) = \frac{4}{\pi^3} \sin(\pi x)$ es la solución de Galerkin en el espacio $V = \{\alpha \sin(x) \mid \alpha \in \mathbb{R}\}$.

Nota. Cuando la solución exacta está en el subespacio V , entonces el método de Galerkin da como resultado la solución exacta.

La idea del método de los elementos finitos es usar funciones base con soporte pequeño (es decir diferentes de cero en un conjunto pequeño).

Ejemplo 3.2.2. *Elementos finitos lineales por partes.*

1. Obtenga una triangulación (división del $[a, b]$ en subintervalos o elementos) del intervalo $[a, b]$. Se genera una triangulación formada por los elementos $[x_0, x_1], [x_1, x_2], \dots, [x_N, x_{N+1}]$, donde x_0, x_{N+1} son llamados los vértices frontera y x_i con $i = 1, \dots, N$ son llamados los vértices interiores.
2. Para cada vértice interior x_i , defina la función base $\varphi_i(x)$ como siendo,

$$\varphi_i(x) = \begin{cases} \varphi_i(x_i) = 1, & x = x_i, \\ \varphi_i(x_j) = 0, & x = x_j, \\ \varphi(x) \text{ es lineal si } x \in [x_j, x_j + 1]. \end{cases}$$

3.3. Introducción a los elementos finitos (1D)

Primero se presentará el espacio de elementos finitos *lineales por partes* con respecto a una triangulación (división del dominio en intervalos), para esto se presentarán la metodología que se empleará inicialmente.

- 1) Subdividir el dominio $[a, b]$ sobre el cual está definida la ecuación diferencial ordinaria (EDO) en n subintervalos denominados elementos, que van desde x_1 hasta x_{n-1} , a este arreglo se le unen los puntos $x_0 = a$ y $x_n = b$ tal que $x_j < x_{j+1}$ para $j = 0, 1, \dots, n-1$. Los elementos se numeran del 1 hasta el n donde el elemento (i) va desde x_{i-1} hasta x_i con $i = 1, 2, \dots, n$.
- 2) Aplicar el método de Galerkin por separado en cada elemento para interpolar $u(x)$ que son las aproximaciones de la verdadera solución de la ecuación.
- 3) Cuando se ha aplicado el método de Galerkin a cada elemento se obtienen ecuaciones que involucran todas las constantes desconocidas de $u(x)$, el proceso de combinar las ecuaciones de los elementos por separado se denomina ensambla el sistema.
- 4) El sistema obtenido se ajusta para las condiciones frontera y se resuelve para obtener la aproximación $u(x)$ de $y(x)$.

Con el siguiente ejemplo será más fácil observar la metodología que se emplea.

Ejemplo 3.3.1. Resolver la ecuación de Laplace por elementos finitos en una dimensión, donde la ecuación de Laplace viene dada por:

$$\begin{cases} -u''(x) = -1, & 0 < x < 1 \\ u(0) = 1, u(1) = 1. \end{cases} \quad (3.11)$$

Entonces para el anterior sistema se quiere encontrar $u(x) : [0, 1] \rightarrow \mathbb{R}$ tal que se satisfagan las condiciones de (3.11).

Para la formulación débil de la ecuación, se toma $v \in C_0^\infty[0, 1]$ una función fija pero arbitraria, tal que,

$$\begin{cases} -\int_0^1 u''(x)v(x)dx = -\int_0^1 v(x)dx, & 0 < x < 1 \text{ y } v(x) \in C_0^\infty[0, 1] \\ u(0) = 1, u(1) = 1. \end{cases} \quad (3.12)$$

Considere $\frac{d}{dx}(u'(x)v(x)) = v(x)\frac{du'(x)}{dx} + u'(x)\frac{dv(x)}{dx}$, que al integrar en ambas partes de la igualdad en un intervalo $[a, b]$ se convierte en:

$$\begin{aligned} [u'(b)v(b) - u'(a)v(a)] &= \int_a^b u''(x)v(x)dx + \int_a^b u'(x)v'(x)dx \\ -\int_a^b u''(x)v(x)dx &= \int_a^b u'(x)v'(x)dx - [u'(b)v(b) - u'(a)v(a)]. \end{aligned}$$

Dado que $v(x)$ es cero en la frontera de $[a, b]$, es decir, $v(a) = 0 = v(b)$, así reemplazando la ecuación (3.12) se obtiene:

$$\begin{cases} \int_0^1 u'(x)v'(x)dx = -\int_0^1 v(x)dx, & 0 < x < 1 \text{ y } v(x) \in C_0^\infty[0, 1] \\ u(0) = 1, u(1) = 1. \end{cases} \quad (3.13)$$

Sea $u^*(x) \in H^1([0, 1])$ una aproximación de $u(x)$ definida como $u^*(x) = \sum_{i=1}^n \alpha_i \varphi_i(x)$, donde $\varphi_i \in H_0^1([0, 1])$ y $\alpha_i \in \mathbb{R}$ para $i = 1, 2, \dots, n$. Así, reemplazando en (3.13) se obtiene la siguiente formulación:

$$\begin{cases} \sum_{i=1}^n \int_{x_1}^{x_2} \alpha_i \varphi_i'(x)v_i'(x)dx = -\int_{x_1}^{x_2} v_i(x)dx, & 0 < x < 1 \text{ y } \varphi_i \in H_0^1([x_1, x_2]) \text{ } i = 1, 2, \dots, n \\ u^*(0) = 1, u^*(1) = 1. \end{cases} \quad (3.14)$$

Ahora triangulando el intervalo $[0, 1]$ de manera uniforme en cuatro partes, con puntos $x_0 = 0$, $x_1 = 1/3$, $x_2 = 2/3$ y $x_3 = 1$ y con elementos $K_1 = (0, 1/3)$, $K_2 = (1/3, 2/3)$ y $K_3 = (2/3, 1)$.

Tomando funciones $\varphi_i = \frac{x_2-x}{x_2-x_1}$ y $\varphi_j = \frac{x-x_1}{x_2-x_1}$ para $i = 1, 3, 5$, $j = 2, 4, 6$ en un intervalo (x_1, x_2) , con esto se plantea la formulación de Galerkin para cada elemento, tomando $v_1 = \varphi_1$ y $v_2 = \varphi_2$ en cada K_i , $i = 1, 2, 3$, como sigue:

- Para el elemento K_1 , $\varphi_1(x) = -3x + 1$ y $\varphi_2(x) = 3x$, que al sustituir en (3.14) se obtiene el sistema:

$$\begin{aligned}\alpha_1 \int_0^{1/3} 9dx - \alpha_2 \int_0^{1/3} 9dx &= - \int_0^{1/3} (-3x + 1)dx. \\ -\alpha_1 \int_0^{1/3} 9dx + \alpha_2 \int_0^{1/3} 9dx &= - \int_0^{1/3} 3xdx.\end{aligned}$$

Al solucionar las integrales y planteando las matrices, resulta:

$$A_{K_1} = 3 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad y \quad \mathbf{b}_{K_1} = -\frac{1}{6} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

- Para el elemento K_2 , $\varphi_2(x) = -3x + 2$ y $\varphi_3(x) = 3x - 1$, que al sustituir en (3.14) se obtienen las matrices:

$$A_{K_2} = 3 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad y \quad \mathbf{b}_{K_2} = -\frac{1}{6} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

- Para el elemento K_3 , $\varphi_3(x) = -3x + 3$ y $\varphi_4(x) = 3x - 2$, que al sustituir en (3.14) se obtienen las matrices:

$$A_{K_3} = 3 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad y \quad \mathbf{b}_{K_3} = -\frac{1}{6} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

El sistema a solucionar es $A\alpha = \mathbf{b}$, donde $\alpha = [\alpha_1 = u^*(x_0) \quad \alpha_2 = u^*(x_1) \quad \alpha_3 = u^*(x_2) \quad \alpha_4 = u^*(x_3)]^T$, $A = \sum_{i=1}^3 R_i^T A_{K_i} R_i$ y $\mathbf{b} = \sum_{i=1}^3 R_i^T \mathbf{b}_{K_i}$. Los R_i vienen dados por:

$$R_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$R_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$R_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Con esto se obtiene el sistema:

$$3 \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} u^*(x_0) \\ u^*(x_1) \\ u^*(x_2) \\ u^*(x_3) \end{bmatrix} = -\frac{1}{3} \begin{bmatrix} \frac{1}{2} \\ 1 \\ 1 \\ \frac{1}{2} \end{bmatrix}. \quad (3.15)$$

Utilizando las condiciones de frontera descritas en (3.14)

$$\begin{bmatrix} u^*(x_0) \\ u^*(x_1) \\ u^*(x_2) \\ u^*(x_3) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ u^*(x_1) \\ u^*(x_2) \\ 0 \end{bmatrix}. \quad (3.16)$$

Reemplazando la igualdad (3.16) en (3.15), el sistema se reescribe como:

$$\begin{aligned} R_1 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \\ R_2 &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\ R_3 &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Con esto se obtiene el sistema:

$$3 \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ u^*(x_1) \\ u^*(x_2) \\ 0 \end{bmatrix} = -\frac{1}{6} \begin{bmatrix} 19 \\ -16 \\ -16 \\ 19 \end{bmatrix}.$$

Solucionando el sistema se tiene $\alpha = [u^*(x_0) \ u^*(x_1) \ u^*(x_2) \ u^*(x_3)]^T = [1 \ 8/9 \ 8/9 \ 1]^T$. En este ejemplo la función es la generada al unir los puntos $(0, 1)$, $(1/3, 8/9)$, $(2/3, 8/9)$ y $(1, 1)$.

A continuación se presentará la idea de la implementación de elementos finitos en dos dimensiones.

3.4. Introducción a los elementos finitos (2D)

La aplicación del método de elementos finitos en dos dimensiones es muy similar al de una dimensión, una de las diferencias está en que ya no se empleará la integración por partes presentada en la sección anterior para la formulación del método de Galerkin. En vez de la integración por partes se hará uso de la primera identidad de Green presentada a continuación, que se deriva del teorema de la divergencia aplicado a un campo vectorial $\mathbf{F} = \psi \nabla \varphi$ definido sobre un espacio Ω .

$$\int_{\Omega} \psi \Delta \varphi dV = \oint_{\partial \Omega} \psi (\nabla \varphi \cdot \mathbf{n}) dS - \int_{\Omega} (\nabla \varphi \cdot \nabla \psi) dV, \quad (3.17)$$

donde $\varphi \in C^2$, $\psi \in C^1$ en Ω funciones escalares y $\Delta = \nabla^2$ (Laplaciano).

La anterior igualdad se obtiene de la identidad $\nabla \cdot \mathbf{F} = \nabla \cdot (\psi \nabla \varphi) = \nabla \psi \cdot \nabla \varphi + \psi \Delta \varphi$.

Otra diferencia será la triangulación que se realiza sobre el dominio que en este caso es sobre un área, ésta triangulación puede ser para diferentes elementos, entre los cuales pueden ser triángulos o cuadrados, como se presenta en las Figuras 3.2 y 3.3 respectivamente. Como antes se pueden usar diferentes funciones base, en particular se usará los espacios de los elementos finitos lineal por partes con respecto a una triangulación, ver [11].

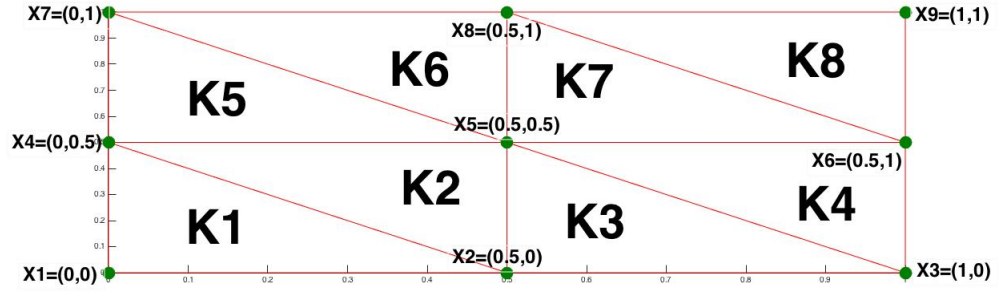


FIGURA 3.2. Triangulación del dominio $[0,1] \times [0,1]$ en triángulos.

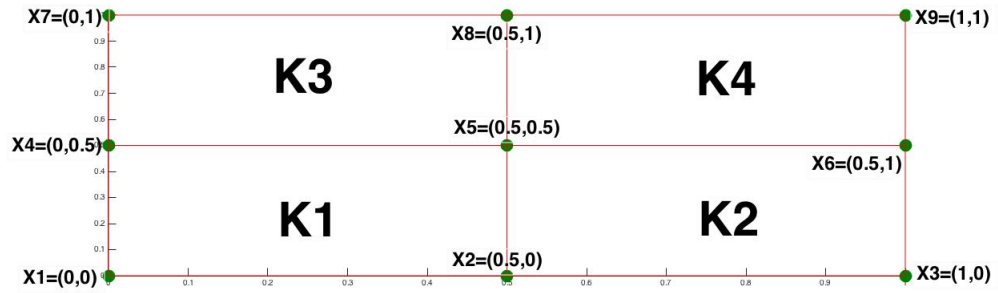


FIGURA 3.3. Triangulación del dominio $[0,1] \times [0,1]$ en cuadrados.

Problema de impedancia tomográfica

En este capítulo se presentará el problema de impedancia de tomográfica, presentando el problema directo con la formulación fuerte y débil de la EDP que es generada por el problema físico, al igual que la metodología para tratar el problema inverso.

La fuente del problema inverso en la impedancia tomográfica viene sustentada en las ecuaciones de la ley de Ohm y las leyes de Kirchhoff que son leyes físicas presentes en los cuerpos. Para esto se darán algunas definiciones de la física que permitirán esclarecer el concepto de las anteriores leyes mencionadas, ver [22].

Definición 4.0.1. *Se presentan las siguientes definiciones de términos que serán utilizados posteriormente;*

1. **Campo eléctrico** (\vec{E}). *Es un campo vectorial con propiedades físicas, tales como la fuerza, que son generadas por una carga en el espacio. Se representa la magnitud del campo con la letra E .*
2. **Campo magnético.** *Es un campo vectorial que con por lo menos dos fuentes de carga a una distancia máxima que lo originan (la distancia máxima depende de la carga de las fuentes). Éste campo depende de dos fuentes que la originan, una de ella es el campo eléctrico y la otra es una corriente de desplazamiento, si ambas no varían durante el tiempo generan un campo magnético estático. Se verá que el nombre de campo magnético se aplica a dos magnitudes, las cuales son inducción magnética (\vec{B}) o simplemente campo magnético y a la excitación magnética (\vec{H}) o simplemente campo \vec{H} .*
3. **Excitación magnética** (\vec{H}). *También conocida como campo \vec{H} , es un campo magnetizante equivalente a un campo de inducción magnética salvo el producto con una función. Se toma frecuentemente relacionada con solenoides (tubos conductores de campos magnético en su interior) o cargas libres (que se mueven en el medio).*
4. **Inducción magnética** (\vec{B}). *En la actualidad la inducción magnética es simplemente llamada campo magnético y está dada por $\vec{B} = \mu\vec{H}$, donde μ es la permeabilidad magnética del material que es la capacidad de una sustancia o medio para atraer y hacer pasar a través de ella campos magnéticos.*

5. **Densidad de corriente eléctrica** (\vec{J}). Es la intensidad por unidad de área o “cantidad” de corriente que pasa a través de un área, su magnitud se denota por J .
6. **Conductividad eléctrica** (γ). Es la capacidad de un material o sustancia para dejar pasar la corriente a través de él. Su inverso multiplicativo es conocido como **resistividad eléctrica** (ρ) que como su nombre lo indica es la resistencia que presenta un cuerpo o sustancia a la energía que atraviesa a través de él.

Estas definiciones brevemente explicadas serán de ayuda para comprender la terminología utilizada, y será de gran importancia para comprender las variables que se presentan en las formulas a continuación que serán la base física de la comprensión del problema de impedancia tomográfica.

1. **Ley de Ohm.** Ésta ley relaciona la intensidad de corriente eléctrica o simplemente corriente (I) que es el flujo de carga eléctrica que recorre un material, con el voltaje (V) o diferencia de potencial eléctrico (es el trabajo que debe realizar un campo electrostático para mover una carga positiva desde dicho punto hasta el punto de referencia) entre dos puntos con la impedancia (Z) que es la oposición que presenta un circuito a una corriente cuando se aplica una tensión (voltaje), ésta ley está presentada como sigue,

$$V = ZI. \quad (4.1)$$

La ley de Ohm en un campo eléctrico viene dada por una relación entre la densidad de corriente \vec{J} , el campo eléctrico \vec{E} y la conductividad γ , dada por:

$$\vec{J} = \gamma \vec{E}. \quad (4.2)$$

2. **Leyes de Kirchhoff.** Estas leyes son también llamadas leyes de conservación de la energía y la carga en circuitos eléctricos.

- a) En cualquier nodo, la suma de las corrientes que entran en ese nodo es igual a la suma de las corrientes que salen. De forma equivalente, la suma de todas las corrientes que pasan por el nodo es igual a cero:

$$\sum_{k=1}^n I_k = I_1 + I_2 + I_3 + \cdots + I_n = 0. \quad (4.3)$$

- b) En un lazo cerrado, la suma de todas las caídas de tensión es igual a la tensión total suministrada. De forma equivalente, la suma algebraica de las diferencias de potencial eléctrico en un lazo es igual a cero:

$$\sum_{k=1}^n V_k = V_1 + V_2 + V_3 + \cdots + V_n = 0. \quad (4.4)$$

3. **Ley de Ampere-Maxwell.** Ésta ley expresa que un campo magnético tiene fuentes vectoriales, por lo tanto el campo magnético no deriva de un potencial escalar (función sobre \mathbb{R}).

$$\nabla \times \vec{H} = \vec{J} + \frac{\partial D}{\partial t}. \quad (4.5)$$

Dicha ecuación en un campo magnético estático se puede escribir como:

$$\nabla \times \vec{H} = \vec{J}. \quad (4.6)$$

Las anteriores ecuaciones serán de utilidad para poder deducir una ecuación matemática que contenga la comprensión física del problema de impedancia tomográfica. Para ello se representa el campo eléctrico \vec{E} en términos del potencial eléctrico escalar u , siendo de la forma:

$$\vec{E}(x) = -\nabla u(x) \text{ con } x \in \Omega. \quad (4.7)$$

Al sustituir (4.7) en la ley de Ohm (4.2), se tiene:

$$\vec{J} = \gamma \vec{E} = -\gamma \nabla u(x) \text{ con } x \in \Omega. \quad (4.8)$$

Sustituyendo el termino obtenido en (4.8) en la Ley de Ampere-Maxwell para un campo magnético estático (4.6),

$$\nabla \times \vec{H} = -\gamma \nabla u(x) \text{ con } x \in \Omega. \quad (4.9)$$

Aplicando divergencia en ambos lados de la igualdad dada en (4.9), se obtendrá un término igualado a cero, puesto que la divergencia del rotacional de un campo vectorial es igual a cero.

$$\underbrace{\nabla \cdot (\nabla \times \vec{H})}_0 = \nabla \cdot (-\gamma \nabla u(x)) \text{ con } x \in \Omega. \quad (4.10)$$

Así se tiene por (4.10) que el potencial eléctrico u en el cuerpo Ω con una función γ satisface la ecuación,

$$\nabla \cdot (\gamma(x) \nabla u(x)) = 0. \quad (4.11)$$

La anterior ecuación (4.11) se puede solucionar mediante el método de elementos finitos, donde los parámetros de entrada dependen de la solución de regularización obtenida de Tikhonov para distribuciones propuestas sobre los parámetros dados para así recuperar la imagen de la tomografía.

Una distribución de conductividad $\gamma : \Omega \rightarrow \mathbb{C}$ está en la clase admisible de conductividades, denotada por $A = A(\Omega)$, si satisface las siguientes condiciones:

- Para algún $N \geq 0$, hay una familia $\{\Omega_j\}_{j=1}^N$ de conjuntos disyuntos, $\Omega_j \subset \Omega$, tiene una frontera suave por partes y para la cual

$$\bar{\Omega} = \bigcup_{j=1}^N \bar{\Omega}_j.$$

Además, es necesario que $\gamma|_{\Omega_j} \in C(\bar{\Omega}_j)$, $1 \leq j \leq N$, es decir γ restringido a cada subconjunto Ω_j admite una extensión continua a la frontera del conjunto.

- Para algunas constantes c y C , se tiene

$$0 < c \leq \gamma(x) \leq C < \infty.$$

En la Figura 4.1 se ilustra la situación.

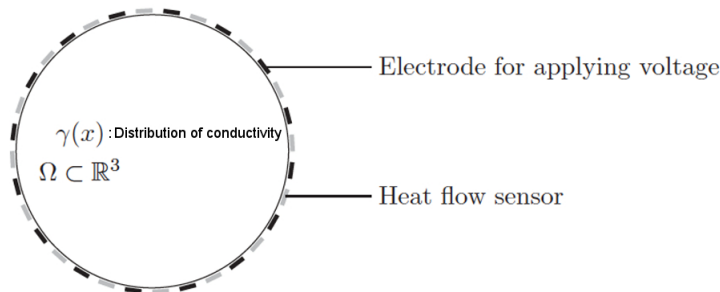


FIGURA 4.1. Esquema del problema de impedancia tomográfica. (Tomado y modificado de “<https://www.aip2015.fips.fi/intro-part5.png>” from Katya Krupchyk)

En aplicaciones medicas, los subconjuntos Ω_j en el problema directo pueden representar los órganos. En el problema inverso, el conjunto de conductividades admisibles proporciona una discretización natural de la base. Debido a posibles discontinuidades de $\gamma \in A$, la ecuación (4.11) debe ser interpretada en el sentido débil (forma alternativa en que dichas ecuaciones se escriben en forma integral, dando lugar a ecuaciones tratables mediante los métodos del álgebra lineal sobre un espacio vectorial).

Se nota $\mathbb{H} = H^1(\Omega) \oplus \mathbb{R}^L$ donde $H^1(\Omega)$ [9] es una base L^2 del espacio de Sóbolev. También se nota $\mathbb{K} = \mathbb{H}/\mathbb{R}$.

Proposición. Sea $\gamma \in A(\Omega)$. El problema (4.11) tiene una solución débil única $(u, U) \in \mathbb{K}$ en el siguiente sentido. Existe una única $(u, U) \in \mathbb{K}$ que satisface la ecuación

$$B_{\gamma,z}((u, U), (v, V)) = \sum_{l=1}^L I_l V_l \quad (4.12)$$

para todo $(v, V) \in \mathbb{K}$, donde la forma cuadrática $B_{\gamma,z}$ está dada por,

$$B_{\gamma,z}((u, U), (v, V)) = \int_{\Omega} \gamma \nabla u \cdot \nabla v dx + \sum_{l=1}^L \frac{1}{Z_l} \int_{e_l} (u - U_l)(v - V_l) dS. \quad (4.13)$$

Además, la forma cuadrática es coercitiva en \mathbb{K} , es decir, se satisface la desigualdad

$$\alpha_0 \|(u, U)\|_{\mathbb{K}}^2 \leq B_{\gamma,z}((u, U), (v, V)) \leq \alpha_1 \|(u, U)\|_{\mathbb{K}}^2 \quad (4.14)$$

para algunas constantes $0 < \alpha_0 \leq \alpha_1 < \infty$. Donde $\|(u, U)\|_{\mathbb{K}} = \inf_{c \in \mathbb{R}} \|(u - c, U - c)\|_{\mathbb{K}}$.

A partir de los datos de corriente y voltaje se quiere estimar γ y de aquí se estima la impedancia, esto se puede interpretar como la imagen del interior de omega, para mayores detalles, la anterior proposición fue tomada de [17].

4.1. Problema directo de impedancia tomográfica en formulación fuerte

El problema directo de impedancia tomográfica hace referencia a que dada la conductividad, para un conjunto de corrientes aplicadas se puede encontrar el conjunto de voltajes

asociado. Se fija un dominio Ω y electrodos $e_l \subseteq \partial\Omega$, $l = 1, 2, \dots, L$, $I \in \mathbb{R}^L$ representa los valores promedio de corriente y $U \in \mathbb{R}^L$ representa valores de voltaje.

Dado γ , $z = (z_l)$ para el problema directo, se construye un operador lineal $T : \mathbb{R}^L \rightarrow \mathbb{R}^L$.

1 Note que $T = T(\gamma, z)$, es decir, T depende de z y γ .

2 T transforma corrientes $I \in \mathbb{R}^L$ en voltajes $U = TI$.

3 Además se quiere que si $\sum_{l=1}^L I_l = 0$ entonces $\sum_{l=1}^L U_l = 0$ por las leyes de Kirchhoff (4.4) y (4.3).

A continuación se describe el operador T .

Dado $I \in \mathbb{R}^L$, $U = TI$, si existe una función $u \in H^1(\Omega)$ tal que,

$$\begin{cases} -\operatorname{div}(\gamma \nabla u) = 0 & \text{en } \Omega. \\ \int_{e_l} \gamma \frac{\partial u}{\partial \eta} = I_l, & l = 1, 2, \dots, L. \quad (1) \\ \gamma \frac{\partial u}{\partial \eta} = 0 & \text{en } \partial\Omega \setminus \cup e_l. \quad (2) \\ \left(u + z_l \gamma \frac{\partial u}{\partial \eta} \right) |_{e_l} = U_l, & l = 1, 2, \dots, L. \quad (3) \end{cases} \quad (4.15)$$

Ahora se escribe la formulación débil de la ecuación diferencial,

$$-\operatorname{div}(\gamma \nabla u) = 0.$$

Multiplicando por $v \in H^1$, integrando y aplicando integración por partes, se tiene:

$$\begin{aligned} - \int_{\Omega} \operatorname{div}(\gamma \nabla u) v &= \int_{\Omega} 0v \\ \int_{\Omega} \gamma \nabla u \nabla v - \int_{\partial\Omega} \gamma \frac{\partial u}{\partial \eta} v &= 0 \end{aligned}$$

Por la condición (2) de la ecuación (4.15) se tiene que $\int_{\partial\Omega} \gamma \frac{\partial u}{\partial \eta} v = \sum_{l=1}^L \int_{e_l} \gamma \frac{\partial u}{\partial \eta} v$, con lo que se obtiene que,

$$\int_{\Omega} \gamma \nabla u \nabla v - \sum_{l=1}^L \int_{e_l} \gamma \frac{\partial u}{\partial \eta} v = 0.$$

Haciendo uso de la condición (1) de la ecuación (4.15) se toma $\sum_{l=1}^L \left(\int_{e_l} \gamma \frac{\partial u}{\partial \eta} - I_l \right) V_l = 0$, para agregar este término a la ecuación anterior obteniendo como resultado,

$$\begin{aligned} \int_{\Omega} \gamma \nabla u \nabla v - \sum_{l=1}^L \int_{e_l} \gamma \frac{\partial u}{\partial \eta} v + \sum_{l=1}^L \left(\int_{e_l} \gamma \frac{\partial u}{\partial \eta} - I_l \right) V_l &= 0 \\ \int_{\Omega} \gamma \nabla u \nabla v - \sum_{l=1}^L \int_{e_l} \gamma \frac{\partial u}{\partial \eta} (v - V_l) &= \sum_{l=1}^L I_l V_l. \end{aligned}$$

De la condición (1) de la ecuación (4.15) tomando $-\gamma \frac{\partial u}{\partial \eta} = \frac{1}{z_l} (u - U_l)$ finalmente se obtiene,

$$\int_{\Omega} \gamma \nabla u \nabla v + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u - U_l) (v - V_l) = \sum_{l=1}^L I_l V_l. \quad (4.16)$$

Donde se obtiene la formulación débil del operador T como se sigue. Dado $I \in \mathbb{R}^L$ con $\sum_{l=1}^N I_l = 0$, se tiene que $U = TI$ si vale,

$$\begin{cases} \text{Existe } u \in H^1(D) \text{ con } \int_{\Omega} u = 0 \text{ tal que} \\ \int_{\Omega} \gamma \nabla u \nabla v + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u - U_l)(v - V_l) = \sum_{l=1}^L I_l V_l \\ \text{para todo } v \in H^1(D) \text{ y todo } v \in \mathbb{R}^L. \end{cases}$$

Ahora se presentará la formulación matricial del problema directo.

4.1.1. Formulación matricial del problema directo

Dado γ ($\gamma(x) \geq \gamma_0 > 0$), $z \in \mathbb{R}^L$, $I \in \mathbb{R}^{L-1}$ encontrar $u \in H^1(\Omega)$ y $U \in \mathbb{R}^L$ tal que,

$$\int_{\Omega} \gamma \nabla u \nabla v + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} (u - U_l)(v - V_l) = \sum_{l=1}^L I_l V_l. \text{ para toda } v \in H^1(\Omega) \text{ y } V \in \mathbb{R}^L.$$

Método de Galerkin: Buscar $u \in \{\varphi_i\}_{i=1}^N$, (espacio generado por φ_i con $i=1, \dots, N$.) $U \in \mathbb{R}^L$. Así se obtiene que $u = \sum_{j=1}^N \alpha_j \varphi_j$ y $U = [U_1, \dots, U_N]^T$. De lo anterior se tiene:

a) Si $v = \varphi_i$, para $i = 1, 2, \dots, N$ y $V = 0$. Se tiene, reemplazando en la ecuación (4.16),

$$\int_{\Omega} \gamma \nabla \left(\sum_{j=1}^N \alpha_j \varphi_j \right) \nabla \varphi_i + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} \left(\sum_{j=1}^N \alpha_j \varphi_j - U_l \right) (\varphi_i - 0) = \sum_{l=1}^L (0) I_l, \text{ para } i = 1, \dots, N.$$

Reagrupando los términos anteriores,

$$\sum_{j=1}^N \left(\underbrace{\int_{\Omega} \gamma \nabla \varphi_j \nabla \varphi_i}_{a_{ij}} \right) \alpha_j + \sum_{j=1}^N \left(\underbrace{\sum_{l=1}^L \frac{1}{z_l} \int_{e_l} \varphi_j \varphi_i}_{m_{ij}} \right) \alpha_j - \sum_{l=1}^L \frac{1}{z_l} U_l \underbrace{\int_{e_l} \varphi_i}_{c_{il}}$$

Así se obtiene la siguiente ecuación,

$$\sum_{j=1}^N (a_{ij} + m_{ij}) \alpha_j - \sum_{l=1}^L c_{il} U_l = 0, \quad i = 1, 2, \dots, N. \quad (4.17)$$

b) Si $v = 0$ y $V = E^{(k)} = [0, \dots, 0, 1, 0, \dots, 0]^T$ donde el 1 se encuentra en la k -ésima posición (es decir $E_l^{(k)} = \delta_{kl}$). Se tiene, reemplazando en la ecuación (4.16),

$$\underbrace{\int_{\Omega} \gamma \nabla \left(\sum_{j=1}^N \alpha_j \varphi_j \right) \nabla 0}_0 + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} \left(\sum_{j=1}^N \alpha_j \varphi_j - U_l \right) (0 - E_l^{(k)}) = \sum_{l=1}^L E_l^{(k)} I_l, \text{ para } i = 1, \dots, N.$$

Reagrupando los términos se obtiene la siguiente ecuación,

$$-\sum_{j=1}^L \left(\underbrace{\sum_{l=1}^L \frac{1}{z_l} \int_{e_l} \varphi_j \varphi_l E_l^{(k)}}_{\frac{1}{z_k} \int_{e_k} \varphi_j} \right) \alpha_j + \underbrace{\sum_{l=1}^L \frac{1}{z_l} \int_{e_l} U_l E_l^{(k)}}_{\left(\frac{1}{z_k} \int_{e_k} 1\right) U_k} = \underbrace{\sum_{l=1}^L E_l^{(k)} I_l}_{I_k}.$$

Así,

$$\left(\underbrace{\frac{1}{z_k} \int_{e_k} \varphi_j}_{b_{kj}} \right) \alpha_j + \underbrace{\left(\frac{1}{z_k} \int_{e_k} 1 \right)}_{d_k} U_k = I_k, \quad k = 1, 2, \dots, L.$$

Obteniendo finalmente la siguiente ecuación,

$$-\sum_{j=1}^N b_{kj} \alpha_j + d_k U_k = I_k, \quad k = 1, 2, \dots, L. \quad (4.18)$$

Agrupando las ecuaciones (4.17) y (4.18) en notación matricial los anteriores numerales de la siguiente forma:

$$\begin{bmatrix} A + M & -C \\ -B & D \end{bmatrix} \begin{bmatrix} \alpha \\ U \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix},$$

donde para $i \in \{1, 2, \dots, N\}$ y $l \in \{1, 2, \dots, L\}$ se tiene,

- $A = [a_{ij}]$, $a_{ij} = \int_{\Omega} \gamma \nabla \varphi_i \nabla \varphi_j$ (Note que $A = A(\gamma)$).
- $M = [m_{ij}]$, $m_{ij} = \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} \varphi_j \varphi_l$.
- $C = [c_{ij}]$, $c_{ij} = \frac{1}{z_l} \int_{e_l} \varphi_i$.
- $B = [b_{kj}]$, $b_{kj} = \frac{1}{z_k} \int_{e_k} \varphi_j = c_{kj}$.
- $D = \text{diag}\{d_{kk}\}_{k=1}^L$, $d_{kk} = \frac{1}{z_k} \int_{e_k} 1$.
- $\alpha = [\alpha_1, \dots, \alpha_N]^T$ y $u = \sum_{j=1}^N \alpha_j \varphi_j$.

Note que $B = C^T$, entonces el sistema matricial queda la la forma:

$$\begin{bmatrix} A + M & -B^T \\ -B & D \end{bmatrix} \begin{bmatrix} \alpha \\ U \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix},$$

donde A y M son matrices de tamaño $N \times N$, B es una matriz de tamaño $L \times N$ y D una matriz de tamaño $L \times L$. Recuerde que $\sum_{l=1}^N I_l = 0$, se quiere $\sum_{l=1}^N U_l = 0$ y $\int u = 0$. Así se plantea el siguiente problema:

1. $\sum_{l=1}^N U_l = \underbrace{[1, 1, \dots, 1]}_{\vec{1}} \cdot U$ (se busca $\vec{1} \cdot u = 0$).

2. $\int_{\Omega} u = \sum_{j=1}^N \alpha_j \int_{\Omega} \varphi_j = \sum_{j=1}^N \alpha_j g_j = g \cdot \alpha$ (se busca $g \cdot \alpha = 0$), donde $g = [g_1, g_2, \dots, g_N]$.

Usando multiplicadores de Lagrange para imponer estas dos condiciones, se obtiene la siguiente forma matricial:

$$\left[\begin{array}{cc|cc} A + M & g_{N \times 1}^T & -B^T & 0_{L \times 1} \\ g_{1 \times N} & 0_{1 \times 1} & 0_{1 \times L} & 0_{1 \times 1} \\ \hline -B & 0_{L \times 1} & D & \vec{1} \\ 0_{1 \times N} & 0_{1 \times 1} & \vec{1} & 0_{1 \times 1} \end{array} \right] \begin{bmatrix} \alpha \\ \lambda \\ U \\ \mu \end{bmatrix} = \begin{bmatrix} 0_{N \times 1} \\ 0_{1 \times 1} \\ I \\ 0_{1 \times 1} \end{bmatrix}.$$

Tomando las siguientes matrices,

$$E = \begin{bmatrix} A + M & g \\ g & 0 \end{bmatrix}, \quad \beta = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \quad \hat{I} = \begin{bmatrix} I \\ 0 \end{bmatrix} \quad y \quad W = \begin{bmatrix} U \\ \mu \end{bmatrix},$$

se obtiene el siguiente sistema,

$$\begin{bmatrix} E & -F^T \\ -F & G \end{bmatrix} \begin{bmatrix} \beta \\ W \end{bmatrix} = \begin{bmatrix} 0 \\ \hat{I} \end{bmatrix}.$$

En ecuaciones el anterior sistema queda:

$$1) \quad E\beta - F^T W = 0, \quad (4.19)$$

$$2) \quad -F\beta + GW = \begin{bmatrix} I \\ 0 \end{bmatrix}. \quad (4.20)$$

Despejando β en (4.19) se tiene $\beta = E^{-1}F^T W$, y sustituyendo este término en (4.20) se obtiene la igualdad $-FE^{-1}F^T W + GW = [I, 0]^T$. Agrupando este término se obtiene

$$\underbrace{(G - FE^{-1}F^T)}_S W = [I, 0]^T.$$

Luego

$$SW = [I, 0]^T,$$

y

$$[U, \mu]^T = W = S^{-1}[I, 0]^T.$$

Note que U es la primera componente de W y S es una matriz $(L + 1) \times (L + 1)$ que depende de γ y z . Esto es,

$$S = S(\gamma, z).$$

El operador T es una submatriz de S^{-1} de tamaño $L \times L$ (menor principal) y transforma I (corriente) en U (voltaje). El operador T es conocido como la matriz de resistencias.

4.2. Problema inverso de impedancia tomográfica con planeamiento bayesiano

El problema inverso de impedancia tomográfica consiste en estimar la distribución de conductividad de un conjunto de mediciones en pares de corriente y voltaje, es decir, dados los patrones de corriente $(I_l)_{l=1}^L$ y las mediciones de los patrones de voltaje aplicados en los electrodos $(U_l)_{l=1}^L$ se desea recuperar la función de conductividad γ . En el problema inverso se asume que los datos son experimentales y tienen un error asociado a una distribución, razón por la cual se escoge el método bayesiano para la formulación y solución del problema.

4.2.1. Construcción de la distribución apriori de γ

El objetivo de ésta sección es asignar una distribución apriori para la conductividad γ . Ésto se realizará teniendo en cuenta una discretización del espacio, tal como se presenta en [2].

Considere una discretización uniforme en el cuadrado $[0, 1] \times [0, 1]$ con n^2 vértices $\{(i, j)\}_{i,j=1}^n$ y localización $\{(s_i, t_i)\}_{i,j=1}^n$ donde $s_k = t_k = kh$ con $h = \frac{1}{n+1}$. Sea ∂_i los nodos $i \neq j$ de la partición para los cuales $[Q]_{ij} \neq 0$ en (2.26), se tiene entonces que:

$$\partial_i = \{j \neq i | \beta_{ij} \neq 0\}.$$

Más aún, si $X_{\partial_i} = \{x_j | j \in \partial_i\}$ se tiene que $\pi(x_i | X_{-i}) = \pi(x_i | X_{\partial_i})$ para todo i , entonces,

$$x_i | X_{\partial_i} \sim \text{Nor} \left(\sum_{j \in \partial_i} \beta_{ij} x_j, k_i^{-1} \right). \quad (4.21)$$

Se asume la condición de frontera de Dirichlet dada por $x_{0,j} = x_{n+1,j} = x_{i,0} = x_{i,n+1} = 0$. Las vecindades del sistema están dadas por:

$$\partial_{i,j} = \{(i-1, j), (i+1, j), (i, j-1), (i, j+1)\}, \quad \text{para } i, j = 1, \dots, n.$$

Donde $n_{i,j} = |\partial_{i,j}| = 4$, suponga que existe $\delta \in \mathbb{R}$ tal que se satisface $k_i = \frac{\delta * n_{ij}}{h^2}$, así (4.21) queda de la forma,

$$x_i | X_{\partial_i} \sim \text{Nor} \left(\bar{x}_{\partial_{i,j}}, \left(\frac{\delta * n_{ij}}{h^2} \right)^{-1} \right), \quad (4.22)$$

donde $\bar{x}_{\partial_{i,j}} = \frac{1}{n_{ij}} \sum_{(r,s) \in \partial_{i,j}} x_{rs}$.

Así $Q = \delta L$ y por (2.26) se obtendrá,

$$L = \frac{1}{h^2} \begin{bmatrix} \tilde{L} & -I & o & \cdots & o \\ -I & \tilde{L} & -I & \ddots & \vdots \\ o & \ddots & \ddots & \ddots & o \\ \vdots & \ddots & -I & \tilde{L} & -I \\ o & \cdots & o & -I & \tilde{L} \end{bmatrix}, \quad (4.23)$$

donde I es la matriz identidad de tamaño $n \times n$, o es la matriz de ceros de tamaño $n \times n$, y

$$\tilde{L} = \begin{bmatrix} 4 & -1 & 0 & \cdots & 0 \\ -1 & 4 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 4 & -1 \\ 0 & \cdots & 0 & -1 & 4 \end{bmatrix}. \quad (4.24)$$

Así para la distribución apriori de γ se asume una distribución normal con media 0 y varianza $Q^{-1} = (\delta L)^{-1}$, note que el parámetro δ es desconocido, así que se asume con una distribución gamma de parámetros (1, 1). Así:

$$\pi_{apr}(\gamma) \propto \exp\left(-\frac{\delta}{2}\gamma^T L\gamma\right), \quad (4.25)$$

$$\pi_{apr}(\delta) \propto \exp(-\delta). \quad (4.26)$$

4.2.2. Distribución posterior de γ

Para la distribución posterior se asume para la verosimilitud que los datos obtenidos son de la forma (2.14), donde el resultado obtenido es 2.16, así adaptando este problema para las aprioris para la conductividad (4.25) y el parámetro δ 5.11, se obtiene una función de densidad posterior para γ y δ dada por:

$$\pi_{pos}(\gamma, \delta | U, I) \propto \exp\left(-\frac{1}{2}\|T(\gamma)I - U\|^2 - \frac{\delta}{2}\gamma^T L\gamma - \delta\right), \quad (4.27)$$

donde $T(\gamma)$ es un operador lineal que transforma corrientes I en voltajes U ($T(\gamma)I = U$).

Resultados y aplicaciones de la metodología bayesiana

En éste capítulo se presentan resultados obtenidos en distintos experimentos aplicando una metodología bayesiana, se muestra como dicha metodología permite obtener información (en algunos casos precisa) de parámetros desconocidos y no siempre fáciles de encontrar con los métodos convencionales. En la sección 5.1 se presenta la aplicación a la recuperación de imágenes, asociado al problema de convolución, en la sección 5.2 se presenta una aplicación a la estimación de parámetros de una ecuación diferencial estocástica utilizando como ejemplo los valores reales de un activo en un periodo de tiempo, y en la sección 5.3 se presenta la aplicación de la metodología bayesiana a la estimación de parámetros una ecuación diferencial parcial, de forma más precisa, la solución del problema inverso de EIT para una impedancia conocida. Los códigos para las aplicaciones acá presentadas se encuentran disponibles para su uso libre en: “<https://drive.google.com/open?id=1Qg3B5fWhI7wTHvYAxVdGJxfk3tj2AiCG>”.

5.1. Aplicación a la recuperación de imágenes

En la metodología de recuperación de imágenes ya se aplico previamente una metodología determinista llamada la regularización de Tikhonov, dicha metodología se observo que no siempre es la más adecuada (como se observa en la Figura 1.4).

Para recuperar una imagen borrada se plantea que se tome cada pixel de la imagen como un parámetro a recuperar, suponiendo que se tiene un problema de la forma:

$$Y = AX + E. \tag{5.1}$$

Donde A es una matriz de convolución, X es la imagen real que se desea recuperar a partir de Y que es la imagen que se tiene, donde Y posee un ruido aditivo E . Se asume que el ruido aditivo tiene una distribución normal de media cero y varianza conocida σI_n ($E \sim N(0, \sigma I_n)$) donde I_n es la matriz identidad de tamaño $n \times n$.

Para la distribución apriori de la imagen X se asume que ésta tiene una distribución normal de media cero y varianza desconocida αI_n ($X \sim N(0, \alpha I_n)$). Dado que se desconoce α se asume para ésta una distribución gamma de parámetros $(1, 1)$ ($\alpha \sim \Gamma(1, 1)$), creando

así un modelo jerárquico que permita recuperar la imagen original antes de ser borrada. Así:

$$\pi(y|x, \alpha) \propto \exp\left(-\frac{1}{2\sigma}(y - Ax)^T(y - Ax)\right), \quad (5.2)$$

$$\pi_{apr}(x) \propto \exp\left(-\frac{1}{2\alpha}x^T x\right), \quad (5.3)$$

$$\pi_{apr}(\alpha) \propto \exp(-\alpha). \quad (5.4)$$

Utilizando las funciones de densidad presentadas en (5.2), (5.3) y (5.4) se obtiene la posterior:

$$\pi_{pos}(x, \alpha) \propto \exp\left(-\frac{1}{2\alpha}\|x\|^2 - \frac{1}{2\sigma}\|y - Ax\|^2 - \alpha\right). \quad (5.5)$$

Como ejemplo práctico se tiene la imagen que se muestra en la Figura 5.1, que es una imagen de 5×5 píxeles, a ésta se le realiza un borrado de la forma 5.1 y satisface las mismas condiciones. Con $\sigma = 0.0821$ se obtiene la imagen borrada que se presenta en la Figura 5.2.

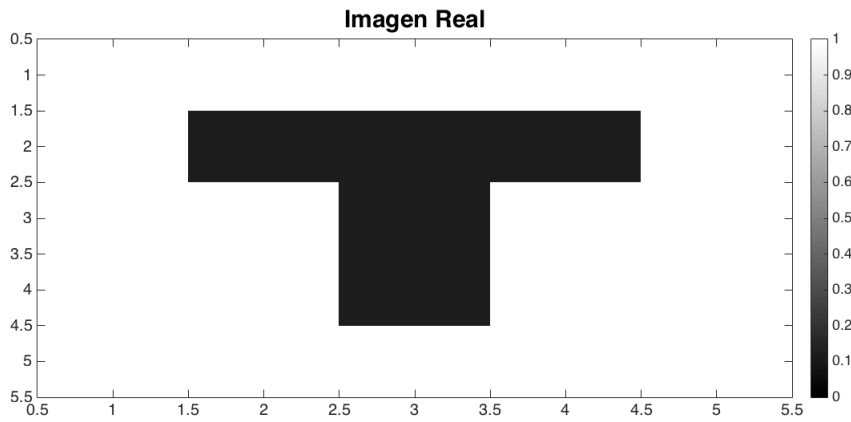


FIGURA 5.1. Imagen de 5×5 píxeles.

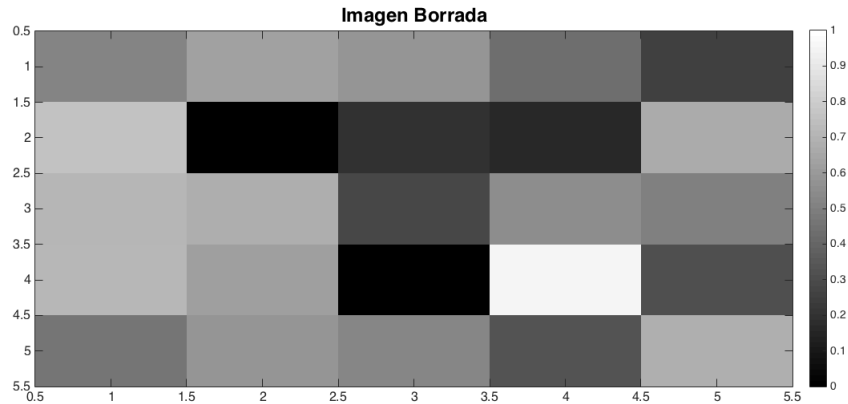


FIGURA 5.2. Imagen de la Figura 5.1 con un borrado de la forma 5.1 para $\sigma = 0.0821$.

Aplicando la metodología bayesiana para recuperar la imagen X a partir de una única imagen borrada Y con la posterior presentada en (5.5) se presenta una comparación entre el máximo posterior con el resultado obtenido al aplicar la regularización de Tikhonov en la Figura 5.3, para así observar las similitudes de éstas imágenes con la imagen real.

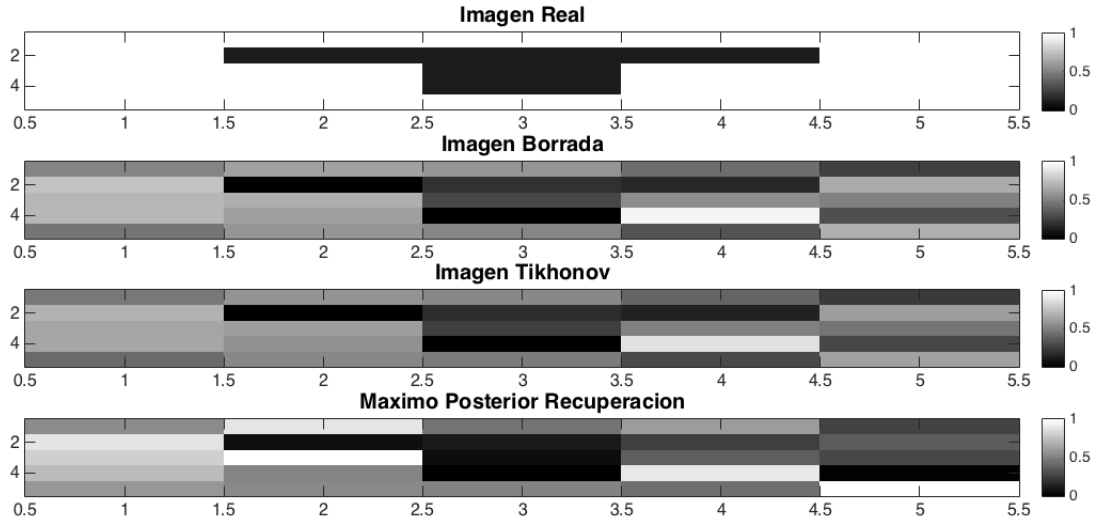


FIGURA 5.3. Comparación entre la imagen real, imagen borrada, resultado obtenido a partir de la regularización de Tikhonov y el máximo posterior de la función de densidad (punto más probable) con función de densidad dada por 5.5 para $\sigma = 0.0821$ con cien mil (100000) iteraciones del modelo MCMC.

En la Figura 5.4 se muestra el logaritmo de la probabilidad de aceptación del MCMC para las muestras de la función posterior (5.5), en este se nota una convergencia aceptable dada por el método, obteniendo cierta estacionariedad entre los valores veinte negativo (-20) y cinco negativo (-5), con esto se podría pensar que el método alcanzó la convergencia para la posterior planteada.

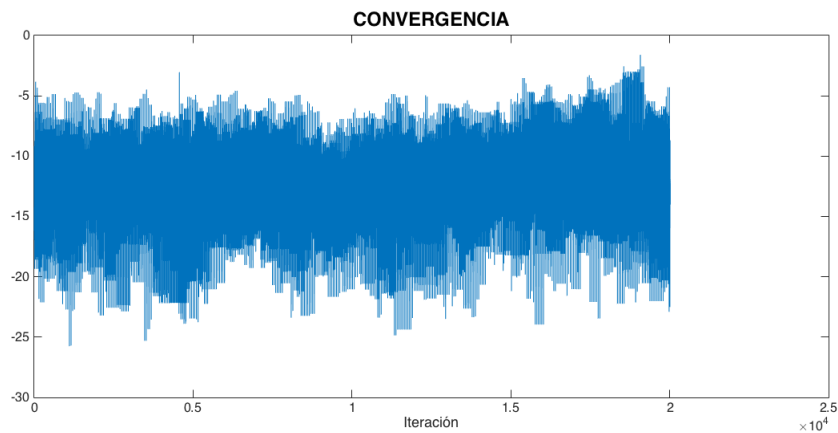


FIGURA 5.4. Logaritmo de la probabilidad de aceptación del MCMC para la posterior (5.5) con cien mil (100000) iteraciones.

En la Figura 5.3 se puede observar que la recuperación de la imagen por el estimador máximo posterior es mejor que la recuperación dada por el modelo de regularización de Tikhonov y la imagen borrada, sin embargo aún está muy lejos de la imagen real. Esto se debe a que se requiere recuperar el valor de 25 píxeles con una posición asociada y se toma como punto de partida una única imagen con 25 píxeles. Es decir para recuperar cada píxel en su posición solo se cuenta con un dato, razón por la cual será casi imposible recuperar la imagen original sin ayudas como Machine Learning entre otras.

Se replantea el ejemplo partiendo de un conjunto de imágenes borradas de la misma imagen para poder recuperar una sola imagen, se considere que con mayor información será posible obtener mejores resultados.

Considere un conjunto de variables aleatorias independientes e igualmente distribuidas Y_i con $i = 1, \dots, n$, y cada Y_i es de la forma $Y_i = AX + E$, donde A es una matriz de convolución, X es la imagen real que se desea recuperar a partir de cada imagen Y_i dada. Además cada Y_i con $i = 1, \dots, n$, posee un ruido aditivo E con distribución normal de media cero y varianza conocida σI_n ($E \sim N(0, \sigma I_n)$). Para la distribución a priori de la imagen X se asume que ésta tiene como función de densidad (5.3) y $\alpha \sim \Gamma(1, 1)$.

Sea $Y = (Y_1, Y_2, \dots, Y_n)$, Y tendrá como función de densidad conjunta o verosimilitud una función de densidad normal al ser cada Y_i independientes e igualmente distribuidas. En (5.7) presenta la función de densidad conjunta de Y .

$$\pi(y|x, \alpha) \propto \exp\left(-\frac{1}{2\alpha} \sum_{i=1}^n \|y_i - Ax\|^2\right). \quad (5.6)$$

Con la verosimilitud dada en (5.6), (5.3) y (5.4) se obtiene la posterior:

$$\pi_{pos}(x, \alpha) \propto \exp\left(-\frac{1}{2\alpha} \|x\|^2 - \frac{1}{2\sigma} \sum_{i=1}^n \|y_i - Ax\|^2 - \alpha\right). \quad (5.7)$$

Teniendo ésta nueva función de densidad posterior se plantea nuevamente el ejemplo tomando $\sigma = 0.0821$ se generan 30 imágenes de la Figura 5.1 con el ruido mencionado y una matriz de convolución presentadas en la Figura 5.5.

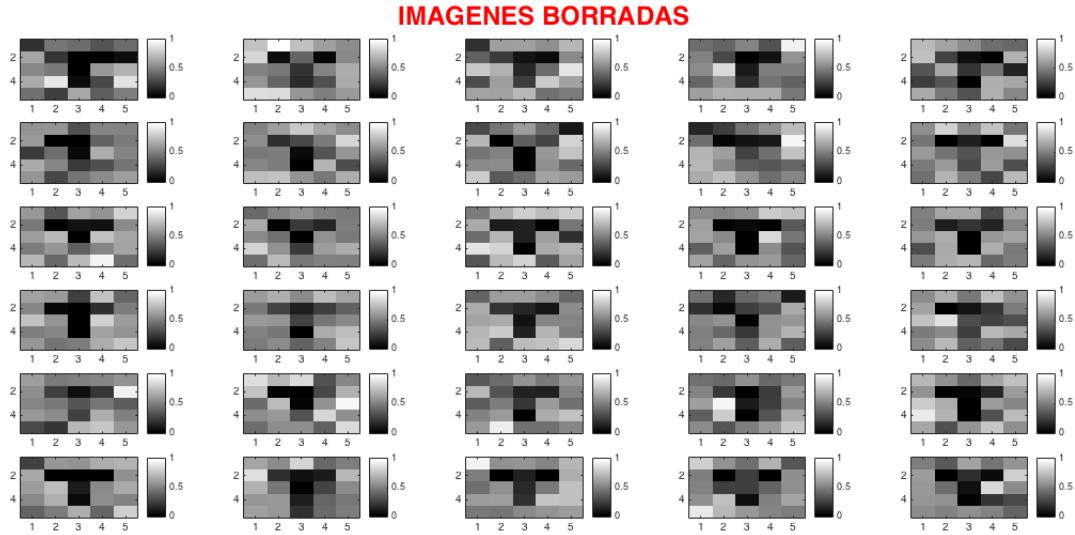


FIGURA 5.5. Imágenes de la Figura 5.1 generadas con un ruido gaussiano presentado en (5.1) con media cero y varianza 0.0821.

Se toman las imágenes de la Figura 5.5 como datos para recuperar la imagen, aplicando el modelo GWCMC con dos millones (2000000) de iteraciones y una varianza $\sigma = 0.0821$, se obtienen los resultados presentados en la Figura 5.6.

Nota. Los parámetros de entrada del modelo GWCMC fueron el logaritmo de la función de densidad posterior (5.7), dos millones (2000000) de iteraciones y una matriz `minit` de valores iniciales con tamaño 26×52 , utilizando el comando `randn` para generar dicha matriz con media 1 y varianza 0.0821. El comando utilizada para ésta matriz es `minit=1+sqrt(0.0821)*randn(26,52)`.

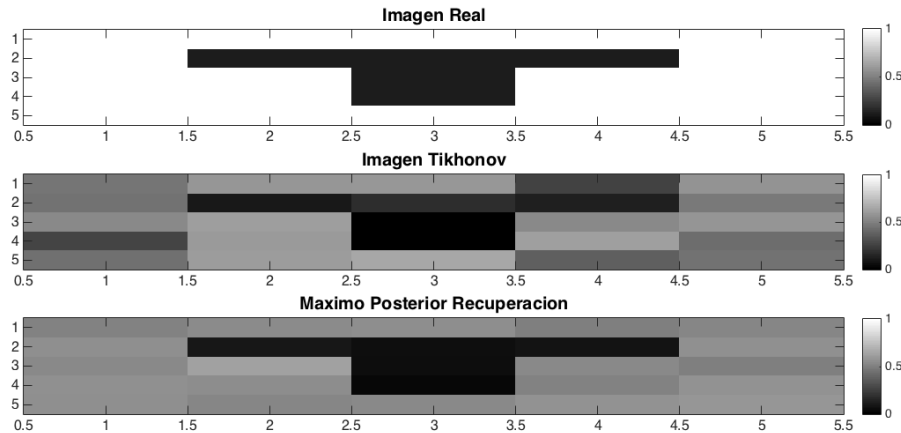


FIGURA 5.6. Comparación de la recuperación de la imagen de la Figura 5.1 por medio del promedio de la regularización de Tikhonov a cada imagen con el máximo posterior de las muestras obtenidas para la función de densidad posterior (5.7).

En la Figura 5.6 se muestra el resultado de aplicar la regularización de Tikhonov a cada una de las 30 imágenes y promediar cada pixel de las 30 imágenes obtenidas, y el resultado obtenido al tomar el estimador máximo posterior de la función de densidad

(5.7), en ambos casos los resultados mejorar significativamente, no obstante la metodología bayesiana presenta mejores resultados en la recuperación de la imagen puesto que el ruido queda casi por completo eliminado. Aunque se la imagen obtenida en por el estimador MAP es más opaca que la imagen original existen métodos deterministas para obtener una imagen con mayor brillo o claridad, como por ejemplo sumar una constante entre cero y uno a cada pixel de la imagen para que ésta obtenga más brillo pues los valores cercanos a cero son más oscuros que los cercanos a uno.

También se puede ver los la matriz de varianzas 5.8 de cada pixel, donde se puede observar que los datos obtenidos no están muy dispersos, pues la mayor varianza que se presenta es de 0.5673.

$$\text{var}(X) = \begin{bmatrix} 0.4277 & 0.4010 & 0.4316 & 0.5440 & 0.4133 \\ 0.3653 & 0.0375 & 0.0738 & 0.0803 & 0.4156 \\ 0.4619 & 0.4400 & 0.0494 & 0.4655 & 0.5233 \\ 0.4622 & 0.3570 & 0.0314 & 0.4248 & 0.4229 \\ 0.5673 & 0.4341 & 0.4087 & 0.4800 & 0.4450 \end{bmatrix} \quad (5.8)$$

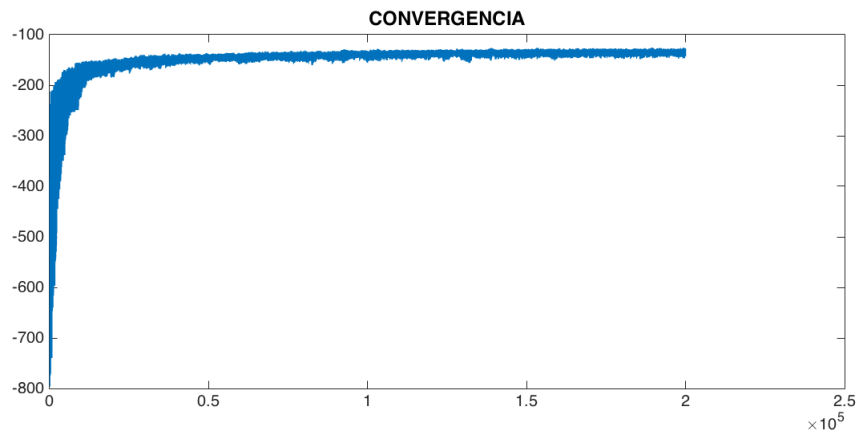


FIGURA 5.7. Logaritmo de la probabilidad de aceptación de cada muestra obtenida por el modelo MCMC para la función de densidad posterior (5.7) con 2 millones (2000000) de iteraciones.

En la Figura 5.7 se observa el logaritmo de la probabilidad de aceptación de cada muestra obtenida por el modelo MCMC para la función de densidad posterior (5.7), donde se evidencia que las muestras obtenidas se estacionalizan a partir del punto 1.5×10^5 , con lo que se puede concluir que los resultados de las muestras convergen para ese número de iteraciones. Para ello se toman otros estimadores como la moda y la media condicional de las muestras obtenidas y así comparar estos resultados con la imagen original, éstos resultados se presentan en la Figura 5.8.

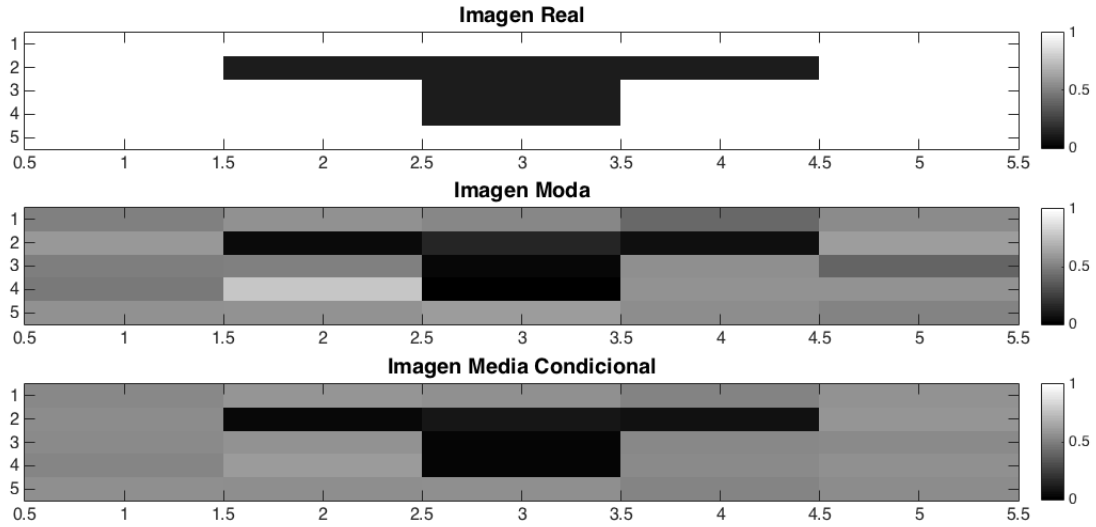


FIGURA 5.8. Comparación de la recuperación de la imagen de la Figura 5.1 utilizando la moda y la media condicional de las muestras obtenidas por el modelo MCMC aplicado a la función de densidad posterior (5.7).

En la Figura 5.8 se puede observar que tanto la media condicional como la moda arrojan mejores resultados que el modelo determinista de regularización de Tikhonov, teniendo la media condicional al igual que el MAP los mejores resultados al momento recuperar la imagen 5.1.

5.2. Aplicación a la estimación de parámetros de ecuaciones diferenciales estocásticas (EDE's)

El modelo CEV o Constant Elasticity of Variance es un modelo de volatilidad estocástica que usa un efecto de apalancamiento. Fue desarrollado por John Cox en 1975 y viene dado por la EDE que se presenta a continuación:

$$dS_t = \alpha S_t dt + \zeta S_t^\beta dB_t, \quad (5.9)$$

donde son parámetros constantes α , ζ y β y satisfacen las condiciones $\beta > 0$ y $\zeta > 0$. El parámetro β controla la relación entre la volatilidad y el precio, siendo la principal característica del modelo. Cuando $\beta < 1$ se observa el llamado efecto de palanca, comúnmente observado en los mercados de valores donde la volatilidad de la acción aumenta a mediada que su precio cae. Por el contrario, en los productos de mercado básico el efecto de palanca es inverso, por lo que la volatilidad del precio de un producto tiende a aumentar a medida que la varianza decrece. Para aplicar el modelo en un activo, éste debe satisfacer las siguientes hipótesis:

- La tasa de interés instantánea α es conocida y constante.
- El precio de la acción es solución a una EDE, ζ y β son conocidos.
- No se pagan dividendos de la acción durante la vida de la opción.

- No hay costos de transacción, impuestos diferenciales, restricciones o las ventas son al descubierto.
- Es posible operar con cualquier fracción de las acciones u opción.

El objetivo es estimar los parámetros del modelo CEV para aproximar la solución mediante el método de milstein para EDE's ver [18, pág 98]. Para ello los parámetros α y σ se estimaran como la media y la desviación estándar de los log-retornos del activo a trabajar, el parámetro β será estimado mediante el uso de estadística bayesiana por algún método de metropolis hastings.

Definición. Se definen los retornos logarítmicos o *log-retornos* (R_t) de un activo como $R_t = \ln\left(\frac{S_t}{S_{t-1}}\right)$ donde S_t es el valor de activo en el tiempo t .

5.2.1. Estimación de β para el modelo CEV

Se cuentan con n datos del valor del activo donde estos datos provienen de una distribución normal con media $S(t_i, \beta)$ y desviación estándar dada por $\sigma\delta$ con σ la desviación muestral de los datos y δ una parámetro desconocido tal que $\delta > 0$ (se asume datos independientes e igualmente distribuidos). En lo anterior $S(t_i, \beta)$ representa la media de las trayectorias simuladas del activo en un periodo de tiempo establecido ya conocido e $i = 1, 2, \dots, n$.

Estos datos al provenir de distribuciones probabilísticas cuentan con una función de densidad conjunta (verosimilitud) dada por:

$$\begin{aligned}\pi(x) &= \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi} \cdot \sigma\delta} \exp \left[-\frac{1}{2} \left(\frac{x - S(t_i, \beta)}{\sigma\delta} \right)^2 \right] \right) \\ &= \frac{1}{(\sqrt{2\pi} \cdot \sigma\delta)^n} \exp \left[\frac{1}{2} \sum_{i=1}^n \left(\frac{x - S(t_i, \beta)}{\sigma\delta} \right)^2 \right].\end{aligned}\quad (5.10)$$

Dado que los parámetros β y δ son desconocidos, se plantea como metodología bayesiana asumir una distribución apriori para dichos parámetros, como se conoce de antemano que dichos parámetros son positivos, se toman las siguientes posibilidades:

1. Para δ se asume una distribución Gamma con parámetros $\alpha = 1$ y $\lambda = 1$. Con lo que se obtiene una apriori dada por:

$$\pi_{apr}(\delta) = e^{-\delta}. \quad (5.11)$$

2. Se asume una distribución Gamma con parámetros $\alpha = 1$ y $\lambda = 1$. Con lo que se obtiene una apriori dada por:

$$\pi_{apr}(\beta) = e^{-\beta}. \quad (5.12)$$

Una vez dadas las distribuciones apriori y la verosimilitud se procede a encontrar la distribución posterior, así se tiene la posterior:

$$\begin{aligned}\pi_{pos}(\delta, \beta|x) &= \frac{1}{(\sqrt{2\pi} \cdot \sigma\delta)^n} \exp \left[-\frac{1}{2} \sum_{i=1}^n \left(\frac{x - S(t_i, \beta)}{\sigma\delta} \right)^2 \right] \cdot e^{-\delta} \cdot e^{-\beta} \\ &= \frac{1}{(\sqrt{2\pi} \cdot \sigma\delta)^n} \exp \left[-\frac{1}{2} \sum_{i=1}^n \left(\frac{x - S(t_i, \beta)}{\sigma\delta} \right)^2 - \delta - \beta \right].\end{aligned}\quad (5.13)$$

Para probar la metodología se toma el valor del dolar en Colombia en un periodo de 2 meses, con un total de 58 datos, desde el 16 de Junio a 12 de Agosto de 2017, buscando realizar una predicción para así obtener una predicción hasta el día 11 de septiembre de 2017.

Una vez aplicada la metodología se observa que en la Figura 5.9 el histograma para la posterior de β , en el que se ve que la media condicional y el máximo posterior (punto más probable) están muy cercanos, por otra parte la moda de los datos obtenidos ésta bastante alejada de estos dos.

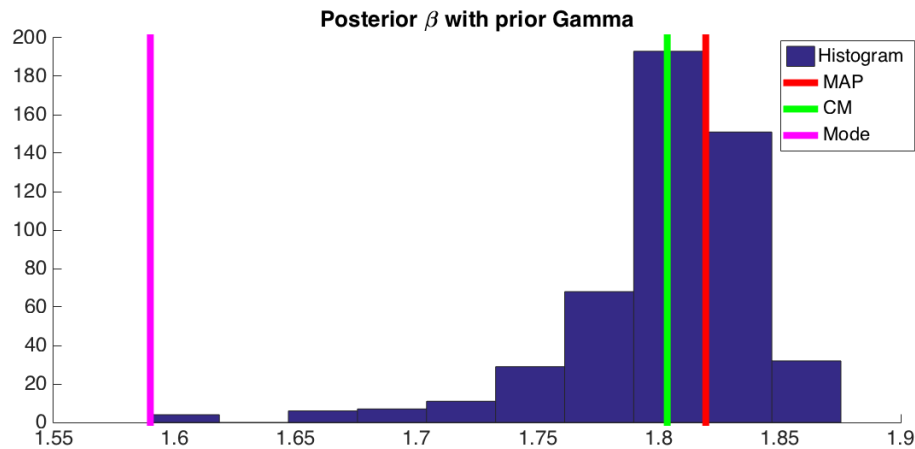


FIGURA 5.9. Histograma de la función posterior para β dada por (5.13) para 5000 iteraciones.

Los resultados para las estimaciones de β son presentados en la Tabla 5.1.

Estimador	MAP	CM	MODA
Beta	1.8195	1.8036	1.5900

TABLA 5.1. Estimaciones de β para los estimadores MAP (máximo posterior), CM (media condicional) y moda respectivamente.

Es importante ver las probabilidades de aceptación de cada muestra, para ello en la Figura 5.10 se muestran los logaritmos de las probabilidades de aceptación de cada una de las muestras aceptadas por el método MCMC empleado, en dicha imagen se cortó el 30% de los datos aceptados. En la Figura 5.10 se observa una buena convergencia para los valores estimados, al observarse que el logaritmo de las probabilidades de aceptación no sigue creciendo, pues se está volviendo estacional al rededor de -315.

Para verificar el ajuste del modelo CEV se compara la media de 30 trayectorias generadas con el método de milstein [18] y parámetros $\alpha = 0.0001841$ y $\sigma = 0.0000243$ con los datos reales del dolar del 16 de Junio al 11 de Septiembre de 2017 (88 datos), periodo en el que se comprende los datos iniciales y la predicción del activo.

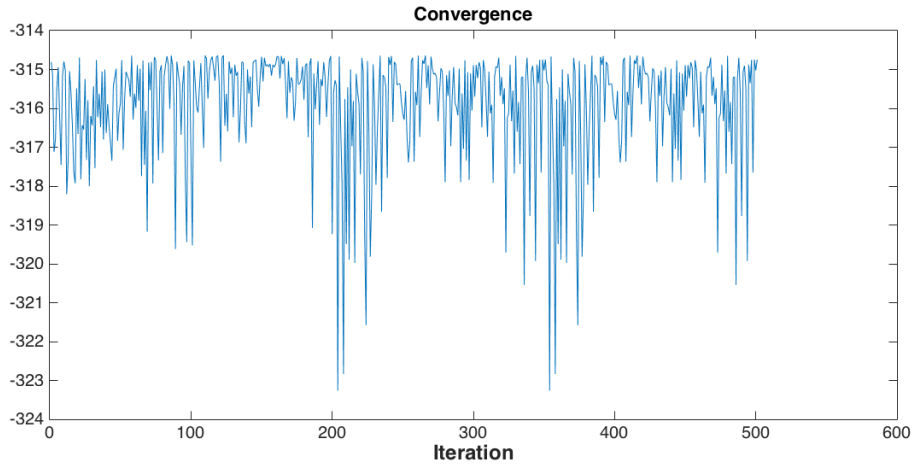


FIGURA 5.10. Logaritmo natural de la probabilidad de aceptación por el modelo MCMC de la función posterior para β dada por (5.13) con 5000 iteraciones.

En la Figura 5.11 se muestra el ajuste del modelo con $\beta = MAP$, y un intervalo de confianza del 95 %, en este ajuste la predicción para el día 11 de Septiembre fue de \$3082.90 siendo el valor real \$2909.2, teniéndose un error absoluto de \$170.77 y relativo de 5.8 %.

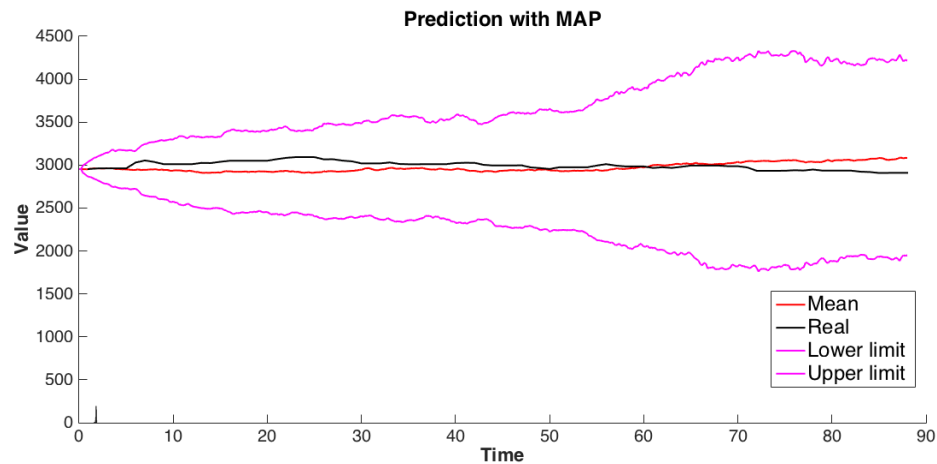


FIGURA 5.11. Ajuste y predicción de del dólar en Colombia usando el modelo CEV para β dado por el máximo posterior.

En la Figura 5.12 se muestra el ajuste del modelo con $\beta = CM$, y un intervalo de confianza del 95 %, en este ajuste la predicción para el día 11 de Septiembre fue de \$3071.30 siendo el valor real \$2909.20, teniéndose un error absoluto de \$162.15 y relativo de 5.5 %.

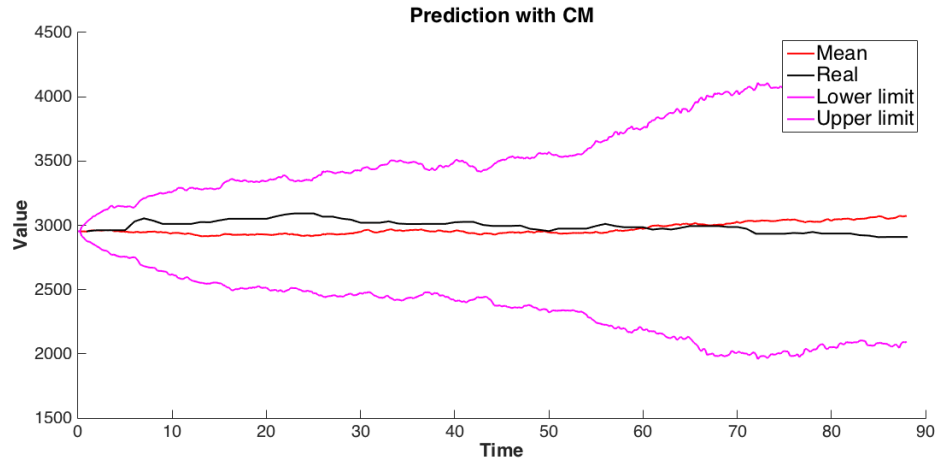


FIGURA 5.12. Ajuste y predicción de del dólar en Colombia usando el modelo CEV para β dado por la media condicional.

En la Figura 5.13 se muestra el ajuste del modelo con $\beta = MODA$, y un intervalo de confianza del 95 %, en este ajuste la predicción para el día 11 de Septiembre fue de \$3012.90 siendo el valor real \$2909.2, teniéndose un error absoluto de \$103.70 y relativo de 3.5 %.

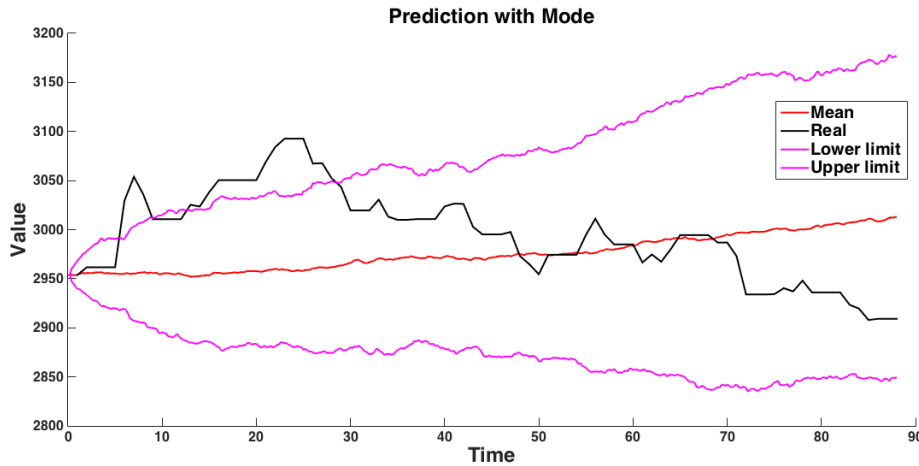


FIGURA 5.13. Ajuste y predicción de del dólar en Colombia usando el modelo CEV para β dado por la moda.

Los datos anteriormente mencionados se presentan en la Tabla 5.2.

	MAP	CM	MODA
β	\$1.8195	\$1.8036	\$1.5900
Predicción	\$3082.90	\$3071.30	\$3012.90
Error Absoluto	\$170.77	\$162.15	\$103.70
Error relativo	5.8 %	5.5 %	3.5 %

TABLA 5.2. Estimaciones del dólar para el 11 de Septiembre de 2017 con el modelo CEV y diferentes estimaciones de β .

Aunque la moda genera la mejor predicción de los estimadores presentados, el intervalo de confianza no contiene a los datos reales en un intervalo, razón por la cual no se considere un buen estimador para este ejemplo.

5.3. Problema inverso de impedancia tomográfica

Para los resultados del problema inverso de impedancia tomográfica se asume que la impedancia es conocida y se buscará estimar la conductividad γ a partir de la posterior presentada en (4.27), se considerarán diferentes casos para observar la eficiencia de la estadística bayesiana en los problemas inversos.

Primero se generan datos sintéticos del problema directo para una conductividad γ en un área dada por el cuadrado $[0, 1] \times [0, 1]$, tomando 25 pixeles, en las que la inclusión del centro es diferente de las demás.

$$\gamma = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 9 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \quad (5.14)$$

Se observa que la conductividad tiene una inclusión en el centro de la matriz, que en forma gráfica se puede observar en la Figura 5.14.

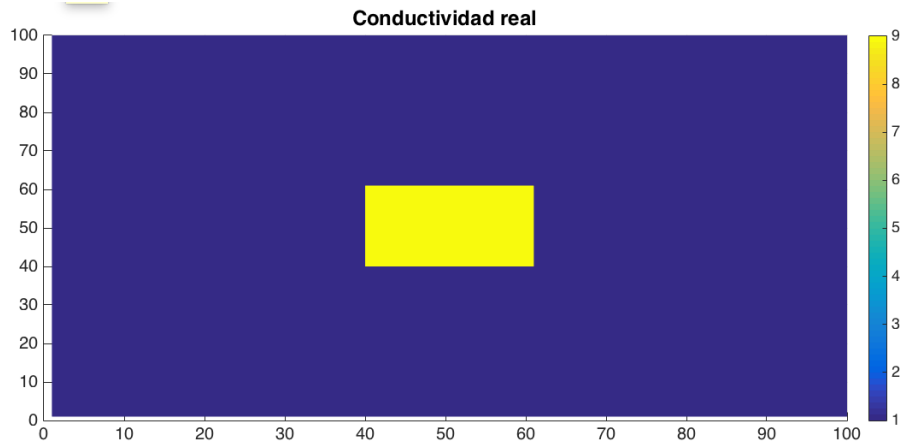


FIGURA 5.14. Conductividad con inclusión en el centro.

Conociendo la conductividad se genera una matriz de corrientes que satisface (4.3) sumando las filas y sus columnas son linealmente independientes,

$$I = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}. \quad (5.15)$$

Posteriormente se soluciona el problema directo de tomografía de impedancia eléctrica sobre el dominio $[0, 1] \times [0, 1]$ con 10 electrodos sobre el borde del dominio y una discretización muy fina de 100000 pixeles como se muestra en la Figura 5.15, ésta discretización se realiza tomando 25 grandes pixeles donde se tiene un valor de conductividad y sobre cada uno de estos 25 pixeles se toma un subdominio de 20×20 .

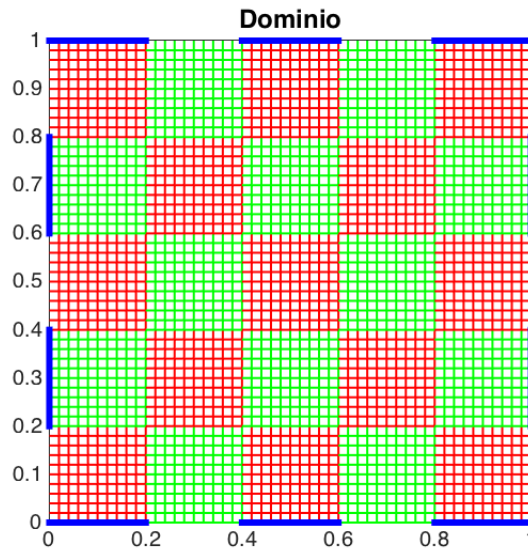


FIGURA 5.15. Dominio del problema EIT con cien particiones en el eje X y cien particiones en el eje Y para un total de cien mil pixeles, tomando una discretización cuadrada como en 3.3 con 10 electrodos en los bordes ocupando 20 pixeles cada electrodo.

Se asume una impedancia conocida de 1, es decir $T(\gamma, z) = T(\gamma, 1) = T(\gamma)$, así al solucionar el problema directo dado por 4.15 mediante elementos finitos asumiendo que la

solución está en el espacio Q_1 ver [11], se obtiene la matriz $T^{-1}(\gamma)$ dada por:

$$T^{-1}(\gamma) = \begin{bmatrix} 0.164 & -0.003 & 0.002 & 0.003 & 0.004 & 0.004 & 0.004 & 0.003 & 0.001 & -0.009 \\ -0.003 & 0.175 & -0.003 & -0.001 & 0.003 & 0.004 & 0.004 & 0.004 & 0.003 & -0.001 \\ 0.002 & -0.003 & 0.164 & -0.009 & 0.001 & 0.003 & 0.004 & 0.004 & 0.004 & 0.003 \\ 0.003 & -0.001 & -0.009 & 0.173 & -0.001 & 0.001 & 0.003 & 0.004 & 0.004 & 0.003 \\ 0.004 & 0.003 & 0.001 & -0.001 & 0.173 & -0.009 & -0.001 & 0.003 & 0.003 & 0.004 \\ 0.004 & 0.004 & 0.003 & 0.001 & -0.009 & 0.164 & -0.003 & 0.002 & 0.003 & 0.004 \\ 0.004 & 0.004 & 0.004 & 0.003 & -0.001 & -0.003 & 0.175 & -0.003 & -0.001 & 0.003 \\ 0.003 & 0.004 & 0.004 & 0.004 & 0.003 & 0.002 & -0.003 & 0.164 & -0.009 & 0.001 \\ 0.001 & 0.003 & 0.004 & 0.004 & 0.003 & 0.003 & -0.001 & -0.009 & 0.173 & -0.001 \\ -0.009 & -0.001 & 0.003 & 0.003 & 0.004 & 0.004 & 0.003 & 0.001 & -0.001 & 0.174 \end{bmatrix} \quad (5.16)$$

Con el operador lineal $T(\gamma) = \left(T^{-1}(\gamma)\right)^{-1}$ (5.16) y la matriz de corrientes (5.15) se halla la matriz de voltajes, dada por:

$$U = \begin{bmatrix} 2.977 & 0.127 & 0.005 & 0.027 & 0.019 & -0.039 & 0.002 & -0.0487 & -0.1933 \\ -2.811 & 2.809 & 0.035 & 0.065 & 0.024 & -0.019 & 0.016 & -0.034 & -0.0664 \\ -0.123 & -2.981 & 2.876 & 0.198 & 0.051 & 0.001 & 0.041 & -0.020 & -0.0281 \\ -0.087 & -0.141 & -2.735 & 2.867 & 0.069 & 0.019 & 0.044 & -0.018 & -0.0155 \\ -0.045 & -0.017 & -0.067 & -2.863 & 2.735 & 0.138 & 0.086 & -0.002 & 0.0124 \\ -0.039 & 0.001 & -0.046 & -0.193 & -2.880 & 2.984 & 0.121 & 0.003 & 0.0247 \\ -0.026 & 0.018 & -0.030 & -0.066 & -0.035 & -2.815 & 2.813 & 0.038 & 0.0659 \\ -0.003 & 0.041 & -0.017 & -0.027 & -0.002 & -0.132 & -2.975 & 2.877 & 0.1949 \\ 0.017 & 0.047 & -0.013 & -0.017 & 0.005 & -0.091 & -0.141 & -2.731 & 2.8649 \\ 0.142 & 0.093 & -0.005 & 0.010 & 0.013 & -0.043 & -0.010 & -0.063 & -2.8594 \end{bmatrix} \quad (5.17)$$

Los datos del voltaje (5.17) obtenidos de la solución del problema directo se pueden observar en la Figura 5.16.

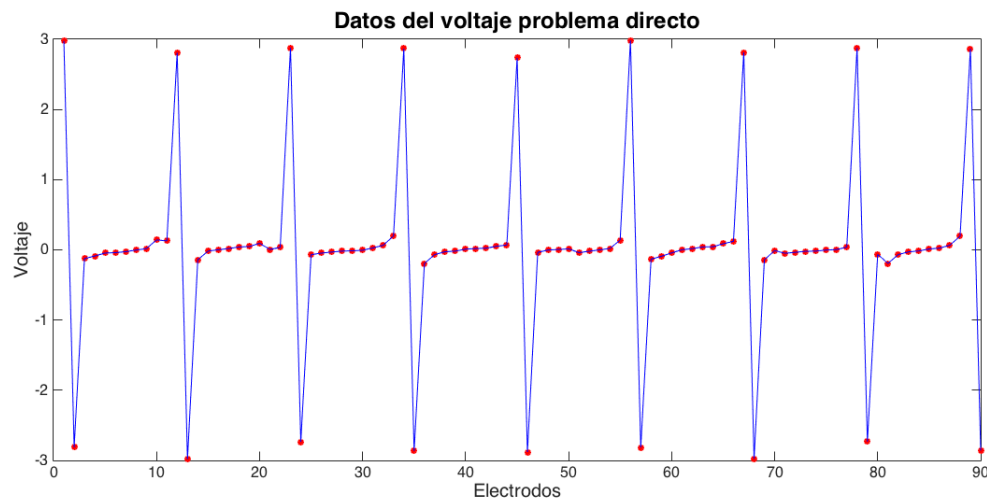


FIGURA 5.16. Voltajes obtenidos de la solución del problema directo mediante la aplicación del método de los elementos finitos.

Una vez se tienen los datos sintéticos (5.17) y (5.15) se procede a realizarles un borrado gaussiano obteniendo las gráficas para voltajes y corrientes presentadas respectivamente en la Figuras 5.17 y 5.18. Se estima el valor de la conductividad para una impedancia conocida de $z = 1$ en todo el dominio.

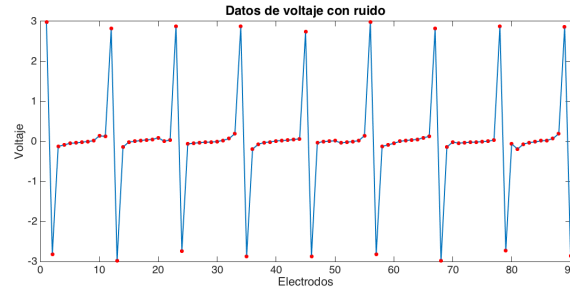


FIGURA 5.17. Voltajes con ruido.

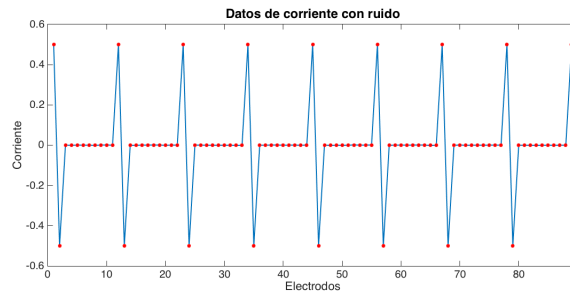


FIGURA 5.18. Corrientes con ruido.

FIGURA 5.19. Datos iniciales con ruido gaussiano de media cero (0) y varianza 0.00000022.

Una vez se realizó la construcción de los datos sintéticos para la validación del modelo, se obtuvieron los siguientes datos;

1. Se supone que se conoce donde se encuentra la inclusión y se quiere estimar dos parámetros de γ , el de la inclusión y el exterior de la inclusión, obteniendo los resultados presentados en la Tabla 5.3

	MAP	MODA	CM	REAL	VARIANZA
Inclusión	9.6370	9.0107	9.5963	9	0.0588
Exterior	0.9973	1.0032	0.9974	1	0.0000

TABLA 5.3. Resultados obtenidos para la conductividad γ con 20000 iteraciones en una malla de 50×50 .

Los resultados de la Tabla 5.3 muestran que el mejor estimador es la moda en este caso, aunque gráficamente los resultados muestran una clara inclusión en el centro de la malla como se puede observar en la Figura 5.20.

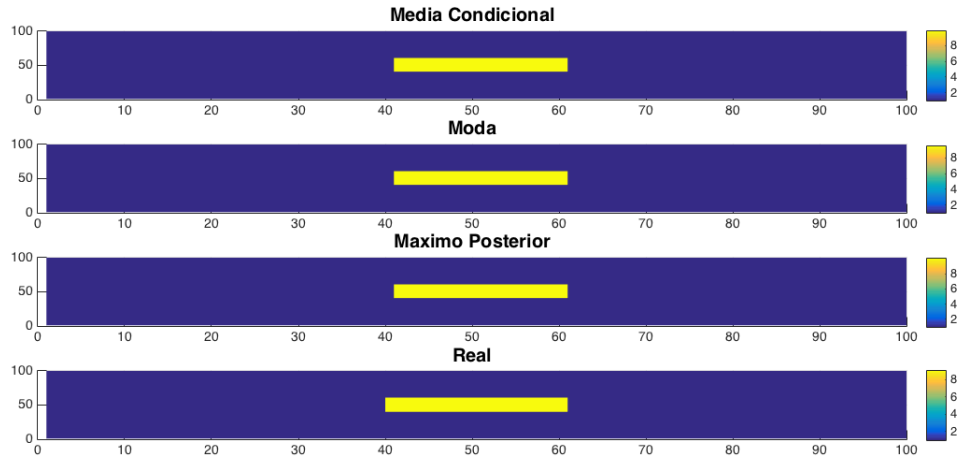


FIGURA 5.20. Versus entre los estimadores de la media condicional, moda, máximo posterior y el valor real de la inclusión de la conductividad γ .

Se puede ver en la Figura 5.21 la convergencia en probabilidad para las muestras aceptadas, observando que la convergencia tiende a estacionalizarse, lo que gráficamente es un buen indicador que las muestras están convergiendo en probabilidad.

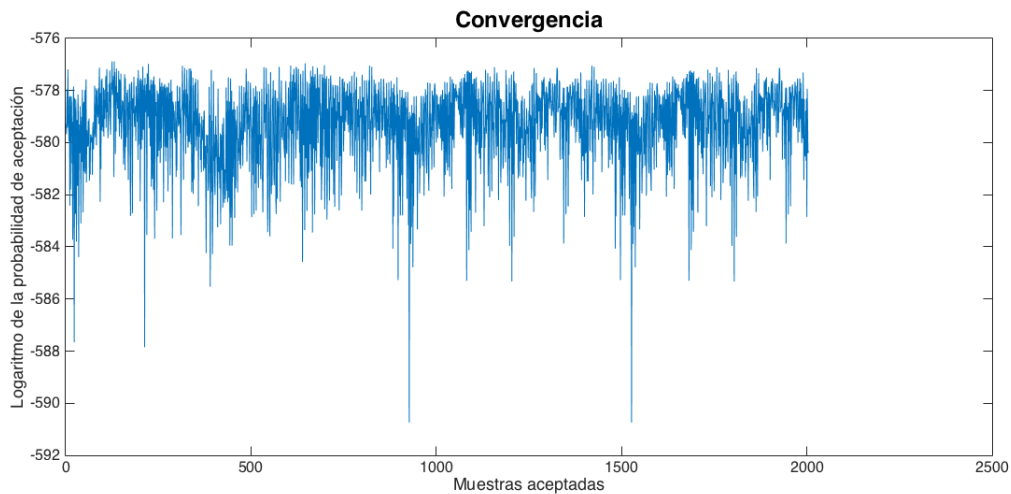


FIGURA 5.21. Logaritmo de la probabilidad de aceptación de cada muestra obtenida por el modelo MCMC para la 20000 iteraciones.

Las distribuciones marginales para los parámetros obtenidos se pueden observar en la Figura 5.24, observando que para los pixeles que no corresponden al centro se tiene una mayor aceptación en su valor real que en el pixel de inclusión central.

Se obtuvo buenos resultados para 20000 iteraciones pues se estimaron solo los parámetros, en realidad no siempre es fácil conocer la inclusión, para ello se toman diferentes metodologías que serán presentadas a continuación.

Se puede dividir el dominio en un número finito subdominios que no supere la cantidad de electrodos para ver en cuál de éstos se presenta una mayor conductividad que el el resto, para así identificar en con un análisis posterior donde se presenta la mayor conductividad

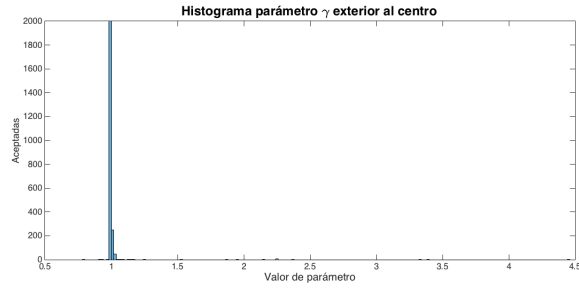


FIGURA 5.22. Función de densidad marginal para el exterior de la inclusión.

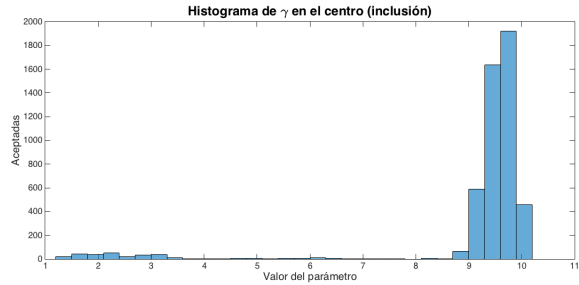


FIGURA 5.23. Función de densidad marginal para la inclusión en el centro del dominio.

FIGURA 5.24. Funciones de densidad marginales obtenidas para los parámetros estimados por el método bayesiano.

en ese dominio dando solución al problema inverso de EIT. Para éste caso se dividirá el dominio en cinco subdominios disyuntos como se muestra en la Figura 5.25 para recuperar cinco parámetros (uno en cada subdominio).



FIGURA 5.25. División del subdominio en cinco “subdominios verticales” donde se recuperarán 5 parámetros.

Los electrodos se asumen de la misma forma que en la Figura 5.15, así se obtienen los resultados presentados en la Figura 5.26 para cien mil (100000) iteraciones con una malla 50×50 , donde se puede ver que existe una inclusión en el dominio central.

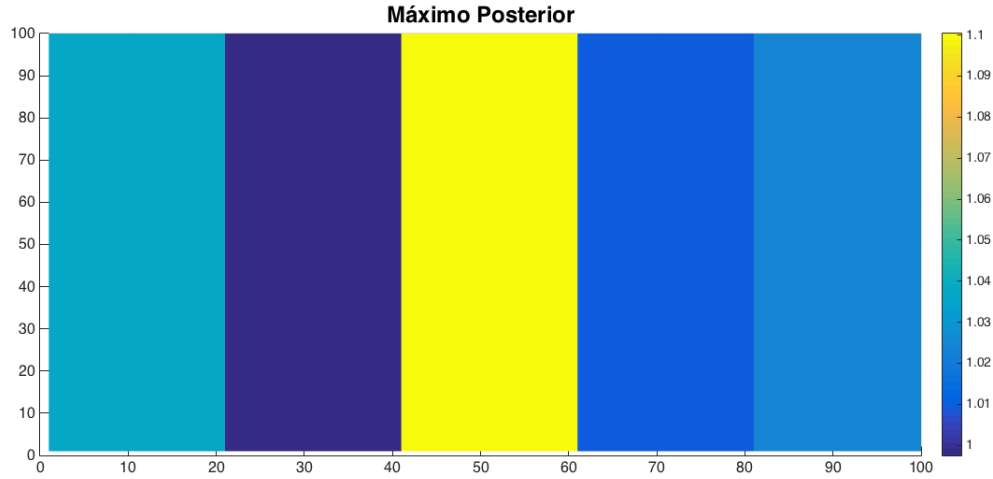


FIGURA 5.26. Máximo posterior de los cinco parámetros estimados.

En la Tabla 5.4 se puede observar los resultados presentados para las estimaciones, donde también es claro que la mayor conductividad se encuentra en el centro y es donde se deben estimar más parámetros, también se puede ver que en los dominios contiguos a la inclusión se presenta una mayor varianza de los datos. En la Figura 5.27 se puede ver la convergencia en probabilidad de ésta estimación.

MAP	Parámetro 1	Parámetro 2	Parámetro 3	Parámetro 4	Parámetro 5
Estimación	1.036	0.9975	1.2004	1.0096	1.0247
Varianza	0.0461	1.2183	0.2545	0.6003	0.0203

TABLA 5.4. Resultados obtenidos para la conductividad γ con 50000 iteraciones en una malla de 50×50 sobre el dominio 5.25 con los parámetros presentados en el mismo orden.

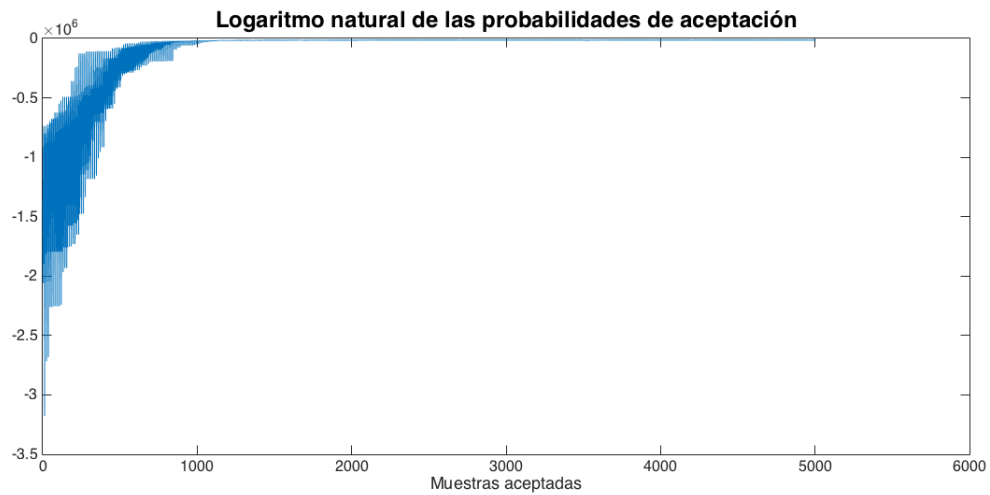


FIGURA 5.27. Logaritmo natural de las probabilidades de aceptación de las muestras para los parámetros de la Figura 5.26.

Al igual que en el ejemplo anterior también se podría crear un dominio de franjas horizontales y contrastar las soluciones para ser más precisos en cuanto a la región que se debe estudiar con un mayor detalle. A medida que se incrementan el número de parámetros a estimar se hace más difícil obtener el resultado más cercano a la inclusión, en la Figura 5.28 se muestra el resultado de la media condicional obtenido de estimar nueve parámetros y poder detectar la inclusión. Para ésto se tomó cincuenta mil iteraciones con una malla 25×25 para eficiencia computacional.

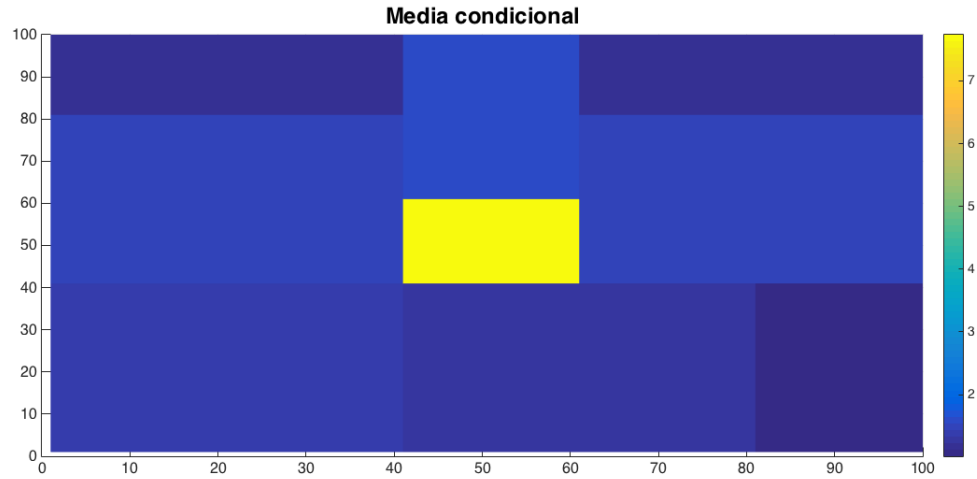


FIGURA 5.28. Media condicional de la conductividad obtenida al estimar 9 parámetros sobre el dominio.

En la Figura 5.29 se tiene la convergencia en probabilidad que aunque no es algo óptimo, presenta una convergencia aceptable como muestra la gráfica de forma logarítmica.

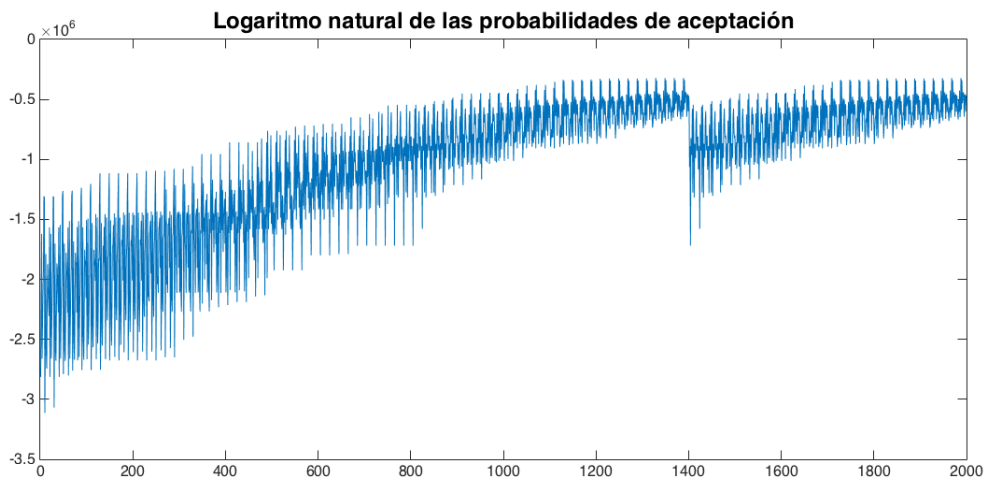


FIGURA 5.29. Logaritmo natural de las probabilidades de aceptación de las muestras para 9 parámetros.

Para éste caso el valor que recupero de la inclusión fue 7.7335 y los valores de la no inclusión oscilaron entre 1.0173 y 1.6073, lo que evidencia que éste resultado se aleja más de los resultados obtenidos en la Tabla 5.3.

Por último se presenta una estimación para 25 pixeles en el dominio, con los mismo 10 electrodos, los resultados en éste caso no son buenos pues para estimar más parámetros se necesita más información, es decir, se necesitarían más electrodos sobre el dominio para tener más resultados de corriente y voltaje. En la Figura 5.30 se muestran los resultados obtenidos y en la Figura 5.31 se presenta la convergencia de éstos parámetros recuperados, estos resultados se tomaron sobre una malla 50×50 con cien mil iteraciones.

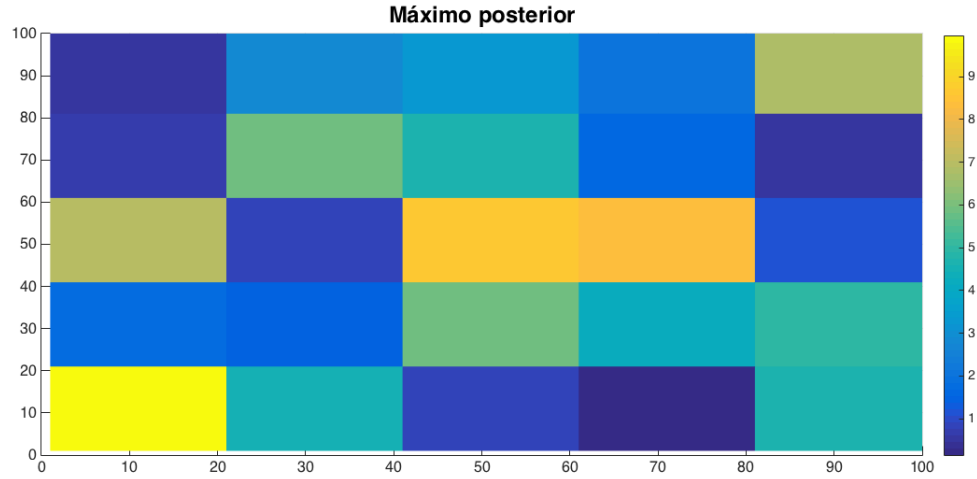


FIGURA 5.30. Máximo posterior de la conductividad obtenida al estimar 25 parámetros sobre el dominio.

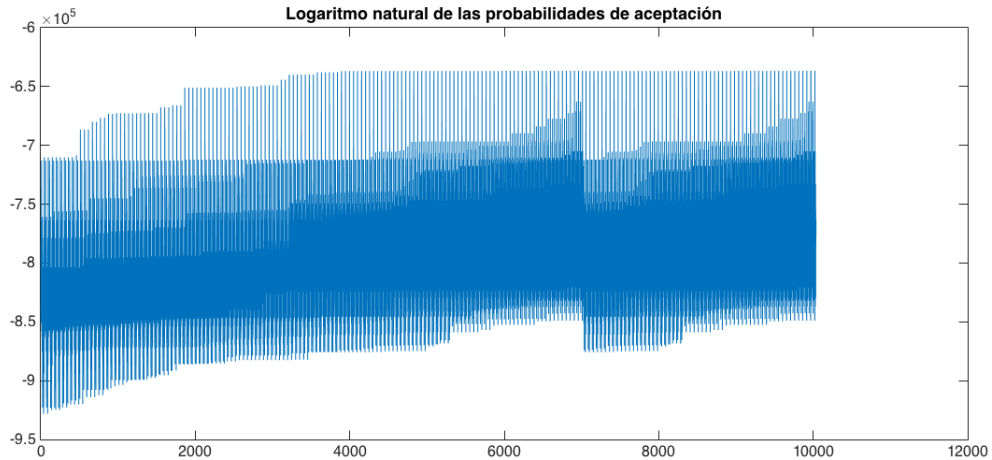


FIGURA 5.31. Logaritmo natural de las probabilidades de aceptación de las muestras para 25 parámetros.

Algoritmos numéricos

En éste apéndice se presentan los algoritmos métodos de Runge-Kutta para ecuaciones diferenciales ordinarias y Milstein para ecuaciones diferenciales estocásticas.

A.1. Método Runge - Kutta de orden 4

El método de Runge-Kutta (ver [7]) llamado así en honor a los matemáticos alemanes Runge y Kutta es un algoritmo que resuelve de forma eficaz una ecuación diferencial y coincide con los primeros n términos del desarrollo de la serie de Taylor. Así una ecuación diferencial de la forma $dy/dx = f(x, y)$ tiene una solución aproximada dada por:

$$\begin{aligned}y_{n+1} &= y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \\k_1 &= hf(x_n, y_n), \\k_2 &= hf\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1\right), \\k_3 &= hf\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_2\right), \\k_4 &= hf(x_n + h, y_n + k_3).\end{aligned}$$

La aproximación de la solución y está dada en un intervalo (a, b) y $h = \frac{b-a}{2}$.

A.2. Método de Milstein

El método de Milstein es utilizado para aproximar la solución de una ecuación diferencial estocástica. Se obtiene truncando la serie de Taylo-Itô (ver [18]), es un método muy utilizado pues no presenta restricción en la forma de la ecuación estocástica a diferencia de los métodos de orden superior utilizados para éstas ecuaciones.

Una ecuación diferencial estocástica de la forma,

$$dX_t = \alpha(t, X_t)dt + \sigma(t, X_t)dB_t,$$

tiene una solución aproximada en un intervalo de tiempo $[0, T]$ dada por:

$$X_j = X_{j-1} + \alpha(X_{j-1})\Delta t + \beta(X_{j-1})\Delta B_{\tau_j} + \frac{1}{2}\beta(X_{j-1})\beta^{(1)}(X_{j-1})\left((\Delta B_{\tau_j})^2 - \Delta t\right), \quad j = 1, 2, \dots, N,$$

con $\Delta t = T/N$, N nodos igualmente espaciados en el tiempo y ΔB_{τ_j} la variación j -ésima en la trayectoria browniana.

Conclusiones, recomendaciones y trabajo futuro

En este trabajo se implemento la metodología bayesiana en diversas aplicaciones enfocadas a resolver problemas inversos provenientes de modelos físicos, entre ellas el de recuperación de imágenes y problemas de tipo financiero como predicción de activos. Se presentaron metodologías para resolver ecuaciones lineales como el problema de convolución para imágenes (5.1), ecuaciones diferenciales ordinarias (EDO's) como recuperar los parámetros del modelo logístico (2.18), ecuaciones diferenciales parciales (EDP's) del problema de EIT (4.11) y ecuaciones diferenciales estocásticas (EDE's) como el modelo CEV (5.9), mostrando así una combinación entre métodos numéricos y estadística bayesiana y como la combinación de éstas dos áreas permite obtener ecuaciones confiables que permitan modelar el comportamiento de un conjunto de datos.

En el caso de recuperación de imágenes (5.1) se puede ver que al tenerse un borrado es complicado obtener buenos resultados si la imagen borrada no es “muy informativa” con respecto a la imagen real, es decir, si el borrado es muy grande, o si la distribución a priori asociada a la imagen no es indicada. Los mejores resultados se obtuvieron al tomarse varias imágenes borradas del mismo elemento, algo que en realidad es muy complicado de obtener en cualquier proyecto pues solo se tendría una única imagen borrada. Podría llegar a ser útil usar métodos de recuperación de imágenes que combinan la minimización de funcionales como es el problema de Tikhonov (1.4) en combinación con métodos bayesianos.

En el caso de la ecuación diferencial estocástica (5.9) se presenta el problema de ajuste versus suavidad, que es un problema muy común al ajustar datos, pues en la metodología propuesta para la distribución posterior (5.13) se considera la media de las trayectorias, al tener una gran cantidad de trayectorias (usualmente más de 50) se pierde ajuste pero se gana suavidad pues la media tiende a ser una recta, y para pocas trayectorias se genera un mejor ajuste pero se pierde suavidad en la trayectoria. Ésto dificulta obtener un valor que genere un buen ajuste con una trayectoria suave. Si se desea solo obtener una predicción en un periodo final es preferible generar un gran número de trayectorias, pues se observa que ésta forma reduce el error en la predicción, para solucionar el problema del mínimo de trayectorias que se desea, se puede tomar como referencia el teorema del límite central y es conocido que para 30 o más valores éste tiene un buen comportamiento, así que es recomendable usar 30 trayectorias para tener una buena relación entre suavidad y ajuste.

En el problema de EIT (4.11) se uso una a priori CAMG con inclusión en el centro que es el caso más complicado a estimar, dado que el centro es el punto más lejano de los electrodos y la corriente no llegará con la misma intensidad a éste pixel en comparación con los demás pixeles de la malla. Se utilizó en método de los elementos finitos con mallas cuadradas de tal forma que la longitud total cubierta por los electrodos represente la mitad

del perímetro de la frontera del dominio, éstos electrodos se distribuyen uniformemente sobre la frontera. La inclusión se tomó considerablemente mayor con respecto a los demás píxeles para verificar que tanta información se podría recuperar a partir de los datos. También se observó que la distribución apriori es fundamental para obtener algoritmos prácticos, y en combinación con un modelo que resuelva el problema directo puede hacer una gran diferencia en la aceptación de muestras de la distribución posterior.

Se puede continuar desarrollando éste trabajo en varias direcciones. Entre ellas se puede incluir, en el caso de EIT, donde se puede considerar la eficiencia de la solución del problema directo y, más difícil de tratar, la estimación de la dimensión máxima del espacio de parámetros informado por los datos. ésto relacionado con modelos de reducción de escala y métodos de elementos finitos multiescala, ver [8]. También se podría considerar una apriori más general que las CAMG, que son generadas de forma iterativa como se muestra en [2].

Bibliografía

- [1] Richard C Aster, Brian Borchers, and Clifford H Thurber, *Parameter estimation and inverse problems*. 2005.
- [2] Johnathan M Bardsley and Jari Kaipio, *Gaussian markov random field priors for inverse problems.*, *Inverse Problems & Imaging* **7** (2013), no. 2.
- [3] Tan Bui-Thanh, Omar Ghattas, James Martin, and Georg Stadler, *A computational framework for infinite-dimensional bayesian inverse problems part i: The linearized case, with application to global seismic inversion*, *SIAM Journal on Scientific Computing* **35** (2013), no. 6, A2494–A2523.
- [4] George Casella and Roger L Berger, *Statistical inference*, vol. 2, Duxbury Pacific Grove, CA, 2002.
- [5] J Andrés Christen, Colin Fox, et al., *A general purpose sampling algorithm for continuous distributions (the t-walk)*, *Bayesian Analysis* **5** (2010), no. 2, 263–281.
- [6] Martin Crowder, *An intermediate course in probability, by allan gut*, *International Statistical Review* **78** (2010), no. 1, 154–154.
- [7] GERALD Curtis and WHEATLEY PATRICK, *Análisis numérico con aplicaciones*, México: Editorial Prentice-Hall (2000).
- [8] Yalchin Efendiev, Juan Galvis, and Thomas Y Hou, *Generalized multiscale finite element methods (gmsfem)*, *Journal of Computational Physics* **251** (2013), 116–135.
- [9] WD Evans, *Partial differential equations*, *Bulletin of the London Mathematical Society* **20** (1988), no. 4, 375–376.
- [10] Daniel Foreman-Mackey, David W Hogg, Dustin Lang, and Jonathan Goodman, *emcee: the mcmc hammer*, *Publications of the Astronomical Society of the Pacific* **125** (2013), no. 925, 306.
- [11] Carlos Humberto Galeano, Juan Miguel mantilla, and Juan Carlos galvis, *El método de los elementos finitos: Un enfoque teórico práctico*, Universidad Nacional de Colombia, 2016.
- [12] Jonathan Goodman, Jonathan Weare, et al., *Ensemble samplers with affine invariance*, *Communications in applied mathematics and computational science* **5** (2010), no. 1, 65–80.

-
- [13] W Keith Hastings, *Monte carlo sampling methods using markov chains and their applications*, (1970).
- [14] Robert V Hogg and Allen T Craig, *Introduction to mathematical statistics. (5 edition)*, Upper Saddle River, New Jersey: Prentice Hall, 1995.
- [15] Valéria Iório, *Edp un curso de graduación*, IMCA, 1999.
- [16] Jari Kaipio and Erkki Somersalo, *Statistical and computational inverse problems*, vol. 160, Springer Science & Business Media, 2006.
- [17] Jari P Kaipio, Ville Kolehmainen, Erkki Somersalo, and Marko Vauhkonen, *Statistical inversion and monte carlo sampling methods in electrical impedance tomography*, *Inverse problems* **16** (2000), no. 5, 1487.
- [18] Oscar Alberto Rodríguez Melendez, *Serie estocástica de taylor-itô y métodos numéricos para ecuaciones diferenciales estocásticas*, *Boletín de Matemáticas* **23** (2016), no. 1, 81–103.
- [19] Helio S Migon, Dani Gamerman, and Francisco Louzada, *Statistical inference: an integrated approach*, CRC press, 2014.
- [20] Christian Robert and George Casella, *Introducing monte carlo methods with r*, Springer Science & Business Media, 2009.
- [21] Andrew M Stuart, *Inverse problems: a bayesian perspective*, *Acta Numerica* **19** (2010), 451–559.
- [22] Paul Allen Tipler and Gene Mosca, *Física para la ciencia y la tecnología: Electricidad y magnetismo, luz, física moderna*, vol. 2, Reverté, 2005.
- [23] Zhou Zhou, Emma Malone, Gustavo Sato dos Santos, Nan Li, Hui Xu, and David Holder, *Comparison of different quadratic regularization for electrical impedance tomography*, 6th European Conference of the International Federation for Medical and Biological Engineering, Springer, 2015, pp. 200–203.