

UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
POSTGRADO EN MATEMÁTICA



“Consistencia de un Estimador para el Oscilador Armónico
Estocástico No Lineal. Datos Totalmente Observados.”

AUTOR: Roberto Ruggiero

TUTORES: Dr. José Rafael León y Dr. Luis Ángel Rodríguez

Tesis Doctoral
Presentada ante la Ilustre
Universidad Central de Venezuela
Para Optar al Título de
Doctor en Ciencias
Mención Matemática

Caracas, 17 de julio de 2018



VEREDICTO

Quienes suscriben, miembros del jurado designado por el Consejo de la Facultad de Ciencias y el Consejo de Estudios de Postgrado de la Universidad Central de Venezuela, para examinar la **TESIS DOCTORAL** presentada por el estudiante **ROBERTO GIOVANNI RUGGIERO PALLADINO**, C.I.V-6.850.966, bajo el título: "**CONSISTENCIA DE UN ESTIMADOR PARA EL OSCILADOR ARMÓNICO ESTOCÁSTICO NO LINEAL. DATOS TOTALMENTE OBSERVADOS**", a fin de cumplir con el requisito legal para optar al grado académico de **DOCTOR EN CIENCIAS, MENCIÓN MATEMÁTICA**, dejan constancia de lo siguiente:

1.- Leído como fue dicha tesis por cada uno de los miembros del jurado, se fijó el día **martes 17 de julio de 2018** a las **2:00 p.m.**, para que el autor lo defendiera en forma pública, lo que éste hizo en la **Sala de Postgrado en Matemática de la Facultad de Ciencias, Universidad Central de Venezuela**, mediante un resumen oral de su contenido, luego de lo cual respondió satisfactoriamente a las preguntas que le fueron formuladas por el jurado, todo ello conforme con lo dispuesto en el Reglamento de Estudios de Postgrado.

2.- El jurado decidió **APROBARLO**, sin hacerse solidario con la ideas expuestas por el autor, porque se ajusta a lo dispuesto y exigido en el Reglamento de Estudios de Postgrado.

Para dar este Veredicto, el jurado estimó que el trabajo examinado reúne los requisitos de presentación, originalidad, revisión bibliográfica e impacto necesarios para constituir una **Tesis Doctoral**, ya que el autor considera la estimación paramétrica para el oscilador armónico estocástico no lineal basando su estimación en un método de aproximación lineal local y una discretización de la ecuación. A partir de esto se establece una aproximación de la función de verosimilitud la cual permite construir una función de contraste cuadrática. A partir de este contraste se demuestra la consistencia asintótica de los parámetros del modelo. La técnica utilizada necesita un control de la varianza del proceso discretizado para evitar la degeneración. Estos resultados demuestran en el autor un dominio profundo de elementos de cálculo estocástico y propiedades probabilísticas del modelo.

3.- El jurado por unanimidad decidió otorgar la calificación de **EXCELENTE** al presente trabajo por considerar que presenta contribuciones valiosas a los temas abordados que se sustentan por opiniones de investigadores reconocidos en el área.

En fe de lo cual se levanta la presente ACTA, a los **17 días del mes de julio del año 2018**, conforme a lo dispuesto en el Reglamento de Estudios de Postgrado. Actuaron como Tutores los Profesores **Dr. JOSÉ RAFAEL LEÓN (UCV-Universidad de la República-Uruguay) (Tutor) (Coordinador del Jurado)** y **Dr. LUIS RODRÍGUEZ (UC)**.

Dra. MAIRENE COLINA (UCV)

C.I.V-12.761.954

Jurado designado por Consejo Facultad

Dr. LISANDRO FERMÍN (Univ. Valparaíso, Chile)

C.I.V-13.087.352

Jurado designado por Consejo de Facultad

Dra. ZORAIDA MARTÍNEZ (USB)

C.I.V-11.165.566

Jurado designado por Consejo de Estudios de Postgrado

Dr. RICARDO RÍOS (UCV)

C.I.V-3.949.476

Jurado designado por Consejo de Estudios de Postgrado

DR. JOSÉ RAFAEL LEÓN

(UCV-Universidad de la República-Uruguay) (Tutor) (Coordinador)

C.I.V-3.846.337

Consistencia de un Estimador para el Oscilador Armónico Estocástico No Lineal. Datos Totalmente Observados.

Resumen

En este trabajo demostramos, para el oscilador armónico estocástico no-lineal, la consistencia de un estimador de los parámetros basado en el método de linealización local, considerando que los datos se observan totalmente. Se realizan simulaciones para comparar los resultados con otros autores y con el caso cuando los datos son observados parcialmente. Además, se presenta una aplicación que consiste en un oscilador estocástico lineal con datos parcialmente observados por medio del cual se simula la dinámica de la glucosa en la sangre durante una prueba oral de tolerancia a la glucosa.

Palabras Claves: Oscilador, Estimador, Consistencia, Máxima Verosimilitud.

Consistency of a Likelihood Estimator for Stochastic Damping Hamiltonian Systems. Totally Observed Data.

Abstract

In this work we prove the consistency of an estimator for a stochastic damping Hamiltonian system considering that both position and velocity are observed. Next we perform some simulations, including the case when only position is available, to see how the estimators work numerically and then compare the obtained results with those obtained by other authors. Besides, we present an application by means of a stochastic linear model, considering partially observed data, to simulate the dynamics of the glucose in the blood stream, during an oral glucose tolerance test.

Keywords: Two dimensional hypoelliptic diffusion, Stochastic damping Hamiltonian systems, likelihood estimator, consistency, discretely observed data.

Índice general

Introducción	1
1. Marco Teórico	4
1.1. El modelo principal y las hipótesis	4
1.2. Modelo Localmente Linealizado	8
2. CONSISTENCIA	13
3. SIMULACIONES	30
3.1. Descripción de las simulaciones	30
3.2. Algoritmos	32
3.3. Resultados	33
3.4. Estimación y comparación	35
4. APLICACIÓN:Modelo estocástico para la detección de Diabetes	42
4.1. Preliminares	42

4.2. Modelo Estocástico	44
4.2.1. Modelo continuo	44
4.2.2. Modelo discreto	47
5. CONCLUSIONES Y RECOMENDACIONES	51
5.1. Conclusiones	51
5.2. Recomendaciones	52
6. Apéndice	53
7. Bibliografía	71

INTRODUCCIÓN

En las últimas décadas muchos investigadores se han interesado en obtener estimadores consistentes y normalmente asintóticos con la idea de identificar un modelo estocástico el cual se ajuste a algunos datos observados. En muchos casos, los datos son observados totalmente, es decir; se observa la posición y la velocidad. En otros casos, no es posible medir ambas características y por lo tanto se dice que se observan parcialmente, es decir; sólo la posición o sólo la velocidad. En muchas aplicaciones como mecánica aleatoria (en ingeniería), neurofisiología, finanzas, química, etc. (ver por ejemplo los trabajos de Robert y Spanos [8], Lindner y Schimansky–Geier [3], Nicolau [12], Lelièvre, Rousset, Stoltz [18]), el modelo estocástico consiste en osciladores estocásticos no lineales (o sistemas de vibraciones aleatorias no lineales) que tienen la forma

$$\ddot{x}_t + g_1(x_t; a) \dot{x}_t + g_2(x_t; b) x_t = \sigma dW_t, \quad (1)$$

donde g_1 y g_2 son funciones no lineales de la posición x_t y que dependen de unos parámetros a, b . El término $g_1(x_t; a)$ en la ecuación, representa el coeficiente de la fuerza de amortiguación del sistema y el término $g_2(x_t; b) x_t$ representa la fuerza de deriva, la cual se supone que es conducida por el potencial $V(x)$. En este trabajo, dicho potencial es tal que su derivada viene dada por $V'(x) = g_2(x; b) x$. El término dW_t es un ruido blanco Gaussiano y $\sigma > 0$ representa su magnitud o varianza local. Para un estudio sobre este tipo de modelos, referimos al lector revisar el libro de Gitterman [11]. El modelo (1) puede ser representado como un sistema Hamiltoniano amortiguado estocástico mediante

$$\begin{cases} dx_t = y_t dt \\ dy_t = \sigma dW_t - [g_1(x_t; a)y_t + g_2(x_t; b)x_t] dt \end{cases} \quad (2)$$

Bajo algunas hipótesis sobre el potencial $V(x)$, si el coeficiente de amortiguación satisface $g_1(x; a) = c > 0$ y la amplitud del ruido cumple que $\sigma > 0$, se demuestra

Introducción

en el trabajo de Abbaoui y Bendjeddou [1] que existe una solución estacionaria (en el tiempo) de la ecuación (2). En el artículo de Wu [21] fueron demostradas algunas propiedades asintóticas del modelo (2) tales como la ergodicidad exponencial en la convergencia a la distribución invariante, bajo supuestos mucho más generales que los asumidos por Abbaoui y Bendjeddou, sobre el potencial $V(x)$ y la amplitud del ruido $g_1(x; a)$. Con respecto al tema de la inferencia estadística para difusiones clásicas, ésta ha sido estudiada ampliamente y en la mayoría de los casos la estimación paramétrica es llevada a cabo hallando una función contraste realizando una aproximación de Euler directamente sobre el modelo (ver por ejemplo el trabajo de Genon-Catalot, Jeantheau y Laredo [7] para más información). El sistema (2) también es considerado en la literatura como una difusión hipo-elíptica estocástica bivariada. Esta condición de hipo-elipticidad es demostrada por Malliavin en su importante trabajo [10]. En el caso de difusiones hipo-elípticas, es mucho más complicado realizar una aproximación de Euler directamente sobre el modelo debido a que la matriz de varianza-covarianza de una difusión hipo-elíptica bivariada es degenerada. No obstante, varios investigadores se dedicaron a resolver este problema: Ozaki [13] propone el método de linealización local para obtener un modelo discreto cuya matriz de varianza-covarianza no es degenerada y a partir de ese modelo discreto obtiene un estimador por máxima verosimilitud. Talay [19] utilizó un esquema de Euler implícito sobre el sistema y demuestra la convergencia exponencial hacia la medida invariante. En el trabajo de Pokern, Stuart y Wiberg [16] se enfocan en difusiones hipo-elípticas de orden superior y proponen una aproximación empírica de la verosimilitud basada en expansiones de Itô y un muestreador de Gibbs híbrido, obteniendo un estimador Bayesiano. En el artículo de Samson and Thieullen [17], se utiliza una aproximación de Euler sólo en la segunda ecuación del sistema (2) lo cual les permite construir un estimador consistente.

En el contexto de estimación no-paramétrica los trabajos son pocos. En dos recientes publicaciones de Cattiaux, León y Prieur [5] y [15], una estimación no-paramétrica de la densidad invariante y del término de deriva fueron obtenidas usando estimadores del kernel. A partir de allí logran establecer un teorema del límite central para tales estimadores, bajo observaciones parciales de los datos.

En este trabajo se demuestra la consistencia del estimador basado en la máxima verosimilitud obtenido a partir de un modelo local linealizado, para datos totalmente observados. Este trabajo está fuertemente inspirado en el artículo de Ozaki [13]. En el citado artículo, Ozaki señala que existe un sesgo que viene de estimar los parámetros usando el modelo linealizado discreto a partir de observaciones

Introducción

del modelo continuo. A partir de lo que señaló Ozaki, al estudiar el problema encontramos que dicha solución dista mucho de ser trivial. Una recompensa inesperada al estudiar estos modelos es que fuimos llevados de manera natural a proponer un modelo estocástico para simular la dinámica de la glucosa en el plasma sanguíneo, que se encuentra en completa sintonía con nuestro desarrollo.

El trabajo está organizado de la siguiente manera: en el capítulo 2 presentamos el modelo principal y estableceremos algunas notaciones y las hipótesis que usaremos a lo largo del trabajo. Finalmente, en la última sección se hace una breve descripción del método de linealización local de Ozaki, estableciendo algunos resultados importantes. En el capítulo 3 se demuestra el resultado principal de este trabajo: la consistencia, en probabilidad, del estimador considerando que los datos se observan totalmente. En el capítulo 4 se realizan las simulaciones, tanto para datos completamente observados como para datos parcialmente observados (supondremos que se observa sólo la posición). Esto con la idea de ver cómo se comporta el estimador en ambos casos y además, para comparar con los resultados obtenidos por otros autores: Ozaki [13], Samson *et al.* [17] y Pokern *et al.* [16]. En el capítulo 5 se presenta una aplicación interesante del oscilador estocástico lineal que consiste en un modelo para simular la dinámica de la glucosa en el plasma sanguíneo, durante una prueba de tolerancia a la glucosa. Inicialmente se presenta el modelo estocástico continuo y su solución analítica. Luego se obtiene el modelo estocástico discreto que aproxima al modelo continuo. A partir del modelo discreto, obtenemos un estimador basado en la máxima verosimilitud. En el futuro, a partir de los parámetros estimados se podría establecer un criterio que permita decidir si el paciente tiene diabetes tipo II. Finalmente, se presentan las conclusiones y trabajos futuros.

Capítulo 1

Marco Teórico

En este capítulo presentamos el modelo principal y algunos conceptos relacionados con el mismo. Además se establecerá tanto la notación que será utilizada a lo largo de todo el manuscrito como las hipótesis generales sobre el modelo. Finalmente, se desarrollará brevemente el modelo localmente linealizado propuesto por Ozaki [13].

1.1. El modelo principal y las hipótesis

En este trabajo utilizaremos el sistema de ecuaciones diferenciales estocásticas dado por:

$$\begin{cases} dx(t) = y(t)dt \\ dy(t) = \sigma dW_t - [g_1(x(t); a)y(t) + g_2(x(t); b)x(t)] dt \end{cases}, \quad (1.1)$$

donde $z(t) := (x(t), y(t))$, $t \geq 0$ es un vector cuyas coordenadas representan la posición $x(t)$ y la velocidad $y(t)$ de una partícula del sistema en el instante de tiempo t . Los parámetros serán denotados por a y b . Cada uno de ellos representa un vector de parámetros $a = (a_1, \dots, a_q)$ y $b = (b_1, \dots, b_p)$, los cuales aparecen en las expresiones de las funciones g_1 y g_2 respectivamente. En relación con los parámetros supondremos que

$$a \in \Theta_1 \subset \mathbb{R}^q \quad \text{y} \quad b \in \Theta_2 \subset \mathbb{R}^p,$$

donde Θ_1 y Θ_2 son conjuntos compactos. Además, denotaremos por $\theta = (a, b) \in \Theta \subset \mathbb{R}^{p+q}$ a los parámetros del modelo, $\Theta = \Theta_1 \times \Theta_2$ como un conjunto compacto que contiene a los parámetros. Como es usual, los verdaderos parámetros se denotarán por medio de $\theta_0 = (a_0, b_0) \in \Theta^\circ$ (el interior de Θ). Supondremos que la amplitud del ruido, denotada por σ , es una constante positiva. Para $z = (x, y) \in \mathbb{R}^2$ definimos la función $P : \mathbb{R}^2 \times \Theta \rightarrow \mathbb{R}$ mediante

$$P(z; \theta) := -g_1(x; a)y - g_2(x; b)x. \quad (1.2)$$

Las derivadas parciales de P serán denotadas por los vectores $\nabla_z P(z; \theta) = (P_x(z; \theta), P_y(z; \theta)) \in \mathbb{R}^2$ y $\nabla_\theta P(z; \theta) = (P_a(z; \theta), P_b(z; \theta)) \in \mathbb{R}^{p+q}$. A menos que se indique otra cosa, a lo largo de este trabajo usaremos constantes positivas genéricas \mathbf{C} o \mathbf{C}_r (en este último caso, la constante depende del subíndice), que pueden variar de una línea del manuscrito a otra.

Para el sistema (1.1), supondremos válidas las siguientes hipótesis:

- H1
1. Para todo $N > 0 : \sup_{|x| \leq N} |g_1(x; a)| < \infty$.
 2. Existen constantes positivas c, L tales que $|g_1(x; a)| \geq c$, para todo $|x| > L$.
 3. $g_1(x; a) \in \mathcal{C}^2(\mathbb{R} \times \Theta_1^\circ)$.
 4. Existe un polinomio $q_1(|x|) = \mathbf{C}_\theta |x|^{\gamma_1}$, $\mathbf{C}_\theta > 0$ tal que

$$|g_1(x; a)| \leq \mathbf{C}_\theta |x|^{\gamma_1} \quad \text{y} \quad |g_1'(x; a)| \leq \mathbf{C}_\theta |x|^{\gamma_1},$$

para todo $x \in \mathbb{R}$.

- H2
1. El potencial $V(x; b) \in \mathcal{C}^2(\mathbb{R} \times \Theta_2^\circ)$ y es acotado inferiormente. Su derivada tiene la forma $V'(x; b) = g_2(x; b)x$.
 2. Supondremos que la condición de amortiguamiento

$$+\infty \geq \liminf_{|x| \rightarrow +\infty} \frac{x^2 \cdot g_2(x; b)}{|x|} \geq v > 0,$$

vale para alguna constante v .

3. El potencial tiene crecimiento exponencial para $|x| \rightarrow \infty$. Además, existe un polinomio $q_2(|x|) = \mathbf{C}_\theta |x|^{\gamma_2}$, $\mathbf{C}_\theta > 0$ tal que

$$|V(x; b)| \leq \mathbf{C}_\theta |x|^{\gamma_2} \quad \text{y} \quad |V'(x; b)| \leq \mathbf{C}_\theta |x|^{\gamma_2},$$

para todo $x \in \mathbb{R}$.

- H3
1. $P : \mathbb{R}^2 \times \Theta \rightarrow \mathbb{R}^2$ es continua.
 2. $P(z; \theta) \in \mathcal{C}^2(\mathbb{R}^2 \times \Theta^o)$; para cada θ , existen constantes positivas \mathbf{C}_θ y γ tales que $|P(z; \theta)| \leq \mathbf{C}_\theta(1 + \|z\|)^\gamma$, para todo $z \in \mathbb{R}^2$.
 3. Para todo $\theta \neq \theta'$, $P(z; \theta) \neq P(z; \theta')$.
 4. Existe una constante positiva γ tal que

$$\sup_{\theta' \in \Theta} \|\nabla_{\theta'} P(z; \theta')\| \leq \sup_{\theta' \in \Theta} \{\mathbf{C}_{\theta'}\} \|z\|^\gamma = \mathbf{C} \|z\|^\gamma.$$

Para cada $\theta \in \Theta$, se infiere de las hipótesis H1 y H2 que las siguientes cotas son válidas

$$|P_x(z; \theta)| \leq \mathbf{C}_\theta \|z\|^{\gamma_3} \quad \text{y} \quad |P_y(z; \theta)| \leq \mathbf{C}_\theta \|z\|^{\gamma_1}, \quad \forall z \in \mathbb{R}^2. \quad (1.3)$$

De acuerdo con el trabajo de Wu [21], las hipótesis H1-H2 implican lo siguiente

- Existe una función de Foster–Lyapunov $\Psi \geq 1$ la cual satisface que $\mathcal{L}\Psi \leq -\alpha\Psi + D 1_K$, para algunas constantes $\alpha, D > 0$ y algún subconjunto compacto K . Aquí \mathcal{L} denota el generador infinitesimal dado por

$$\mathcal{L} = \frac{\sigma^2}{2} \frac{d^2}{dy^2} + y \frac{d}{dx} - (g_1(x; a)y + g_2(x; b)x) \frac{d}{dy}.$$

- Además, Ψ satisface que

$$\log(\Psi(x, y)) \geq D(|y|^2 + |x|),$$

cuando $\|z\| \rightarrow \infty$, para alguna constante adecuadamente escogida $D > 0$.

Por lo tanto se deduce que el proceso $z(t) = (x(t), y(t))$ es recurrente positivo y tiene una única medida invariante denotada por $\pi_c(z) = \pi_c(x, y)$. Note que la medida invariante $\pi_c(z(0))dz(0)$ depende de los parámetros θ que definen el modelo. Por ejemplo, en el caso de oscilador de Duffing, considerando $g_1(x; a) = g_1(a)$ constante, la expresión para la medida invariante es dada, salvo por un factor constante, por

$$\pi_c(x, y)dxdy = \exp\left(-\frac{2g_1(a)}{\sigma^2}\mathbf{H}(x, y)\right)dxdy,$$

donde $\mathbf{H}(x, y) = \frac{1}{2}|y|^2 + V(x)$ es el Hamiltoniano. Sin embargo, en general no se conoce una forma cerrada para la medida invariante tal como ocurre con el oscilador de Van der Pol, en el cual se tiene que $g_1(x; a) = a_1 x^2 - a_2$ con $a_1, a_2 > 0$, $V(x) = \frac{1}{2}b^2 x^2$.

Lo anterior también permite inferir algunas propiedades relacionadas con la independencia asintótica. En efecto, el proceso $z(t)$ es α -mixing y todos los momentos polinomiales de la medida π_c son finitos (ver Cattiaux *et al* [5]). En este trabajo supondremos que la distribución del proceso $z(0) = (x(0), y(0))$ es la medida invariante π_c .

Observación 1.1.1. *La hipo-elipticidad del modelo principal (1.1) fue demostrada a mediados de los años setenta por Malliavin [10]. A mediados del siglo pasado ya Wiener había desarrollado una teoría sobre el movimiento Browniano, Kolmogorov demostró que las probabilidades de transición de un proceso de difusión definen una solución fundamental a una ecuación de calor asociada, e Itô desarrolló un cálculo estocástico que permitía representar a una difusión, con un generador infinitesimal dado, como la solución de una ecuación diferencial estocástica. Con respecto al área de ecuaciones en derivadas parciales, Hörmander había establecido un criterio que permitía establecer si un operador diferenciable con coeficientes variables era hipo-elíptico. De manera muy general: un operador G es hipo-elíptico si, siempre que $G(u)$ es diferenciable para una distribución u definida en algún subconjunto abierto del dominio de G , entonces u es diferenciable. Estos avances permitieron a Malliavin, en el año 1976, establecer una conexión entre el análisis estocástico y las ecuaciones en derivadas parciales: Consideremos la ecuación diferencial estocástica (en el sentido de Stratonovich) dada por*

$$d\xi_t = \sum_{i=1}^n X_i(\xi_t) \circ dw_i(t) + X_0(\xi_t)dt,$$

donde X_0, \dots, X_n son campos vectoriales diferenciables sobre \mathbb{R}^d y $w = (w_1, \dots, w_n)$ es un proceso de Wiener estándar. Denotemos por L el operador diferencial de segundo orden dado por $\frac{1}{2} \sum_{i=1}^n X_i^2 + X_0$. Se sabe que la solución ξ_t de la ecuación mencionada, es un proceso de Markov y sus probabilidades de transición son soluciones fundamentales (en el sentido de distribuciones) de la ecuación de calor $\frac{\partial u}{\partial t} = Lu$. Supongamos que los campos vectoriales X_1, \dots, X_n junto con todos los corchetes de Lie que se expanden por X_0, \dots, X_n generan todo \mathbb{R}^d , en todo punto. Del teorema de hipo-elipticidad de Hörmander se deduce que las probabilidades de transición $p(t, x, dy)$ admiten densidades diferenciables $p(t, x, y)$, para todo $t > 0$. Luego, se deduce que el operador $\frac{\partial}{\partial t} - L$ es hipo-elíptico,

obteniendo así el resultado para ecuaciones diferenciales estocásticas. Para ver más detalles sugerimos el libro de Bramanti [4].

1.2. Modelo Localmente Linealizado

En esta sección presentamos de manera breve el método de linealización local desarrollado por Ozaki en [13, 14]. Además, se calcula la densidad conjunta del proceso $z(t)$.

En el trabajo de Ozaki [13] se obtiene un método de identificación estadística para modelos no lineales de vibraciones aleatorias, basados en la máxima verosimilitud para modelos con tiempo discreto. En el mencionado trabajo Ozaki considera la ecuación diferencial estocástica no lineal a tiempo continuo (1). Esta ecuación se expresa como el sistema dinámico estocástico bivariado dado por

$$\dot{z}(t) = f(z(t); \theta) + d\mathbf{W}(t), \quad (1.4)$$

donde

- $z(t) = (x(t), y(t))'$,
- $f(z(t); \theta) = (y(t), -g_1(x(t); a)y(t) - g_2(x(t); b)x(t))' = (y(t), P(z(t); \theta))'$ y
- $d\mathbf{W}(t) = (0, \sigma dW_t)'$.
- El proceso $\{z(t)\}$ es observado en instantes de tiempo igualmente espaciados $t = ih$, $i = 0, \dots, N$.

La notación A' se utiliza para la traspuesta de un vector o matriz A . La matriz de varianza-covarianza para el ruido blanco bivariado $d\mathbf{W}(t)$ viene dada por

$$\mathbb{E} [d\mathbf{W}(t) d\mathbf{W}(s)'] = \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2 \delta(t - s) \end{bmatrix},$$

donde $\delta(t)$ es la función delta de Dirac. Observamos que la matriz de varianza-covarianza es singular. Para hallar un esquema discreto que aproxime el modelo

continuo (1.4) en el intervalo de tiempo $[ih, (i+1)h]$ con $h > 0$ suficientemente pequeño, se considera en primer lugar la ecuación determinística homogénea

$$\dot{z}(t) = f(z(t); \theta). \quad (1.5)$$

Tomando derivadas se obtiene

$$\ddot{z}(t) = J(z(t); \theta)\dot{z}(t), \quad (1.6)$$

donde

$$J(z(t); \theta) = \left[\frac{\partial f(z; \theta)}{\partial z} \right]_{z=z(t)} = \begin{bmatrix} 0 & 1 \\ g'_1(x; a)y + g'_2(x; b)x + g_2(x; b) & g_1(x; a) \end{bmatrix}.$$

Debido a la importancia que tiene la matriz Jacobiana $J(z(t); \theta)$ en el desarrollo del resto de este trabajo, debemos agregar la siguiente hipótesis

H4 La matriz $\left[\frac{\partial f(z; \theta)}{\partial z} \right]$ es no singular para todo $z \in \mathbb{R}^2$ y $\theta \in \Theta^\circ$, es decir;

$$g'_1(x; a)y + g'_2(x; b)x + g_2(x; b) \neq 0,$$

para $(x, y) \in \mathbb{R}^2$ y $\theta \in \Theta^\circ$.

Ahora bien, suponiendo que f es una función lineal en el intervalo $[ih, (i+1)h]$ la matriz Jacobiana es constante en este intervalo y por lo tanto se tiene que $J(z(s); \theta) = J(z(t); \theta)$ para todo $s \in [ih, (i+1)h]$. Integrando dos veces la ecuación $\ddot{z}(s) = J(z(t); \theta)\dot{z}(s)$ resulta

$$\begin{aligned} z(t+h) &= z(t) + J(z(t); \theta)^{-1}(e^{J(z(t); \theta)h} - I)f(z(t); \theta) \\ &= [I + J(z(t); \theta)^{-1}(e^{J(z(t); \theta)h} - I)F(z(t); \theta)]z(t), \end{aligned}$$

donde

$$F(z(t); \theta) := \begin{bmatrix} 0 & 1 \\ -g_2(x(t); a) & -g_1(x(t); b) \end{bmatrix}$$

es una matriz tal que $F(z(t); \theta)z(t) = f(z(t); \theta)$. Se definen la matriz

$$A(z(t), h; \theta) = I + J^{-1}(z(t); \theta)(e^{J(z(t); \theta)h} - I)F(z(t); \theta)$$

y la matriz $L(z(t); \theta)h$ mediante la relación

$$A(z(t), h; \theta) = e^{L(z(t); \theta)h}.$$

De esta manera se obtiene la ecuación estocástica bivariada

$$\dot{z}(s) = L(z(t); \theta)z(s) + d\mathbf{W}(s). \quad (1.7)$$

Usando la fórmula de variación de parámetros, la solución de la ecuación (1.7) viene dada por

$$z(t+h) = e^{L(z(t); \theta)h} z(t) + \int_t^{t+h} e^{L(z(t); \theta)(t+h-s)} d\mathbf{W}(s). \quad (1.8)$$

Observación 1.2.1. *Se debe señalar que en el paso anterior es donde el ruido se propaga a las dos coordenadas del sistema.*

Cuando el proceso $\{z(t)\}$ es observado en los instantes de tiempo $t = ih$, $i = 1, \dots, N$ se considera que el proceso discreto $\{z(ih)\} = \{z_{ih}\}$ se define por medio de la ecuación

$$z_{(i+1)h} = A_{ih} z_{ih} + \int_{ih}^{(i+1)h} e^{L(z_{ih}; \theta)[(i+1)h-s]} d\mathbf{W}(s). \quad (1.9)$$

En este trabajo consideraremos la siguiente ecuación de observación

$$z((i+1)h) = z(ih) + \int_{ih}^{(i+1)h} f(z(s); \theta) ds + \int_{ih}^{(i+1)h} d\mathbf{W}(s) ds. \quad (1.10)$$

Al calcular la matriz de varianza-covarianza $\Sigma_{ih}(h; \theta)$ para el ruido blanco discreto $\int_{ih}^{(i+1)h} e^{L(z_{ih}; \theta)[(i+1)h-s]} d\mathbf{W}(s)$ obtenemos

$$\begin{aligned} \Sigma_{ih}(h; \theta) &= \frac{\sigma^2}{(\mu_1 - \mu_2)^2} \begin{bmatrix} \frac{1}{2\mu_1} E_1 - \frac{2}{\mu_1 + \mu_2} E_{1,2} + \frac{1}{2\mu_2} E_2 & \frac{1}{2} E_1 - E_{1,2} + \frac{1}{2} E_2 \\ \frac{1}{2} E_1 - E_{1,2} + \frac{1}{2} E_2 & \frac{\mu_1}{2} E_1 - \frac{2\mu_1\mu_2}{\mu_1 + \mu_2} E_{1,2} + \frac{\mu_2}{2} E_2 \end{bmatrix} \\ &:= \frac{\sigma^2}{(\mu_1 - \mu_2)^2} \begin{bmatrix} a'_{11}(z_{ih}, h; \theta) & a'_{12}(z_{ih}, h; \theta) \\ a'_{12}(z_{ih}, h; \theta) & a'_{22}(z_{ih}, h; \theta) \end{bmatrix} = \sigma^2 \begin{bmatrix} a_{11}(z_{ih}, h; \theta) & a_{12}(z_{ih}, h; \theta) \\ a_{12}(z_{ih}, h; \theta) & a_{22}(z_{ih}, h; \theta) \end{bmatrix}, \end{aligned} \quad (1.11)$$

donde

- $\mu_1 = \mu_1(z_{ih}, h; \theta)$ and $\mu_2 = \mu_2(z_{ih}, h; \theta)$ son los autovalores de la matriz $J(z_{ih}; \theta)$,
- $E_1 = e^{2\mu_1 h} - 1$,
- $E_2 = e^{2\mu_2 h} - 1$ y

$$\blacksquare E_{1,2} = e^{(\mu_1 + \mu_2)h} - 1.$$

Observación 1.2.2. *Observe que la matriz de varianza-covarianza del ruido blanco continuo $d\mathbf{W}(t)$ es singular mientras que la matriz de varianza-covarianza del ruido blanco discreto es simétrica y para $h > 0$ fijo tiene rango 2. Más aún, la matriz $\Sigma_{ih}(h; \theta)$ es la verdadera matriz de varianza-covarianza del ruido blanco discreto $\int_{ih}^{(i+1)h} e^{L(z_{ih}; \theta)[(i+1)h-s]} d\mathbf{W}(s)$ y no una aproximación.*

En el siguiente lema se establece un resultado relacionado con el comportamiento asintótico de los elementos de las matrices $\Sigma_{ih}(h; \theta)$ y $\Sigma_{ih}^{-1}(h; \theta)$.

Lema 1.2.1. *Consideremos las matrices $\Sigma_{ih}(h; \theta)$ y $\Sigma_{ih}^{-1}(h; \theta)$. Bajo las hipótesis H1–H3 y H5'' se cumple*

$$\frac{1}{h^3} a_{11}(z_{ih}, h; \theta) \rightarrow \frac{1}{3}, \quad \frac{1}{h} a_{22}(z_{ih}, h; \theta) \rightarrow 1 \quad y \quad \frac{1}{h^2} a_{12}(z_{ih}, h; \theta) \rightarrow \frac{1}{2}, \quad (1.12)$$

y

$$h^3 b_{11}(z_{ih}, h; \theta) \rightarrow 12, \quad h b_{22}(z_{ih}, h; \theta) \rightarrow 4 \quad y \quad h^2 b_{12}(z_{ih}, h; \theta) \rightarrow -6, \quad (1.13)$$

en probabilidad. Más aún,

$$\frac{h^3}{2} \|\Sigma_{ih}^{-1}(h; \theta)\| = O_{\mathbb{P}}(1), \quad (1.14)$$

uniformemente en θ .

La demostración de este lema se puede consultar en el apéndice.

El modelo (1.9) se escribe como

$$z_{(i+1)h} = A_{ih}(h; \theta)z_{ih} + B_{ih}(h; \theta)\Xi_{(i+1)h}, \quad (1.15)$$

donde

- $A_{ih}(h; \theta) = A(z_{ih}, h; \theta)$,
- la condición inicial es $z_{ih} = z(ih)$, es decir; el modelo comienza con una distribución inicial que no es otra que la distribución de la medida invariante del modelo continuo.

- $B_{ih}(h; \theta) = U_{ih}(h; \theta) \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix}$. Aquí $\lambda_1 = \lambda_1(z_{ih}, h; \theta)$ y $\lambda_2 = \lambda_2(z_{ih}, h; \theta)$ denotan los autovalores de la matriz $\frac{1}{\sigma^2} \Sigma_{ih}(h; \theta)$, y las columnas de la matriz $U_{ih}(h; \theta)$ son los autovectores normalizados asociados a cada autovalor.
- $\Xi_{(i+1)h}$ es un ruido Gaussiano bivariado con matriz de varianza–covarianza $\sigma^2 I$ y de media cero.

En el siguiente lema se calcula la densidad conjunta del modelo (1.15).

Lema 1.2.2. *La densidad conjunta $p(z_h, \dots, z_{hN}; \theta, \sigma^2)$ del modelo discreto (1.15) viene dada por*

$$\frac{p(z_h)}{(2\pi)^{\frac{N-1}{2}} \sigma^{2(N-1)}} \prod_{i=1}^{N-1} \frac{1}{|\det B_{(i-1)h}(h; \theta)|} \times \exp \left[-\frac{1}{2\sigma^2} [z_{ih} - A_{(i-1)h}(h; \theta)z_{(i-1)h}]' [B_{(i-1)h}^{-1}(h; \theta)]' [B_{(i-1)h}^{-1}(h; \theta)] [z_{ih} - A_{(i-1)h}(h; \theta)z_{(i-1)h}] \right].$$

La demostración de este lema se encuentra en el apéndice.

Como consecuencia del lema (1.2.2), la log–similitud del modelo (1.15) es

$$l(\theta) = \log p(z_h, \dots, z_{hN}; \theta, \sigma^2). \quad (1.16)$$

Finalmente, manteniendo únicamente los términos que dependen de θ en (1.16), definimos el contraste verosímil aproximado para el modelo discreto (1.15) como

$$M_{N,h}(z_h^{Nh}; \theta) = \frac{1}{2N} \sum_{i=0}^{N-1} \|B_{ih}^{-1}(h; \theta) [z_{(i+1)h} - A_{ih}(h; \theta)z_{ih}]\|^2. \quad (1.17)$$

Capítulo 2

CONSISTENCIA

En este capítulo demostraremos el resultado principal de este trabajo. Para ello comenzaremos por establecer un estadístico auxiliar con el cual trabajaremos en este capítulo. Luego de definir la notación, las hipótesis necesarias y dos lemas que serán muy útiles, comenzaremos por demostrar que para $h \rightarrow 0$ y $hN \rightarrow \infty$, el contraste auxiliar converge en probabilidad, uniformemente con respecto a los datos z_{hN} y a los parámetros θ , a una función determinística. Luego, siguiendo el esquema clásico de Wald se completará la demostración.

Dado que $h = h_N$, para $h \rightarrow 0$, necesitamos suponer que $hN \rightarrow \infty$ y por lo tanto $N \rightarrow \infty$.

Observación 2.0.3. *Consideraremos como contraste el logaritmo del cociente de las verosimilitudes. Este estadístico se puede escribir como*

$$M_{N,h}(z_h^{Nh}; \theta) - M_{N,h}(z_h^{Nh}; \theta_0).$$

El estimador obtenido al minimizar esta última función con respecto de θ es el mismo que se obtiene al minimizar la función $M_{N,h}(z_h^{Nh}; \theta)$ con respecto de θ .

Antes de proseguir estableceremos la siguiente notación

1. $z_h^{Nh} = (z_h, \dots, z_{Nh})$.

2. $\hat{\theta}_{N,h} = \arg \min_{\theta \in \Theta} M_{N,h}(z_h^{Nh}; \theta)$.
3. Un detalle importante del trabajo de Ozaki [13] es que los parámetros del modelo discreto (1.15) son exactamente los mismos que los del modelo continuo (1). Esto nos permite pensar en θ_0 de dos maneras equivalentes
 - $\theta_0 = \arg \min_{\theta \in \Theta} \mathbb{E}_{Q_{d,h}} [M_{N,h}(z_h^{Nh}; \theta)]$. Aquí $Q_{d,h}$ es la medida estacionaria del par $(z_{ih}, z_{(i+1)h})$ dada por $p(z_{(i+1)h} | z_{ih})\pi_d(z_{ih})$, donde $p(\cdot | \cdot)$ es la densidad del núcleo de transición (1.15) y π_d su medida invariante.
 - $\theta_0 = \arg \min_{\theta \in \Theta} \mathbb{E}_{\pi_c} [M_{N,h}(z_h^{Nh}; \theta)]$. Aquí π_c es la medida estacionaria del modelo continuo (1).
4. Para el proceso continuo (1) denotaremos por $Q_{c,h}$ la medida estacionaria del par $(z(t), z(t+h))$, la cual viene dada por $p(z(t+h) | z(t))\pi_c(z(t))$, siendo $p(\cdot | \cdot)$ la densidad del núcleo de transición del modelo continuo (1).

Con la idea de demostrar la consistencia del estimador, es necesario establecer las siguientes hipótesis : para h_0 suficientemente pequeño vale

$$H5 \quad \mathbb{E} [e^{(2h_0 \|J(z_0; \theta)\|)} \|J(z_0; \theta)\|^4 \|f(z_0; \theta)\|^2] < \infty,$$

$$H5' \quad \mathbb{E} [\sup_{\theta \in \Theta} \{e^{(2h_0 \|J(z_0; \theta)\|)} \|J(z_0; \theta)\|^4 \|f(z_0; \theta)\|^2\}] < \infty \text{ y}$$

$$H5'' \quad \mathbb{E} [\sup_{\theta \in \Theta} \{\|J(z_0; \theta)\| e^{h_0 \|J(z_0; \theta)\|}\}] < \infty.$$

Observación 2.0.4. *Si bien no nos atrevemos a afirmar que las hipótesis H5–H5'' son válidas para g_1 y g_2 en el caso general, se demuestra (ver apéndice) que para los osciladores lineales, de Kramer's, de Van der Pol y de FitzHug–Nagumo estas hipótesis se cumplen.*

La idea es demostrar que el estimador $\hat{\theta}_{N,h}$ es consistente en el siguiente sentido: si $h \rightarrow 0$ y $hN \rightarrow \infty$, entonces

$$\lim_{N \rightarrow \infty} \hat{\theta}_{N,h} \rightarrow \theta_0 \quad \text{en probabilidad.} \quad (2.1)$$

Para ello, utilizaremos una técnica similar a la usada por Wald en su trabajo [20] para demostrar la consistencia de un estimador. En primer lugar demostraremos que $\lim_{h \rightarrow 0, hN \rightarrow \infty} \frac{1}{h} [M_{N,h}(z_h^{Nh}; \theta) - M_{N,h}(z_h^{Nh}; \theta_0)]$ existe y es una función determinística. Luego demostramos (2.1). Un antecedente de nuestra prueba para

difusiones ordinarias se puede consultar en el artículo de Genon–Catalot [6].

A continuación establecemos un resultado muy importante que será utilizado en la demostración del teorema principal de este trabajo. Su demostración puede consultarse en el apéndice.

Lema 2.0.3. *Supongamos que $z(s) = (x(s), y(s))'$ satisface la ecuación (1.4). Entonces*

$$\mathbb{E} \left[|x(s) - x(s')|^{2k} \right] = O(|s - s'|^{3k}) \quad \text{and} \quad \mathbb{E} \left[|y(s) - y(s')|^{2k} \right] = O(|s - s'|^k),$$

para todo entero positivo k .

El último detalle antes de establecer el teorema principal consiste en re–escribir el contraste (4.28) de una manera más adecuada. Éste puede ser escrito como

$$\begin{aligned} M_{N,h}(z_h^{Nh}; \theta) &= \frac{1}{2N} \sum_{i=0}^{N-1} \left\| B_{ih}^{-1}(h; \theta) [z_{(i+1)h} - A_{ih}(h; \theta) z_{ih}] \right\|^2 \\ &= \frac{1}{2N} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta) (z_{(i+1)h} - A_{ih}(h; \theta) z_{ih}), z_{(i+1)h} - A_{ih}(h; \theta) z_{ih} \rangle. \end{aligned} \quad (2.2)$$

Con respecto al término $z_{(i+1)h} - A_{ih}(h; \theta) z_{ih}$ en (2.2), dado que $z_{(i+1)h}$ satisface la ecuación de observación

$$z((i+1)h) = z(ih) + \int_{ih}^{(i+1)h} f(z(s); \theta) ds + \int_{ih}^{(i+1)h} d\mathbf{W}(s) ds,$$

con los verdaderos parámetros θ_0 , se tiene

$$z_{(i+1)h} - A_{ih}(h; \theta) z_{ih} = \int_{ih}^{(i+1)h} f(z(s); \theta_0) ds + \begin{bmatrix} 0 \\ \sigma(W_{(i+1)h} - W_{ih}) \end{bmatrix} + (I - A_{ih}(h; \theta)) z_{ih}. \quad (2.3)$$

Usando la expansión de Taylor para el término $(I - A_{ih}(h; \theta))z_{ih}$, se sigue de (2.3)

$$\begin{aligned}
z_{(i+1)h} - A_{ih}(h; \theta)z_{ih} &= \int_{ih}^{(i+1)h} f(z(s); \theta_0) ds + \begin{bmatrix} 0 \\ \sigma(W_{(i+1)h} - W_{ih}) \end{bmatrix} - hf(z_{ih}; \theta) \\
&\quad - \frac{h^2}{2} J(z_{ih}, \theta) f(z_{ih}; \theta) - \sum_{k=3}^{\infty} \frac{J^{k-1}(z_{ih}, \theta)}{k!} h^{k-3} f(z_{ih}; \theta) \\
&= \int_{ih}^{(i+1)h} [f(z(s); \theta_0) - f(z_{ih}; \theta)] ds + \begin{bmatrix} 0 \\ \sigma(W_{(i+1)h} - W_{ih}) \end{bmatrix} \\
&\quad - \frac{h^2}{2} J(z_{ih}, \theta) f(z_{ih}; \theta) - \sum_{k=3}^{\infty} \frac{J^{k-1}(z_{ih}, \theta)}{k!} h^{k-3} f(z_{ih}; \theta).
\end{aligned} \tag{2.4}$$

Más aún, como

$$J(z_{ih}; \theta) = \begin{bmatrix} 0 & 1 \\ -g'_1(x_{ih}; a)y_{ih} - g'_2(x_{ih}; b)x_{ih} - g_2(x_{ih}; b) & -g_1(x_{ih}; a) \end{bmatrix} := \begin{bmatrix} 0 & 1 \\ P_x(z_{ih}; \theta) & P_y(z_{ih}; \theta) \end{bmatrix}$$

obtenemos

$$-\frac{h^2}{2} J(z_{ih}, \theta) f(z_{ih}; \theta) = \frac{h^2}{2} \begin{bmatrix} -P(z_{ih}; \theta) \\ \ell_1(z_{ih}; \theta) \end{bmatrix}, \tag{2.5}$$

donde $\ell_1(z_{ih}; \theta) = g_1(x_{ih}; \theta)P(z_{ih}; \theta) - y_{ih}P_x(z_{ih}; \theta)$. De las hipótesis H1 y H2 se infiere que

$$|\ell_1(z_{ih}; \theta)| \leq \mathbf{C}_\theta (\|z_{ih}\|^{2\gamma_1+1} + \|z_{ih}\|^{\gamma_1+\gamma_2+1} + \|z_{ih}\|^{\gamma_1+2} + \|z_{ih}\|^{\gamma_2+2}) := \mathbf{C}_\theta \|z_{ih}\|^{\gamma_4}, \tag{2.6}$$

para todo $z \in K$. Además, como los momentos polinomiales del proceso $\{z_{ih}\}$ son acotados resulta

$$\mathbb{E} [|\ell_1(z_{ih}; \theta)|] \leq \mathbf{C}_\theta \mathbb{E} [\|z_{ih}\|^{\gamma_4}] < \infty. \tag{2.7}$$

Ahora bien, recordando que $\frac{h^2}{2} = \int_{ih}^{(i+1)h} [(i+1)h - s] ds$ y considerando la ecuación de observación (1.10) obtenemos

$$\begin{aligned}
z_{(i+1)h} - A_{ih}(h; \theta)z_{ih} &= \left[\begin{aligned} &\sigma \int_{ih}^{(i+1)h} [(i+1)h - u] dW_u - \int_{ih}^{(i+1)h} [(i+1)h - u] (P(z(u); \theta_0) - P(z_{ih}; \theta)) du \\ &\int_{ih}^{(i+1)h} (P(z(s); \theta_0) - P(z_{ih}; \theta)) ds + \sigma \int_{ih}^{(i+1)h} dW_s + \frac{h^2}{2} \ell_1(z_{ih}; \theta) \end{aligned} \right] \\
&\quad - \sum_{k=3}^{\infty} \frac{J^{k-1}(z_{ih}, \theta)}{k!} h^k f(z_{ih}; \theta) \\
&:= \mathcal{W}_i(\theta) + \mathcal{K}_i(\theta) - \sum_{k=3}^{\infty} \frac{J^{k-1}(z_{ih}, \theta)}{k!} h^k f(z_{ih}; \theta),
\end{aligned} \tag{2.8}$$

donde

$$\begin{aligned} \mathcal{W}_i(\theta) &= \left[\begin{array}{l} \sigma \int_{ih}^{(i+1)h} [(i+1)h - u] dW_u - \int_{ih}^{(i+1)h} [(i+1)h - u] (P(z(u); \theta_0) - P(z_{ih}; \theta)) du \\ \int_{ih}^{(i+1)h} (P(z(s); \theta_0) - P(z_{ih}; \theta)) ds + \sigma \int_{ih}^{(i+1)h} dW_s \end{array} \right] \\ &:= \begin{bmatrix} \mathcal{W}_{i1}(\theta) \\ \mathcal{W}_{i2}(\theta) \end{bmatrix}. \end{aligned}$$

y

$$\mathcal{K}_i(\theta) = \begin{bmatrix} 0 \\ \frac{h^2}{2} \ell_1(z_{ih}; \theta) \end{bmatrix}.$$

De esta manera, el contraste (2.2) se puede escribir como

$$\begin{aligned} &M_{N,h}(z_h^{Nh}; \theta) \\ &= \frac{1}{2N} \sum_{i=0}^{N-1} \left\langle \Sigma_{ih}^{-1}(h; \theta) \left[\mathcal{W}_i(\theta) + \mathcal{K}_i(\theta) - \sum_{k=3}^{\infty} \frac{J^{k-1}(z_{ih}, \theta)}{k!} h^k f(z_{ih}; \theta) \right], \right. \\ &\quad \left. \mathcal{W}_i(\theta) + \mathcal{K}_i(\theta) - \sum_{k=3}^{\infty} \frac{J^{k-1}(z_{ih}, \theta)}{k!} h^k f(z_{ih}; \theta) \right\rangle \\ &= \frac{1}{2N} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta) (\mathcal{W}_i(\theta) + \mathcal{K}_i(\theta)), \mathcal{W}_i(\theta) + \mathcal{K}_i(\theta) \rangle \\ &\quad + \frac{1}{2N} \sum_{i=0}^{N-1} \left\langle \Sigma_{ih}^{-1}(h; \theta) \sum_{k=3}^{\infty} \frac{J^{k-1}(z_{ih}, \theta)}{k!} h^k f(z_{ih}; \theta), \sum_{k=3}^{\infty} \frac{J^{k-1}(z_{ih}, \theta)}{k!} h^k f(z_{ih}; \theta) \right\rangle \\ &\quad - \frac{1}{N} \sum_{i=0}^{N-1} \left\langle \Sigma_{ih}^{-1}(h; \theta) (\mathcal{W}_i(\theta) + \mathcal{K}_i(\theta)), \sum_{k=3}^{\infty} \frac{J^{k-1}(z_{ih}, \theta)}{k!} h^k f(z_{ih}; \theta) \right\rangle \\ &:= \mathcal{D}_1(\theta) + \mathcal{D}_2(\theta) - \mathcal{D}_3(\theta). \end{aligned} \tag{2.9}$$

A partir de la última expresión establecemos el siguiente resultado

Lema 2.0.4. *Bajo las hipótesis H1–H3 and H5 tenemos que*

$$\frac{1}{h} (\mathcal{D}_2(\theta) - \mathcal{D}_3(\theta)) = o_{\mathbb{P}}(1)$$

Demostración. Trabajando con la expresión $\mathcal{D}_2(\theta)$, tenemos

$$\begin{aligned}
|\mathcal{D}_2(\theta)| &\leq h^6 \|\Sigma_{ih}^{-1}(h; \theta)\| \frac{1}{2N} \sum_{i=0}^{N-1} \left(\sum_{k=3}^{\infty} \frac{\|J(z_{ih}; \theta)\|^{k-1} h^{k-3}}{k!} \|f(z_{ih}; \theta)\| \right)^2 \\
&\leq h^6 \|\Sigma_{ih}^{-1}(h; \theta)\| \frac{1}{2N} \sum_{i=0}^{N-1} \left(\sum_{k=3}^{\infty} \frac{\|J(z_{ih}; \theta)\|^{k-3} h^{k-3}}{k!} \|J(z_{ih}; \theta)\|^2 \|f(z_{ih}; \theta)\| \right)^2 \\
&\leq h^6 \|\Sigma_{ih}^{-1}(h; \theta)\| \frac{1}{2N} \sum_{i=0}^{N-1} e^{(2h\|J(z_{ih}; \theta)\|)} \|J(z_{ih}; \theta)\|^4 \|f(z_{ih}; \theta)\|^2. \tag{2.10}
\end{aligned}$$

De la hipótesis H5 se sigue

$$\begin{aligned}
\mathbb{E} \left[\frac{1}{N} \sum_{i=0}^{N-1} e^{(2h\|J(z_{ih}; \theta)\|)} \|J(z_{ih}; \theta)\|^4 \|f(z_{ih}; \theta)\|^2 \right] &= \mathbb{E}[e^{(2h\|J(z_0; \theta)\|)} \|J(z_0; \theta)\|^4 \|f(z_0; \theta)\|^2] \\
&\leq \mathbb{E}[e^{(2h_0\|J(z_0; \theta)\|)} \|J(z_0; \theta)\|^4 \|f(z_0; \theta)\|^2] < \infty,
\end{aligned}$$

y del lema 1.2.1 se infiere que $\frac{h^3}{2} \|\Sigma_{ih}^{-1}(h; \theta)\| = O(1)$. Estos dos hechos nos permiten deducir, usando la desigualdad de Chebyshev, que

$$\frac{1}{N} \sum_{i=0}^{N-1} e^{(2h\|J(z_{ih}; \theta)\|)} \|J(z_{ih}; \theta)\|^4 \|f(z_{ih}; \theta)\|^2 = O_{\mathbb{P}}(1),$$

es decir; el término $\mathcal{D}_2(\theta)$ es acotado en probabilidad. Más aún, podemos deducir que

$$|\mathcal{D}_2(\theta)| = O(h^3), \tag{2.11}$$

y por lo tanto

$$\frac{|\mathcal{D}_2(\theta)|}{h} = O(h^2).$$

Por otro lado, consideremos el término $\mathcal{D}_3(\theta)$. En la observación 2.0.5 más adelante, se demuestra que

$$\frac{1}{N} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta)(\mathcal{W}_i(\theta) + \mathcal{K}_i(\theta)), \mathcal{W}_i(\theta) + \mathcal{K}_i(\theta) \rangle = O_{\mathbb{P}}(1).$$

Usando la desigualdad de Cauchy-Schwarz dos veces y utilizando la conocida expresión $\langle \Sigma^{-1}w, w \rangle = \langle \Sigma^{-\frac{1}{2}}w, \Sigma^{-\frac{1}{2}}w \rangle = \|\Sigma^{-\frac{1}{2}}w\|^2$, obtenemos

$$\begin{aligned}
|\mathcal{D}_3(\theta)| &\leq \left(\frac{1}{N} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta)(\mathcal{W}_i(\theta) + \mathcal{K}_i(\theta)), \mathcal{W}_i(\theta) + \mathcal{K}_i(\theta) \rangle \right)^{\frac{1}{2}} \\
&\quad \times \left(\frac{1}{N} \sum_{i=0}^{N-1} \left\langle \Sigma_{ih}^{-1}(h; \theta) \sum_{k=3}^{\infty} \frac{J^{k-1}}{k!} h^k f(z_{ih}; \theta), \sum_{k=3}^{\infty} \frac{J^{k-1}(z_{ih}; \theta)}{k!} h^k f(z_{ih}; \theta) \right\rangle \right)^{\frac{1}{2}}.
\end{aligned}$$

Además, usando (6.13) resulta $\frac{|\mathcal{D}_3(\theta)|}{h} = O_{\mathbb{P}}(1)O(h^{\frac{1}{2}}) \rightarrow 0$. De esta manera se obtiene la siguiente aproximación:

$$M_{N,h}(z_h^{Nh}; \theta) = \frac{1}{2N} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta)(\mathcal{W}_i(\theta) + \mathcal{K}_i(\theta)), \mathcal{W}_i(\theta) + \mathcal{K}_i(\theta) \rangle + O_{\mathbb{P}}(h^{\frac{3}{2}}).$$

□

A partir de los cálculos hechos y de los lemas 2.0.4 y 2.0.3 el siguiente resultado es válido:

Teorema 2.0.1. *Sea $\theta \in \Theta$. Bajo las hipótesis H1–H5'', para $h \rightarrow 0$, $N \rightarrow \infty$ y $hN \rightarrow \infty$ se tiene*

$$\frac{1}{h} (M_{N,h}(z_h^{Nh}; \theta) - M_{N,h}(z_h^{Nh}; \theta_0)) \rightarrow \frac{13}{2} \mathbb{E}[(P(z_0, \theta) - P(z_0; \theta_0))^2] := M(\theta),$$

en probabilidad.

Demostración. En virtud del lema 2.0.4, para demostrar el teorema es suficiente estudiar la expresión

$$\frac{1}{2hN} \left(\sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta)(\mathcal{W}_i(\theta) + \mathcal{K}_i(\theta)), \mathcal{W}_i(\theta) + \mathcal{K}_i(\theta) \rangle - \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta_0)(\mathcal{W}_i(\theta_0) + \mathcal{K}_i(\theta_0)), \mathcal{W}_i(\theta_0) + \mathcal{K}_i(\theta_0) \rangle \right).$$

Para cada $\theta \in \Theta$ se desarrolla cada sumando de la primera sumatoria y obtenemos

$$\begin{aligned} \frac{1}{2hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta) \mathcal{W}_i(\theta), \mathcal{W}_i(\theta) \rangle &+ \frac{1}{2hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta) \mathcal{K}_i(\theta), \mathcal{K}_i(\theta) \rangle \\ &+ \frac{1}{hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta) \mathcal{W}_i(\theta), \mathcal{K}_i(\theta) \rangle \end{aligned} \quad (2.12)$$

Comenzaremos con el segundo término en (2.12)

$$\begin{aligned} \frac{1}{2hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta) \mathcal{K}_i(\theta), \mathcal{K}_i(\theta) \rangle &= \frac{1}{2hN} \sum_{i=0}^{N-1} \left\langle \begin{bmatrix} \frac{h^2}{2} b_{12}(h) \ell_1(z_{ih}; \theta) \\ \frac{h^2}{2} b_{22}(h) \ell_1(z_{ih}; \theta) \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{h^2}{2} \ell_1(z_{ih}; \theta) \end{bmatrix} \right\rangle \\ &= h^3 b_{22} \frac{1}{8N} \sum_{i=0}^{N-1} \ell_1^2(z_{ih}; \theta). \end{aligned} \quad (2.13)$$

Tomando esperanza se sigue

$$\frac{1}{2hN} \sum_{i=0}^{N-1} \mathbb{E} [\langle \Sigma_{ih}^{-1}(h; \theta) \mathcal{K}_i(\theta), \mathcal{K}_i(\theta) \rangle] = h^3 b_{22} \frac{1}{8N} \sum_{i=0}^{N-1} \mathbb{E} [\ell_1^2(z_{ih}; \theta)]. \quad (2.14)$$

Dado que el proceso $\{z_{ih}\}$ es estacionario, se sigue de (2.14)

$$h^3 b_{22} \frac{1}{8N} \sum_{i=0}^{N-1} \mathbb{E} [\ell_1^2(z_{ih}; \theta)] = h^2 h b_{22} \frac{1}{8} \mathbb{E}(\ell_1^2(z_0; \theta)) \rightarrow 0,$$

para $h \rightarrow 0$, en la norma L^1 y por lo tanto en probabilidad.

El tercer término en (2.12) puede ser escrito como

$$\frac{1}{hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta) \mathcal{W}_i(\theta), \mathcal{K}_i(\theta) \rangle = \frac{h}{N} \sum_{i=0}^{N-1} (b_{12} \mathcal{W}_{i1}(\theta) \ell_1(z_{ih}; \theta) + b_{22} \mathcal{W}_{i2}(\theta) \ell_1(z_{ih}; \theta)). \quad (2.15)$$

Para el primer sumando en (2.15) tenemos

$$\begin{aligned} \frac{h}{N} \sum_{i=0}^{N-1} b_{12} \mathcal{W}_{i1}(\theta) \ell_1(z_{ih}; \theta) &= \frac{h}{N} \sum_{i=0}^{N-1} b_{12} \int_{ih}^{(i+1)h} [(i+1)h - u] dW_u \ell_1(z_{ih}; \theta) \\ &\quad - \frac{h}{N} \sum_{i=0}^{N-1} b_{12} \int_{ih}^{(i+1)h} [(i+1)h - u] [P(z(u); \theta_0) - P(z_{ih}; \theta)] du \ell_1(z_{ih}; \theta). \end{aligned}$$

Tomando esperanza y usando la desigualdad de Cauchy–Schwarz, la estacionaridad del proceso, la acotación de los momentos polinomiales del proceso z_0 , los estimados obtenidos en (1.3) y (2.7) y del lema 1.2.1 obtenemos

$$\begin{aligned} \mathbb{E} \left[\left| \frac{h}{N} \sum_{i=0}^{N-1} b_{12} \mathcal{W}_{i1}(\theta) \ell_1(z_{ih}; \theta) \right| \right] &\leq \mathbf{C} \left\{ |b_{12}| h^2 [\mathbb{E} \ell_1^2(z_0; \theta)]^{\frac{1}{2}} h^{\frac{1}{2}} + |b_{12}| h^2 [\mathbb{E} P^2(z_0; \theta_0)]^{\frac{1}{2}} \right\} h \\ &= o(1). \end{aligned} \quad (2.16)$$

Además, procediendo de la misma manera que antes se obtiene

$$\begin{aligned} \mathbb{E} \left[\left| \frac{h}{N} \sum_{i=0}^{N-1} b_{22} \mathcal{W}_{i2}(\theta) \ell_1(z_{ih}; \theta) \right| \right] &\leq \mathbf{C} \left\{ |b_{22}| h [\mathbb{E} \ell_1^2(z_0; \theta)]^{\frac{1}{2}} h^{\frac{1}{4}} + |b_{22}| h [\mathbb{E} P^2(z_0; \theta_0)]^{\frac{1}{2}} \right\} h \\ &= o(1). \end{aligned} \quad (2.17)$$

De (2.16) y (2.17) concluimos

$$\mathbb{E} \left[\left| \frac{1}{hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta) \mathcal{W}_i(\theta), \mathcal{K}_i(\theta) \rangle \right| \right] = o(1). \quad (2.18)$$

De esta manera podemos escribir

$$\begin{aligned} \frac{1}{h} (M_{N,h}(z_h^{Nh}; \theta) - M_{N,h}(z_h^{Nh}; \theta_0)) &= \frac{1}{2hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta) \mathcal{W}_i(\theta), \mathcal{W}_i(\theta) \rangle \\ &\quad - \frac{1}{2hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta_0) \mathcal{W}_i(\theta_0), \mathcal{W}_i(\theta_0) \rangle \\ &\quad + o_{\mathbb{P}}(1). \end{aligned}$$

Desarrollemos el término

$$\frac{1}{2hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta) \mathcal{W}_i(\theta), \mathcal{W}_i(\theta) \rangle.$$

Tenemos que

$$\begin{aligned} \frac{1}{2hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta) \mathcal{W}_i(\theta), \mathcal{W}_i(\theta) \rangle &= \frac{b_{11}(h)}{2hN} \sum_{i=0}^{N-1} \mathcal{W}_{i1}^2(\theta) + \frac{b_{12}(h)}{hN} \sum_{i=0}^{N-1} \mathcal{W}_{i1}(\theta) \mathcal{W}_{i2}(\theta) \\ &\quad + \frac{b_{22}(h)}{2hN} \sum_{i=0}^{N-1} \mathcal{W}_{i2}^2(\theta). \end{aligned}$$

Tomando en cuenta la expresión para $\mathcal{W}_{i1}(\theta)$ y $\mathcal{W}_{i2}(\theta)$ en (2.8), denotemos por $\mathcal{G}_{jN}(\theta)$, $j = 1, \dots, 10$ a cada término que se obtiene en el desarrollo de la expresión

anterior. Así

$$\begin{aligned}
\mathcal{G}_{1N}(\theta) &= \frac{b_{11}(h)}{2hN} \sum_{i=0}^{N-1} \left(\int_{ih}^{(i+1)h} [(i+1)h - u][P(z(u); \theta_0) - P(z_{ih}; \theta)] du \right)^2. \\
\mathcal{G}_{2N}(\theta) &= -\frac{\sigma b_{11}(h)}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} \int_{ih}^{(i+1)h} [(i+1)h - u_1][(i+1)h - u_2][P(z(u_2); \theta_0) - P(z_{ih}; \theta)] dW_{u_1} du_2. \\
\mathcal{G}_{3N}(\theta) &= \frac{\sigma b_{12}(h)}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} [(i+1)h - u] dW_u \int_{ih}^{(i+1)h} [P(z(s); \theta_0) - P(z_{ih}; \theta)] ds. \\
\mathcal{G}_{4N}(\theta) &= -\frac{b_{12}(h)}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} [P(z(s); \theta_0) - P(z_{ih}; \theta)] ds \int_{ih}^{(i+1)h} [(i+1)h - u][P(z(u); \theta_0) - P(z_{ih}; \theta)] du. \\
\mathcal{G}_{5N}(\theta) &= -\frac{b_{12}(h)}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} dW_s \int_{ih}^{(i+1)h} [(i+1)h - u][P(z(u); \theta_0) - P(z_{ih}; \theta)] du. \\
\mathcal{G}_{6N}(\theta) &= \frac{b_{22}(h)}{2hN} \sum_{i=0}^{N-1} \left(\int_{ih}^{(i+1)h} [P(z(s); \theta_0) - P(z_{ih}; \theta)] ds \right)^2. \\
\mathcal{G}_{7N}(\theta) &= \frac{b_{22}(h)}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} dW_s \int_{ih}^{(i+1)h} [P(z(s); \theta_0) - P(z_{ih}; \theta)] ds. \\
\mathcal{G}_{8N}(\theta) &= \frac{\sigma^2 b_{11}(h)}{2hN} \sum_{i=0}^{N-1} \left(\int_{ih}^{(i+1)h} [(i+1)h - u] dW_u \right)^2. \\
\mathcal{G}_{9N}(\theta) &= \frac{\sigma b_{12}(h)}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} [(i+1)h - u] dW_u \int_{ih}^{(i+1)h} dW_s. \\
\mathcal{G}_{10N}(\theta) &= \frac{b_{22}(h)}{hN} \sum_{i=0}^{N-1} \left(\int_{ih}^{(i+1)h} dW_s \right)^2.
\end{aligned}$$

Es directo darse cuenta que $\frac{1}{h} (M_{N,h}(z_h^{Nh}; \theta) - M_{N,h}(z_h^{Nh}; \theta_0)) = \sum_{j=1}^7 \mathcal{G}_{jN}(\theta) - \sum_{j=1}^7 \mathcal{G}_{jN}(\theta_0) + o_{\mathbb{P}}(1)$, ya que los términos que dependen sólo del ruido se cancelan al hacer la diferencia, por lo tanto consideraremos únicamente los restantes términos. Comencemos con el término \mathcal{G}_{1N} .

$$\begin{aligned}
\mathcal{G}_{1N}(\theta) &= \frac{h^3 b_{11}(h)}{8N} \sum_{i=0}^{N-1} [P(z_{ih}; \theta_0) - P(z_{ih}; \theta)]^2 \\
&\quad + \frac{b_{11}(h)}{2hN} \sum_{i=0}^{N-1} \left(\int_{ih}^{(i+1)h} [(i+1)h - u] \{P_x(\tilde{z}(u); \theta_0)[x(u) - x_{ih}] + P_y(\tilde{z}(u); \theta_0)[y(u) - y_{ih}]\} du \right)^2 \\
&\quad + \frac{h^2 b_{11}(h)}{2hN} \sum_{i=0}^{N-1} [P(z_{ih}; \theta_0) - P(z_{ih}; \theta)] \\
&\quad \times \int_{ih}^{(i+1)h} [(i+1)h - u] \{P_x(\tilde{z}(u); \theta_0)[x(u) - x_{ih}] + P_y(\tilde{z}(u); \theta_0)[y(u) - y_{ih}]\} du \\
&= \mathcal{G}_{1N,1}(\theta) + \mathcal{G}_{1N,2}(\theta) + \mathcal{G}_{1N,3}(\theta),
\end{aligned}$$

donde $\tilde{z}(u) = z_{ih} + \lambda(z(u) - z_{ih})$ para algún $0 \leq \lambda \leq 1$. Para el primer término, usando el Teorema Ergódico, el comportamiento asintótico de $b_{11}(h)$ (lema 1.2.1) y la estacionaridad del proceso $\{z_{ih}\}$ tenemos

$$\begin{aligned} \mathcal{G}_{1N,1}(\theta) &= \frac{b_{11}(h)}{2hN} \sum_{i=0}^{N-1} [P(z_{ih}; \theta_0) - P(z_{ih}; \theta)]^2 \left(\int_{ih}^{(i+1)h} [(i+1)h - u] du \right)^2 \quad (2.19) \\ &= \frac{h^3 b_{11}(h)}{8N} \sum_{i=0}^{N-1} [P(z_{ih}; \theta_0) - P(z_{ih}; \theta)]^2 \rightarrow \frac{3}{2} E \{ [P(z_0; \theta_0) - P(z_0; \theta)]^2 \}, \end{aligned}$$

para $h \rightarrow 0$ y $N \rightarrow \infty$. Para el segundo término, usando $(a+b)^2 \leq 2(a^2+b^2)$ y la desigualdad de Cauchy–Schwarz obtenemos

$$\begin{aligned} \mathcal{G}_{1N,2}(\theta) &= \frac{b_{11}(h)}{2hN} \sum_{i=0}^{N-1} \left(\int_{ih}^{(i+1)h} [(i+1)h - u] \{ P_x(\tilde{z}(u); \theta_0)[x(u) - x_{ih}] + P_y(\tilde{z}(u); \theta_0)[y(u) - y_{ih}] \} du \right)^2 \\ &\leq \frac{\mathbf{C}_\theta h^3 b_{11}(h)}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} (P_x^2(\tilde{z}(u); \theta_0)[x(u) - x_{ih}]^2 + P_y^2(\tilde{z}(u); \theta_0)[y(u) - y_{ih}]^2) du. \end{aligned}$$

Recordando las cotas superiores para P_x y P_y señaladas en (1.3), se sigue

$$\begin{aligned} \mathcal{G}_{1N,2}(\theta) &\leq \frac{\mathbf{C}_\theta h^3 b_{11}(h)}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} \{ \|\tilde{z}(u)\|^{2\gamma_3} [x(u) - x_{ih}]^2 \\ &\quad + \|\tilde{z}(u)\|^{2\gamma_1} [y(u) - y_{ih}]^2 \} du. \end{aligned}$$

Como $\tilde{z}(u) = \lambda z(u) + (1-\lambda)z_{ih}$, entonces $\|\tilde{z}(u)\| = \|\lambda z(u) + (1-\lambda)z_{ih}\|$. Definimos el polinomio

$$Q(\|z(u)\|, \|z_{ih}\|) = \|\tilde{z}(u)\|^{2\gamma_3} + \|\tilde{z}(u)\|^{2\gamma_1}.$$

Además, como $|x(u) - x_{ih}|^2 \leq \|z(u) - z_{ih}\|^2$ y $|y(u) - y_{ih}|^2 \leq \|z(u) - z_{ih}\|^2$, resulta

$$|\mathcal{G}_{1N,2}(\theta)| \leq \frac{\mathbf{C}_\theta h^3 |b_{11}(h)|}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} |Q(\|z(u)\|, \|z_{ih}\|)| \|z(u) - z_{ih}\|^2 du.$$

Tomando esperanza y usando la desigualdad de Cauchy–Schwarz obtenemos

$$\begin{aligned} \mathbb{E} [|\mathcal{G}_{1N,2}(\theta)|] &\leq \frac{\mathbf{C}_\theta h^3 |b_{11}(h)|}{hN} \sum_{i=0}^{N-1} \left\{ \int_{ih}^{(i+1)h} \mathbb{E} (|Q(\|z(u)\|, \|z_{ih}\|)|)^2 du \right\}^{\frac{1}{2}} \\ &\quad \times \left\{ \int_{ih}^{(i+1)h} \mathbb{E} \|z(u) - z_{ih}\|^2 du \right\}^{\frac{1}{2}}. \end{aligned}$$

Finalmente, dado que los momentos polinomiales de los procesos $z(u)$ y z_{ih} son finitos, usando la cota $\mathbb{E} \|z(u) - z_{ih}\|^2 = O(h)$, la estacionaridad del proceso y las relaciones señaladas en el lema 1.2.1, concluimos que

$$\mathbb{E} [|\mathcal{G}_{1N,2}(\theta)|] \leq \mathbf{C}_\theta h^3 |b_{11}(h)| \left[\mathbb{E} \left(\left| Q'(\|z_0\|) \right| \right)^2 \right]^{\frac{1}{2}} h^{\frac{1}{2}} \rightarrow 0, \quad (2.20)$$

para $h \rightarrow 0$. Para el término $\mathcal{G}_{1N,3}(\theta)$, la convergencia a cero se infiere de los hechos que $\mathcal{G}_{1N,1}(\theta)$ es acotado y $\mathcal{G}_{1N,2}(\theta)$ converge a cero en probabilidad, es decir;

$$\mathcal{G}_{1N,3}(\theta) \rightarrow 0, \quad (2.21)$$

en probabilidad. En vista de los resultados en (2.19), (2.20) y (2.21) se tiene que, para $h \rightarrow 0$ y $N \rightarrow \infty$

$$\mathcal{G}_{1N} \rightarrow \frac{3}{2} E \{ [P(z_0; \theta_0) - P(z_0; \theta)]^2 \}, \quad (2.22)$$

en probabilidad.

Ahora estudiamos el término \mathcal{G}_{2N} . Definimos una aproximación de \mathcal{G}_{2N} mediante

$$\tilde{\mathcal{G}}_{2N,1} = -\frac{\sigma b_{11} h}{2N} \sum_{i=0}^{N-1} [P(z_{ih}; \theta_0) - P(z_{ih}; \theta)] \int_{ih}^{(i+1)h} ((i+1)h - u_1) dW_{u_1}.$$

Condicionando respecto de la σ -álgebra $\mathfrak{F}_{u_1} := \sigma(z_t : t \leq u_1)$ es fácil demostrar que $\mathbb{E}[\tilde{\mathcal{G}}_{2N}] = 0$. Más aún,

$$\mathbb{E}[\tilde{\mathcal{G}}_{2N,1}^2] = b_{11}^2 h^6 \frac{\sigma^2}{8hN} \mathbb{E}[(P(z_0; \theta_0) - P(z_0; \theta))^2] \rightarrow 0.$$

La última afirmación viene del hecho que $hN \rightarrow \infty$. Con la idea de estudiar la variable aleatoria original \mathcal{G}_{2N} , consideremos la aproximación de primer orden

$$P(z(u_2); \theta_0) = P(z_{ih}; \theta_0) + P_x(\tilde{z}(u_2); \theta_0)[x(u_2) - x_{ih}] + P_y(\tilde{z}(u_2); \theta_0)[y(u_2) - y_{ih}],$$

donde $\tilde{z}(u_2) = z_{ih} + \lambda_1(z(u_2) - z_{ih})$ para algún $0 \leq \lambda_1 \leq 1$. De esta manera se tiene

$$\mathcal{G}_{2N} = \tilde{\mathcal{G}}_{2N,1} + \tilde{\mathcal{G}}_{2N,2} + \tilde{\mathcal{G}}_{2N,3},$$

donde

$$\tilde{\mathcal{G}}_{2N,2} = -b_{11} \frac{\sigma}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} \int_{ih}^{(i+1)h} ((i+1)h - u_1)((i+1)h - u_2) P_x(\tilde{z}(u_2); \theta_0)(x(u_2) - x_{ih}) dW_{u_1} du_2,$$

y

$$\tilde{\mathcal{G}}_{2N,3} = b_{11} \frac{\sigma}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} \int_{ih}^{(i+1)h} ((i+1)h-u_1)((i+1)h-u_2) g_1(\tilde{x}(u_2); \theta_0) (y(u_2) - y_{ih}) dW_{u_1} du_2.$$

Usando el lema 2.0.3, las cotas polinomiales para g_1, g_2 y sus derivadas, se demuestra que $\mathbb{E}|\tilde{\mathcal{G}}_{2N,2}| = O(h)$. Para el término $\tilde{\mathcal{G}}_{2N,3}$, si definimos

$$\tilde{\mathcal{G}}_{2N,4} = b_{11} \frac{\sigma}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} \int_{ih}^{(i+1)h} ((i+1)h-u_1)((i+1)h-u_2) g_1(x_{ih}; \theta_0) (y(u_2) - y_{ih}) dW_{u_1} du_2,$$

el mismo procedimiento usado para el término $\tilde{\mathcal{G}}_{2N,2}$ nos permite obtener $\mathbb{E}|\tilde{\mathcal{G}}_{2N,3} - \tilde{\mathcal{G}}_{2N,4}| = o_{\mathbb{P}}(1)$. De la definición arriba señalada se deduce que

$$\tilde{\mathcal{G}}_{2N,4} = b_{11} \frac{\sigma}{hN} \sum_{i=0}^{N-1} g_1(x_{ih}; \theta_0) \int_{ih}^{(i+1)h} \int_{ih}^{(i+1)h} ((i+1)h-u_1)((i+1)h-u_2) (y(u_2) - y_{ih}) dW_{u_1} du_2 + o_{\mathbb{P}}(1).$$

Tomando esperanza se obtiene

$$\mathbb{E}[\tilde{\mathcal{G}}_{2N,4}] \rightarrow 3\sigma \mathbb{E}[g_1(x_0; \theta_0)].$$

Además es fácil demostrar que la varianza tiende a cero. De esta manera

$$\tilde{\mathcal{G}}_{2N,4} \rightarrow 3\sigma \mathbb{E}[g_1(x_0; \theta_0)],$$

en probabilidad. En vista de los resultados obtenidos se tiene que

$$\mathcal{G}_{2N} - 3\sigma \mathbb{E}[g_1(x_0; \theta_0)] = \tilde{\mathcal{G}}_{2N,1} + \tilde{\mathcal{G}}_{2N,2} + (\tilde{\mathcal{G}}_{2N,3} - \tilde{\mathcal{G}}_{2N,4}) + \tilde{\mathcal{G}}_{2N,4} - 3\sigma \mathbb{E}[g_1(x_0; \theta_0)],$$

y cada término tiende a cero en probabilidad. Se concluye que para $h \rightarrow 0$ y $hN \rightarrow \infty$

$$\mathcal{G}_{2N} \rightarrow 3\sigma \mathbb{E}[g_1(x_0; \theta_0)], \quad (2.23)$$

en probabilidad.

En el estudio de los términos restantes se utilizarán procedimientos muy parecidos a los usados hasta ahora, por lo tanto no seremos excesivamente rigurosos con la escritura. Trabajemos el término \mathcal{G}_{3N} . Asintóticamente este término es equivalente a

$$\tilde{\mathcal{G}}_{3N} = -\frac{\sigma^2 b_{12}(h)}{hN} \sum_{i=0}^{N-1} g_1(x_{ih}; \theta_0) \int_{ih}^{(i+1)h} [(i+1)h-u] dW_u \int_{ih}^{(i+1)h} [(i+1)h-s] dW_s.$$

Del lema 1.2.1 tenemos

$$\mathbb{E}[\tilde{\mathcal{G}}_{3N}] = -\frac{\sigma^2 h^2 b_{12}(h)}{3} \mathbb{E}[g_1(x_0; \theta_0)] \rightarrow 2\sigma^2 \mathbb{E}[g_1(x_0; \theta_0)].$$

y su varianza tiende a cero. Entonces, para $h \rightarrow 0$ y $hN \rightarrow \infty$ se tiene

$$\mathcal{G}_{3N} \rightarrow 2\sigma \mathbb{E}[g_1(x_0; \theta_0)], \quad (2.24)$$

en probabilidad.

Para el término \mathcal{G}_{4N} , trabajando de manera similar al término \mathcal{G}_{1N} obtenemos, para $h \rightarrow 0$ y $hN \rightarrow \infty$

$$\mathcal{G}_{4N} \rightarrow 3\mathbb{E}[(P(z_0; \theta_0) - P(z_0; \theta))^2], \quad (2.25)$$

en probabilidad.

Para el término

$$\mathcal{G}_{5N} = -\frac{b_{12}(h)}{hN} \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} dW_s \int_{ih}^{(i+1)h} [(i+1)h - u][P(z(u); \theta_0) - P(z_{ih}; \theta)] du,$$

usando la aproximación

$$\tilde{\mathcal{G}}_{5N} = -\frac{b_{12}(h)}{hN} \sum_{i=0}^{N-1} g_1(x_{ih}; \theta_0) \int_{ih}^{(i+1)h} [(i+1)h - u] dW_u \int_{ih}^{(i+1)h} [(i+1)h - s] dW_s,$$

se infiere, como antes, que para $h \rightarrow 0$ y $hN \rightarrow \infty$

$$\mathcal{G}_{5N} \rightarrow -2\sigma^2 \mathbb{E}[g_1(x_0; \theta_0)], \quad (2.26)$$

en probabilidad.

Con respecto al término \mathcal{G}_{6N} , con el mismo tratamiento que le dimos al término \mathcal{G}_{1N} deducimos que para $h \rightarrow 0$ y $hN \rightarrow \infty$

$$\mathcal{G}_{6N} \rightarrow 2\mathbb{E}[(P(z_0; \theta_0) - P(z_0; \theta))^2], \quad (2.27)$$

en probabilidad.

El término \mathcal{G}_{7N} produce el mismo límite que

$$\tilde{\mathcal{G}}_{7N} = -\frac{\sigma^2 b_{22}(h)}{hN} \sum_{i=0}^{N-1} g_1(x_{ih}; \theta_0) \int_{ih}^{(i+1)h} dW_{s_1} \int_{ih}^{(i+1)h} [(i+1)h - s_2] dW_{s_2}.$$

De esta manera, procediendo de manera similar al término \mathcal{G}_{3N} obtenemos, para $h \rightarrow 0$ y $hN \rightarrow \infty$

$$\mathcal{G}_{7N} \rightarrow -2\sigma^2 \mathbb{E}[g_1(x_0; \theta_0)], \quad (2.28)$$

en probabilidad.

Finalmente, sumando los resultados obtenidos en (2.22)–(2.28) se concluye que

$$\frac{1}{h}(M_{N,h}(z_h^{Nh}; \theta) - M_{N,h}(z_h^{Nh}; \theta_0)) \rightarrow \frac{13}{2} \mathbb{E}[(P(z_0, \theta) - P(z_0; \theta_0))^2],$$

en probabilidad. \square

Observación 2.0.5. *El procedimiento anterior nos brinda un bono extra: para todo $\theta \in \Theta$ vale*

$$\frac{1}{N} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(h; \theta)(\mathcal{W}_i(\theta) + \mathcal{K}_i(\theta)), \mathcal{W}_i(\theta) + \mathcal{K}_i(\theta) \rangle = O_{\mathbb{P}}(1).$$

En efecto, los únicos términos cuyo límite no es cero, son aquellos que dependen del ruido tal y como ocurre, por ejemplo, con el término

$$b_{11} \frac{\sigma^2}{2N} \sum_{i=0}^{N-1} \left(\int_{ih}^{(i+1)h} ((i+1)h - u) dW_u \right)^2 \rightarrow 2\sigma^2,$$

usando la Ley de los Grandes Números. Por lo tanto estos términos pueden ser tratados de manera similar.

Observación 2.0.6. *Notemos que $M(\theta) = \frac{13}{2} \mathbb{E}[(P(z_0; \theta) - P(z_0; \theta_0))^2] \geq 0$ para todo $\theta \in \Theta$ y $M(\theta_0) = 0$. Así, la función $M(\theta)$ tiene un mínimo global en el punto $\theta = \theta_0$ cuyo valor es 0. Supondremos que existe una vecindad U_{θ_0} de θ_0 tal que $U_{\theta_0} \subset \Theta$ y $\nabla_{\theta} P(z_0; \check{\theta}) \neq (0, \dots, 0)$ para todo $\check{\theta} \in U_{\theta_0}$, $\check{\theta} \neq \theta_0$. Esta condición garantiza que θ_0 es el único punto de U_{θ_0} donde M alcanza su mínimo. En los cuatro modelos mencionados en la observación 2.0.4, esta condición se cumple (ver apéndice). Esta condición de identificabilidad será crucial en el próximo teorema.*

En el próximo teorema demostraremos la consistencia del estimador $\hat{\theta}_{N,h}$.

Teorema 2.0.2. *Sea $\theta \in \Theta$, donde Θ es un conjunto compacto. Sea $\{h_N\}_{N \geq 1}$ una sucesión tal que $h_N \rightarrow 0$ y $h_N N \rightarrow \infty$ cuando $N \rightarrow \infty$. Bajo las hipótesis del teorema 2.0.1 se tiene que*

$$\lim_{N \rightarrow \infty} \hat{\theta}_{N,h_N} = \theta_0,$$

en probabilidad.

Antes de comenzar la demostración debemos establecer un resultado que será necesario. Un bosquejo de la prueba puede consultarse en el apéndice.

Lema 2.0.5. *Sea $\Theta \subset \mathbb{R}^{p+q}$ un conjunto compacto. Supongamos que las hipótesis H1–H5 se cumplen. En adición, supongamos que*

$$\mathbb{E}[\sup_{\theta \in \Theta} \ell_1^2(z_{ih}; \theta)] < \infty.$$

Entonces, para h fijo, la sucesión $\left\{ \frac{1}{h_N} \left[M_{N,h_N}(z_{h_N}^{Nh_N}; \theta) \right] \right\}$ converge uniformemente con respecto a θ , en probabilidad.

Observación 2.0.7. *La hipótesis adicional $\mathbb{E}[\sup_{\theta \in \Theta} \ell_1^2(z_{ih}; \theta)] < \infty$ es válida si las constantes C_θ en las ecuaciones (2.6) y (2.7) se pueden acotar uniformemente con respecto de $\theta \in \Theta$ por una constante C . En los cuatro modelos mencionados anteriormente esta hipótesis es válida.*

Demostración del teorema 2.0.2. Definimos

$$\tilde{M}_N(\theta) = \frac{1}{h_N} \left[M_{N,h_N}(z_{h_N}^{Nh_N}; \theta) - M_{N,h_N}(z_{h_N}^{Nh_N}; \theta_0) \right].$$

Note que $\tilde{M}_N(\theta)$ y $M_{N,h_N}(\theta)$ alcanzan el mínimo en el mismo valor $\hat{\theta}_{N,h_N}$. Tenemos que

$$0 \leq M(\hat{\theta}_{N,h_N}) = M(\hat{\theta}_{N,h_N}) - \tilde{M}_N(\hat{\theta}_{N,h_N}) + \tilde{M}_N(\hat{\theta}_{N,h_N}).$$

Como $\hat{\theta}_{N,h_N} = \arg \min_{\theta \in \Theta} \tilde{M}_N(\theta)$, entonces

$$\tilde{M}_N(\hat{\theta}_{N,h_N}) \leq \tilde{M}_N(\theta_0).$$

Esto implica que

$$0 \leq M(\hat{\theta}_{N,h_N}) \leq M(\hat{\theta}_{N,h_N}) - \tilde{M}_N(\hat{\theta}_{N,h_N}) + \tilde{M}_N(\theta_0).$$

De la hipótesis de compacidad, el teorema 2.0.1 y el lema (2.0.5), tomando valor absoluto resulta

$$\begin{aligned} 0 &\leq |M(\hat{\theta}_{N,h_N})| \leq |M(\hat{\theta}_{N,h_N}) - M_N(\hat{\theta}_{N,h_N})| + |M_N(\theta_0)| \\ &\leq 2 \sup_{\theta \in \Theta} |M(\theta) - M_N(\theta)| \rightarrow 0, \end{aligned}$$

en probabilidad. Por lo tanto

$$\lim_{N \rightarrow \infty} M(\hat{\theta}_{N,h_N}) = 0 (= M(\theta_0)) \text{ en probabilidad.} \quad (2.29)$$

Sea θ_* un punto límite de la sucesión $\{\hat{\theta}_{N,h_N}\}$. Dado que Θ es compacto, existe una subsucesión $\{\hat{\theta}_{N_j,h_{N_j}}\} \subseteq \{\hat{\theta}_{N,h_N}\}$ tal que

$$\lim_{j \rightarrow \infty} \hat{\theta}_{N_j,h_{N_j}} = \theta_*.$$

Tomando en cuenta que $M(\theta)$ es continua y por (2.29), se deduce que

$$\lim_{j \rightarrow \infty} M(\hat{\theta}_{N_j,h_{N_j}}) = M(\theta_*).$$

En virtud de la suposición sobre unicidad de θ_0 que hicimos en la observación 2.0.6 obtenemos $\theta_* = \theta_0$ lo cual nos permite concluir que

$$\lim_{N \rightarrow \infty} \hat{\theta}_{N,h_N} = \theta_0,$$

en probabilidad. □

Observación 2.0.8. *Acabamos de demostrar la consistencia del estimador en el caso que los datos son completamente observados. Pensamos que un resultado similar puede obtenerse en el caso en que únicamente se observa la posición, aproximando la velocidad con un esquema de Euler. Definiendo el contraste*

$$M_{N,h}(x_1, \dots, x_N, \theta) = \frac{1}{2(N-1)} \sum_{i=2}^N \left\| B \left(\begin{bmatrix} x(i) \\ \frac{x(i)-x(i-1)}{h} \end{bmatrix}; \theta \right)^{-1} \begin{bmatrix} x(i+1) \\ \frac{x(i+1)-x(i)}{h} \end{bmatrix} - A \left(\begin{bmatrix} x(i) \\ \frac{x(i)-x(i-1)}{h} \end{bmatrix}; \theta \right) \begin{bmatrix} x(i) \\ \frac{x(i)-x(i-1)}{h} \end{bmatrix} \right\|^2,$$

y el estimador

$$\hat{\theta}_{N,h} := \arg \min_{\theta \in \Theta} M_{N,h}(x_1, \dots, x_N, \theta),$$

creemos posible demostrar que $\hat{\theta}_{N,h} \rightarrow \theta_0$ cuando $h \rightarrow 0$ y $Nh \rightarrow \infty$ en algún sentido. Si bien no es el objetivo de este trabajo desarrollar las demostraciones de la afirmación que hacemos, consideramos importante ejecutar simulaciones con este enfoque para comparar con el estimador $\hat{\theta}$ y las estimaciones obtenidas por otros autores. En el próximo capítulo se muestran los resultados.

Capítulo 3

SIMULACIONES

En este capítulo ejecutaremos algunas simulaciones que nos permitirán observar cómo se comporta numéricamente el estimador obtenido por Ozaki, en el caso en que los datos son observados completamente. También llevaremos a cabo algunas simulaciones que nos darán cuenta del comportamiento numérico del estimador en el caso que sólo se observa la posición. Luego compararemos los resultados obtenidos y los contrastaremos con los resultados obtenidos por Ozaki [13], Pokern *et al.* [16] and Samson *et al.* [17].

3.1. Descripción de las simulaciones

Consideraremos los siguientes dos modelos: el oscilador estocástico de Van der Pol

$$\ddot{x} + (ax^2 + bx + c)\dot{x} + dx = \sigma dW_t, \quad (3.1)$$

y el oscilador estocástico de Duffin

$$\ddot{x} + d\dot{x} + (ax^2 + bx + c)x = \sigma dW_t. \quad (3.2)$$

Aquí a, b, c y d son los parámetros a ser estimados. Como la varianza σ^2 también será estimada, los vectores de parámetros a ser estimados tendrán la forma $\theta = (a, b, c, d, \sigma^2)$. Con la idea de obtener datos que nos permitan hacer la estimación, usaremos los siguientes vectores de parámetros, los cuales son los parámetros reales y por lo tanto se denotan como θ_0

Van der Pol model	Duffin model
$\theta_0^{vdP1} = (1, 0, -1, 14, 8, 1)$	$\theta_0^{D1} = (4, 0, 14, 8, 0, 25, 0, 25)$
$\theta_0^{vdP2} = (2, 0, 2, 10, 1)$	$\theta_0^{D2} = (1, 0, -1, 1, 0, 25)$
$\theta_0^{vdP3} = (0, 0, 0, 5, 4, 1)$	$\theta_0^{D3} = (4, 3, -4, 1, 0, 25)$

La elección de estos parámetros obedece, entre otras cosas, a los siguientes criterios:

- 1.- En general todos los parámetros se escogieron con la idea de obtener funciones $g_1(x(t); a)$ y $g_2(x(t); a)$ que satisfagan las hipótesis H1–H5.
- 2.- Para comparar nuestros resultados con los obtenidos por Ozaki en [13], se escogieron θ_0^{vdP1} y θ_0^{D1} igual que en el mencionado trabajo, excepto la varianza en el modelo de Duffin: en este trabajo consideramos $\sigma^2 = 0,5$ mientras que Ozaki consideró $\sigma^2 = 4$. El vector de parámetros θ_{D1} produce el potencial de un sólo pozo $V(x) = x^4 + 7,4x$, esto es, un potencial que tiene un único mínimo.
- 3.- El vector de parámetros θ_0^{D2} produce el llamado oscilador de Kramer. En este caso el potencial obtenido $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$ es llamado un potencial de doble pozo simétrico, es decir, un potencial con un único mínimo, alcanzado en dos puntos. Este tipo de potenciales son considerados en la literatura como la superposición de dos potenciales de pozo simple, que tienen el mismo mínimo.
- 4.- El vector de parámetros θ_0^{D3} produce el potencial de dos pozos no simétrico dado por $V(x) = x^4 + \frac{3x^3}{2} - 2x^2$.
- 5.- El vector de parámetros θ_0^{vdP3} produce el mismo modelo que aparece en los trabajos de Pokern *et al.* [16] y Samson *et al.* [17].

Una vez establecidos los parámetros reales procedemos a obtener, para cada vector de parámetros reales, un conjunto de datos que tienen la forma

$$z(t_i) = \begin{bmatrix} x(t_i) \\ y(t_i) \end{bmatrix}_{i=1, \dots, N}. \quad \text{Aquí } x(t_i) \text{ y } y(t_i) \text{ representan la posición y la velocidad,}$$

respectivamente, de una partícula en el instante t_i en que son observados. Para efectos de la estimación consideraremos varios tamaños del conjunto de datos; $N = 100, 1000, 5000$ y 10000 . Dichos conjuntos de datos se obtendrán a partir del

modelo discreto (1.15) asociado a cada modelo continuo (Van der Pol o Duffin) usando para el primero de los modelos mencionados los parámetros reales cuyo supra-índice es vdP y para el segundo los parámetros reales cuyo supra-índice es D .

3.2. Algoritmos

Para obtener los conjuntos de datos se procede de la siguiente manera:

1. Calculamos la matriz $A(z(t); \theta_0) = I + J^{-1}(z(t); \theta_0)[e^{J(z(t); \theta_0)} - I]F(z(t); \theta_0)$. En la implementación de nuestro algoritmo, en lugar de la matriz $e^{J(z(t); \theta_0)h}$ se usará su aproximación de Taylor de segundo orden dada por $I + J(z(t); \theta_0)h + \frac{(J(z(t); \theta_0)h)^2}{2}$.
2. Calculamos la matriz $B(z(t); \theta_0)$. En nuestro algoritmo esta matriz es reemplazada por uno de los factores que aparece en la descomposición de Cholesky de la matriz $\frac{1}{\sigma^2}\Sigma_{ih}(h; \theta_0)$.
3. El conjunto de datos es obtenido a partir de la relación

$$z(t_i) = A(z(t_i); \theta_0) \cdot z(t_i) + B(z(t_i); \theta_0) \cdot randn(2, 1).$$

En las simulaciones para obtener los conjuntos de datos se utilizará el tamaño del paso $h = 0,1$, excepto para el vector de parámetros θ_0^{vdP3} . En ese caso, para fines de comparación, se usará $h = 0,1, N = 100$, $h = 0,1, N = 1000$ y $h = 0,01, N = 1000$. Para el vector de parámetros reales θ_0^{vdP1} consideramos la condición inicial $z_{t_0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Para el vector de parámetros reales θ_0^{vdP3} tomamos la condición inicial $z_{t_0} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ como en los trabajos de Pokern *et al.* [16] y Samson *et al.* [17]. Para el vector de parámetros reales θ_{vdP2} suponemos las condiciones iniciales $z_{t_0} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Para el caso donde se observa sólo la posición, otros seis

conjuntos de datos son obtenidos con la forma $z(t_i) = \begin{bmatrix} x(t_i) \\ \hat{y}(t_i) \end{bmatrix}$. En este caso, $\hat{y}(t_i) = \frac{x(t_{i+1}) - x(t_i)}{h}$ es la aproximación de la velocidad mediante un esquema de Euler de primer orden estándar. En este caso se harán las mismas consideraciones en cuanto a h y N .

3.3. Resultados

En las figuras 3.1, 3.2, 3.3 y 3.4 que se presentan a continuación se observa, para cada caso, una gráfica de los datos obtenidos:

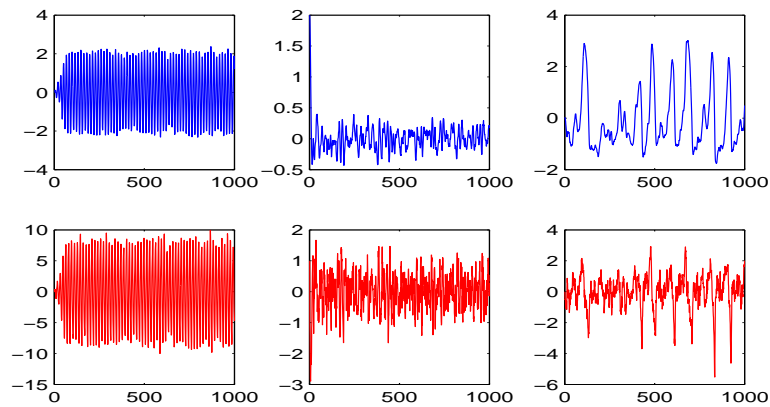


Figura 3.1: Conjunto de datos generados para el modelo de Van der Pol. De izquierda a derecha, en cada columna se muestra la posición (arriba) y la velocidad (abajo) del sistema, obtenidos para los vectores de parámetros reales θ_0^{vdP1} , θ_0^{vdP2} y θ_0^{vdP3} respectivamente.

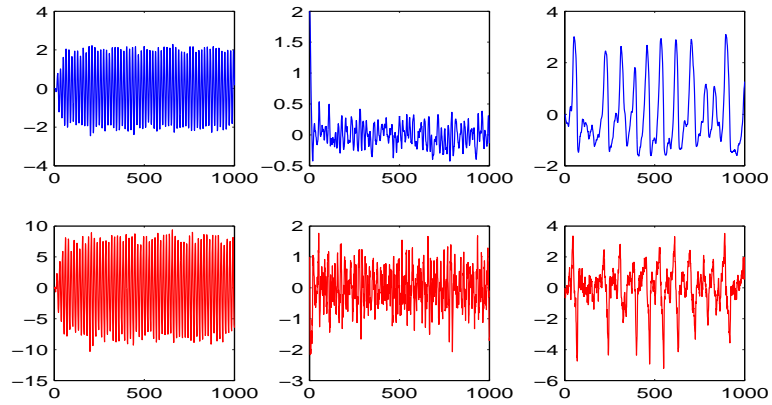


Figura 3.2: Conjunto de datos generados para el modelo de Van der Pol cuando se observa sólo la posición. De izquierda a derecha, en cada columna se muestra la posición (arriba) y la velocidad (abajo) del sistema, obtenidos para los vectores de parámetros reales θ_0^{vdP1} , θ_0^{vdP2} y θ_0^{vdP3} respectivamente.

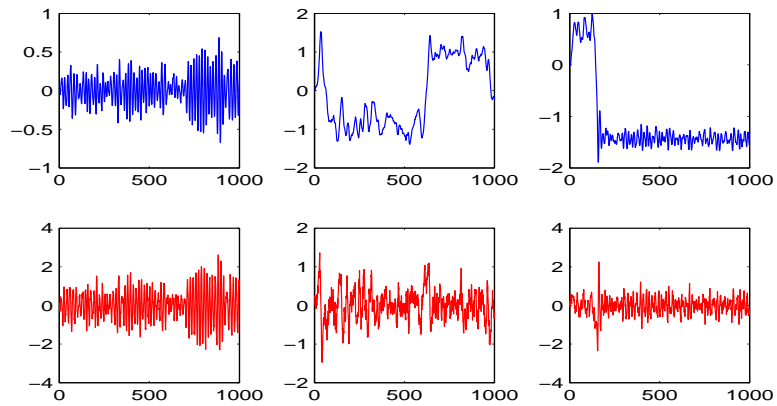


Figura 3.3: Conjunto de datos generados para el modelo de Duffin. De izquierda a derecha, en cada columna se muestra la posición (arriba) y la velocidad (abajo) del sistema, obtenidos para los vectores de parámetros reales θ_0^{D1} , θ_0^{D2} y θ_0^{D3} respectivamente.

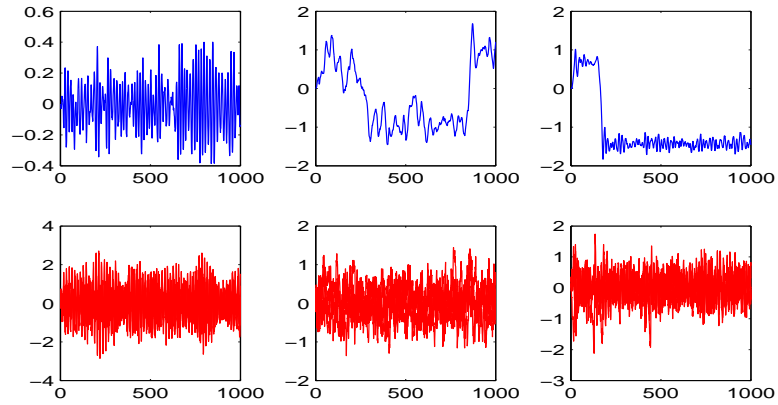


Figura 3.4: Conjunto de datos generados para el modelo de Duffin cuando se observa sólo la posición. De izquierda a derecha, en cada columna se muestra la posición (arriba) y la velocidad (abajo) del sistema, obtenidos para los vectores de parámetros reales θ_0^{D1} , θ_0^{D2} y θ_0^{D3} respectivamente.

3.4. Estimación y comparación

Una vez que tenemos los conjuntos de datos pasamos a la etapa de estimación. Con la idea de evitar algoritmos que hacen uso del gradiente para minimizar la función objetivo, usamos usamos una pequeña variación del algoritmo de Nelder–Mead basado en el método simplex. Para cada conjunto de datos, excepto para los correspondientes al vector de parámetros θ_0^{vdP3} , corremos el algoritmo 100 veces y tomamos el promedio y la desviación estándar de las estimaciones. Para los conjuntos de datos correspondientes al vector de parámetros θ_0^{vdP3} corremos el algoritmo 1000 veces y tomamos el promedio y la desviación estándar. Los parámetros estimados para los casos de datos observados total y parcialmente se denotarán mediante $\widehat{\theta}^C$ y $\widehat{\theta}^P$ respectivamente. Los parámetros estimados en los trabajos de Ozaki, Pokern *et al* y Samson *et al* se identificarán con los sub-índices Oz, Po y Sa respectivamente. La letra E en la primera columna de cada tabla denota el estimador. La segunda columna de cada tabla contiene el valor de los parámetros reales. Para cada vector de parámetros reales, los resultados se presentan a continuación:

E	θ_{vdP1}	N=100	std	N=1000	std	N=5000	std	N=10000	std
$\widehat{a^C}$	1	1.021	0.0906	0.8831	0.0450	1.0139	0.0389	1.0151	0.0754
$\widehat{a^C}_{Oz}$				0.9878					
$\widehat{a^P}$		1.0050	0.0536	1.0445	0.0921	1.0304	0.0430	1.0118	0.1572
$\widehat{b^C}$	0	-0.0122	0.1076	0.0130	0.0845	-0.0101	0.0920	-0.0260	0.1018
$\widehat{b^C}_{Oz}$				-0.0112					
$\widehat{b^P}$		-0.0060	0.1005	0.0133	0.0827	0.0048	0.0992	0.0190	0.1010
$\widehat{c^C}$	-1	-0.9816	0.0939	-0.8447	0.0453	-1.0252	0.1192	-1.0466	0.1894
$\widehat{c^C}_{Oz}$				-0.9742					
$\widehat{c^P}$		-0.9830	0.0917	-0.9983	0.1566	-1.0031	0.0681	-0.9621	0.1468
$\widehat{d^C}$	14.8	14.6709	0.2747	14.8869	0.1764	14.6319	0.3578	14.9214	0.3247
$\widehat{d^C}_{Oz}$				14.8520					
$\widehat{d^P}$		14.6936	0.2851	14.7911	0.4355	14.7890	0.3356	14.8515	0.3634
$\widehat{\sigma^2^C}$	1	0.9966	0.1929	1.1337	0.0302	1.0125	0.1372	1.0819	0.1738
$\widehat{\sigma^2^C}_{Oz}$				1.0141					
$\widehat{\sigma^2^P}$		1.0984	0.1750	1.0389	0.1306	1.0812	0.1597	0.9687	0.2749

Cuadro 3.1: Modelo de Van der Pol 1. Tres estimadores se comparan: datos completamente observados, datos completamente observados obtenido por Ozaki y datos parcialmente observados.

En la tabla 3.1 se observa una buena estimación . Los resultados presentados por Ozaki en [13] corresponden al caso $N = 1000$. En esta caso el estimador de Ozaki funciona mejor que el nuestro, sin embargo cuando N se incrementa nuestros modelos mejoran la estimación y los resultados son similares a los obtenidos por Ozaki. Vale la pena mencionar que en el caso de datos parcialmente observados la estimación funciona muy bien, aún para $N = 100$.

En la tabla 3.2 los resultados son considerablemente buenos para los dos estimadores.

E	θ_{vdP2}	N=100	std	N=1000	std	N=5000	std	N=10000	std
$\widehat{a^C}$		1.6312	0.1953	1.7804	0.3194	2.0858	0.3507	2.2434	0.3919
$\widehat{a^P}$	2	1.8357	0.1993	2.2059	0.3098	1.7038	0.4114	2.1447	0.3257
$\widehat{b^C}$		0.0037	0.0202	0.0024	0.0434	-0.0018	0.0275	-0.0027	0.0341
$\widehat{b^P}$	0	-0.0020	0.0381	0.0010	0.0374	0.0009	0.0267	0.0007	0.0211
$\widehat{c^C}$		2.4118	0.1054	2.1037	0.6030	1.8546	0.2834	2.2821	0.1562
$\widehat{c^P}$	2	1.8926	0.0419	2.2159	0.0804	2.0421	0.3175	1.9150	0.2966
$\widehat{d^C}$		9.8338	0.3999	9.8214	0.6702	9.5326	0.3745	10.4202	0.6696
$\widehat{d^P}$	10	9.2858	0.4746	10.5078	0.6587	9.9742	0.4193	9.4332	0.5790
$\widehat{\sigma^2^C}$		1.0251	0.0051	0.7892	0.0160	0.9284	0.0074	0.8958	0.0249
$\widehat{\sigma^2^P}$	1	0.8941	0.0033	0.9283	0.0073	1.1351	0.0128	1.1284	0.0119

Cuadro 3.2: Modelo de Van der Pol 2. Dos estimadores se comparan: Datos totalmente observados y datos parcialmente observados.

En el trabajo de Pokern *et al.* [16] se estudian tres modelos: modelo de crecimiento estocástico, el oscilador armónico estocástico lineal y el oscilador estocástico trigonométrico. Los mismos tres modelos se estudian en el artículo de Samson *et al.* [17]. Compararemos nuestros resultados con los dos primeros que se mencionaron. Si en el modelo de Van der Pol (3.1) se toma $a = b = 0$ entonces obtenemos el oscilador armónico estocástico lineal. Los resultados se muestran en la tabla 3.3.

E	θ_{vdP3}	N=100		N=1000		N=1000	
		$h = 0,1$	std	$h = 0,1$	std	$h = 0,01$	std
$\widehat{c^C}$ $\widehat{c^C}_{Sa}$ $\widehat{c^P}$ $\widehat{c^P}_{Sa}$	0.5	0.3849	0.0785	0.4866	0.0693	0.5057	0.0876
		1.022	0.098	1.086	0.271	0.678	0.326
		0.3795	0.0619	0.5142	0.0673	0.4284	0.0729
		1.285	0.275	1.215	0.096	0.699	0.330
$\widehat{d^C}$ $\widehat{d^C}_{Sa}$ $\widehat{d^P}$ $\widehat{d^P}_{Sa}$	4	3.8508	0.1000	3.9965	0.0864	3.8189	0.1255
		3.567	0.489	3.488	0.187	4.034	0.642
		3.7767	0.1352	4.0210	0.0741	3.9715	0.1184
		3.588	0.494	3.501	0.188	4.032	0.644
$\widehat{\sigma^2^C}$ $\widehat{\sigma^2^C}_{Sa}$ $\widehat{\sigma^2^P}$ $\widehat{\sigma^2^P}_{Sa}$ $\widehat{\sigma^2^P}_{Po}$	1	0.8033	0.0114	0.9954	0.0033	1.0151	0.0008
		0.980	0.069	0.974	0.021	0.996	0.021
		0.9377	0.0076	1.0176	0.0029	0.9685	0.0005
		0.946	0.074	0.956	0.021	0.994	0.023
		1.154	0.074	1.114	0.025	1.016	0.013

Cuadro 3.3: Modelo de Van der Pol 3. Se comparan: datos total y parcialmente observados, datos total y parcialmente observados obtenidos por Samson *et al.* y la estimación de la varianza obtenida por Pokern *et al.* para datos parcialmente observados.

En la tabla 3.3 se observa que todos los estimadores se comportan relativamente bien. Nuestros resultados mejoran notablemente en el caso $h = 0,1, N = 1000$. Sin embargo, cuando el número de datos aumenta y el tamaño del paso disminuye, nuestro estimador para el caso de datos parcialmente observados desmejora.

Para el oscilador de Duffin correspondiente al primer vector de parámetros reales los resultados se muestran en la tabla 3.4. Se puede apreciar un buen comportamiento en los casos de datos total y parcialmente observados. Además los resultados son similares a los obtenidos por Ozaki.

E	θ_{D1}	N=100	std	N=1000	std	N=5000	std	N=10000	std
$\widehat{a^C}$ $\widehat{a^C}_{Oz}$ $\widehat{a^P}$	4	8.1745	5.5976	3.7635	1.8398	4.1978	0.7339	4.2733	0.6285
				3.7954					
		7.2542	4.6135	4.4201	1.0071	4.3048	1.0117	4.4406	1.6694
$\widehat{b^C}$ $\widehat{b^C}_{Oz}$ $\widehat{b^P}$	0	0.0018	0.0476	-0.0001	0.0216	-0.0015	0.0261	0.0020	0.0280
				0.3551					
		0.0062	0.0562	0.0023	0.0205	-0.0004	0.0155	-0.0018	0.0226
$\widehat{c^C}$ $\widehat{c^C}_{Oz}$ $\widehat{c^P}$	14.8	14.5822	0.4943	14.8029	0.2010	14.7909	0.1550	14.8392	0.1446
				14.9385					
		13.5058	0.3861	14.8850	0.1470	14.7139	0.1352	14.8116	0.1895
$\widehat{d^C}$ $\widehat{d^C}_{Oz}$ $\widehat{d^P}$	0.25	0.2037	0.0282	0.3042	0.0322	0.2328	0.0196	0.2152	0.0164
				0.1589					
		0.3175	0.0312	0.2337	0.0266	0.1990	0.0276	0.3578	0.0238
$\widehat{\sigma^2^C}$ $\widehat{\sigma^2^C}_{Oz}$ $\widehat{\sigma^2^P}$	0.25	0.2253	0.0011	0.2625	0.0004	0.2494	0.0005	0.2343	0.0002
				4.0270					
		0.2016	0.0009	0.2664	0.0003	0.2457	0.0001	0.2564	0.0003

Cuadro 3.4: Modelo de Duffin 1. Se comparan tres estimadores: datos total y parcialmente observados obtenidos por nosotros y datos completamente observados obtenidos por Ozaki.

Para el oscilador de Kramers que corresponde al modelo de Duffin 2, los resultados mostrados en la tabla 3.5 nos muestra que con pocos datos ($N = 100$) nuestros resultados no son buenos. No obstante, cuando ese número aumenta el estimador funciona muy bien.

E	θ_{D2}	N=100	std	N=1000	std	N=5000	std	N=10000	std
$\widehat{a^C}$		0.9080	0.1621	0.9612	0.1105	1.0133	0.2119	1.0906	0.2086
$\widehat{a^P}$	1	0.7809	0.1360	1.0864	0.1215	1.0213	0.0409	1.0123	0.1838
$\widehat{b^C}$		0.0024	0.0264	-0.0062	0.0375	0.0094	0.0493	-0.0023	0.0358
$\widehat{b^P}$	0	0.0054	0.0384	-0.0030	0.0404	-0.0045	0.0350	-0.0017	0.0385
$\widehat{c^C}$		-0.6701	0.1591	-0.9799	0.1461	-0.9673	0.2161	-1.1705	0.2769
$\widehat{c^P}$	-1	-0.8271	0.2117	-1.1448	0.1629	-1.0228	0.0527	-1.0428	0.2730
$\widehat{d^C}$		0.9873	0.1405	0.9477	0.0468	0.9600	0.1052	0.9238	0.2817
$\widehat{d^P}$	1	0.7615	0.1284	1.0387	0.1466	0.9545	0.0443	0.9389	0.1893
$\widehat{\sigma^2^C}$		0.1964	0.0014	0.2437	0.0033	0.2515	0.0068	0.2483	0.0178
$\widehat{\sigma^2^P}$	0.25	0.2499	0.0035	0.2556	0.0038	0.2588	0.0004	0.2434	0.0053

Cuadro 3.5: Modelo de Duffin 2. Se comparan dos estimadores: datos total y parcialmente observados.

Para el modelo de Duffin 3, como se muestra en la tabla 3.6, ambos estimadores no funcionan muy bien para $N = 100, 1000$ and 5000 . Sin embargo, para $N = 10000$ ambos estimadores mejoran.

E	θ_{D3}	N=100	std	N=1000	std	N=5000	std	N=10000	std
$\widehat{a^C}$	4	3.4716	1.9433	3.6678	3.0803	3.4254	3.3269	3.9382	3.2953
$\widehat{a^P}$		3.4407	2.3827	3.7521	3.1696	3.4967	2.4346	4.0809	3.9604
$\widehat{b^C}$	3	1.8289	5.1476	2.2959	8.7992	1.7797	9.2544	2.7084	9.7698
$\widehat{b^P}$		2.0885	6.5927	2.4166	8.7129	1.4950	6.9542	3.1291	11.2185
$\widehat{c^C}$	-4	-4.5941	3.3701	-4.3126	6.2528	-4.5677	6.5279	-4.3060	7.2521
$\widehat{c^P}$		-4.0059	4.5588	-4.3184	6.0181	-5.0932	5.0833	-3.9774	7.9922
$\widehat{d^C}$	1	0.8143	0.3162	0.8946	0.2225	0.8671	0.6516	0.9832	0.3476
$\widehat{d^P}$		0.9104	0.3277	0.8920	0.3492	0.8999	0.4105	0.9610	0.6087
$\widehat{\sigma^2}^C$	0.25	0.2674	0.0417	0.2606	0.0161	0.2856	0.0747	0.2597	0.0357
$\widehat{\sigma^2}^P$		0.3093	0.1138	0.2737	0.0820	0.2594	0.0476	0.2513	0.0439

Cuadro 3.6: Modelo de Duffin 3. Se comparan dos estimadores: datos total y parcialmente observados.

Capítulo 4

APLICACIÓN: Modelo estocástico para la detección de Diabetes

4.1. Preliminares

En el trabajo de Ackerman *et al* [2], se construye un modelo matemático que describe el sistema que regula la cantidad de glucosa en el plasma sanguíneo, durante una prueba oral de tolerancia a la glucosa (POTG). Los autores consideraron dos variables: la cantidad de glucosa en la sangre, en el tiempo t , denotada por $G(t)$ y la concentración de la red de hormonas en la sangre, en el tiempo t , denotada por $H(t)$. La última variable es considerada como el efecto acumulado de todas las hormonas pertinentes. Consideraron además que $G(0) = G_0 > 0$ y $H(0) = H_0 > 0$ representan los niveles óptimos (basales) de glucosa y de la red de hormonas que el paciente tiene al momento de llegar, en ayuna, al laboratorio. El modelo básico es descrito por medio de las ecuaciones

$$\begin{aligned}\frac{dG}{dt} &= F_1(G, H) + J(t) \\ \frac{dH}{dt} &= F_2(G, H),\end{aligned}\tag{4.1}$$

donde F_1 y F_2 son funciones no lineales de G y H . La función $J(t)$ es la tasa externa a la cual se incrementa la concentración de glucosa en la sangre. Como

los autores estaban interesados en saber cuánto se desvían G y H de sus valores óptimos, a través de la sustitución $g = G - G_0$ y $h = H - H_0$ obtienen el sistema

$$\begin{aligned}\frac{dg}{dt} &= F_1(G_0 + g, H_0 + h) + J(t) \\ \frac{dh}{dt} &= F_2(G_0 + g, H_0 + h).\end{aligned}\tag{4.2}$$

Luego de hacer algunas consideraciones de tipo fisiológico y linealizar, obtuvieron el modelo

$$\frac{d^2g}{dt^2} + 2\alpha\frac{dg}{dt} + \omega_0^2g = 0,\tag{4.3}$$

donde ω_0 es la frecuencia natural del sistema. Ahora bien, dependiendo del signo del discriminante $\alpha^2 - \omega_0^2$, tres soluciones se obtienen

1. Si $\alpha^2 - \omega_0^2 > 0$ el sistema es amortiguado y la solución viene dada por

$$G_{ov}(t) = G_{ov}(0) + Ae^{(-\alpha+\omega)t} + Be^{(-\alpha-\omega)t}.\tag{4.4}$$

2. Si $\alpha^2 - \omega_0^2 = 0$ el sistema no es amortiguado y en este caso la solución viene dada por

$$G_{un}(t) = G_{un}(0) + (A + Bt)e^{-\alpha t}.\tag{4.5}$$

3. Si $\alpha^2 - \omega_0^2 < 0$ el sistema es críticamente amortiguado y la solución viene dada por

$$G_{cr}(t) = G_{cr}(0) + Ae^{-\alpha t} \cos(\omega t + \delta).\tag{4.6}$$

Denotando por D_j las mediciones de la cantidad de glucosa en la sangre, en el tiempo t_j , los autores consiguen los valores de los parámetros, en cada caso, para los cuales la función

$$E = \sum_{j=1}^n [D_j - G_{ov,un \ o \ cr}(t_j)]^2\tag{4.7}$$

alcanza un mínimo. Una vez calculados los parámetros, los autores utilizan el período natural del sistema, dado por $T_0 = \frac{2\pi}{\omega_0}$, para establecer un criterio que les permita inferir si el paciente tiene diabetes tipo II: si el valor de T_0 es apreciablemente mayor que 4 horas, entonces el paciente tiene diabetes tipo II.

4.2. Modelo Estocástico

En esta sección presentamos un modelo estocástico para simular la dinámica de la glucosa durante una POTG. En primer lugar, estableceremos un modelo continuo estocástico y su solución analítica. En segundo lugar, construiremos un modelo discreto estocástico que nos permita obtener un estimador de máxima verosimilitud adecuado.

4.2.1. Modelo continuo

Para construir el modelo estocástico, utilizaremos muchos de los argumentos usados por Ackerman *et al* [2]. Sin embargo, nosotros supondremos que la variación en la concentración de la red de hormonas se comporta de manera aleatoria. De esta manera, el sistema (4.2) lo escribimos como

$$\begin{aligned}\frac{dg}{dt} &= F_1(G_0 + g, H_0 + h) + J(t) \\ \frac{dh}{dt} &= F_2(G_0 + g, H_0 + h) + \sigma' dW_t,\end{aligned}\tag{4.8}$$

donde dW_t es un ruido blanco Gaussiano, con media 0 y varianza $\mathbb{E}[dW_t dW_s] = \sigma'^2 \delta(t - s)$. Aquí, $\mathbb{E}[\cdot]$ representa la esperanza, $\delta(t)$ es la función delta de Dirac y σ' es la amplitud del ruido.

Usando la expansión de Taylor de primer orden, para $i = 1, 2$ tenemos que

$$F_i(G_0 + g, H_0 + h) \approx F_i(G_0, H_0) + \frac{\partial F_i(G_0, H_0)}{\partial G} g + \frac{\partial F_i(G_0, H_0)}{\partial H} h + e_i.$$

Considerando que e_i , $i = 1, 2$ son suficientemente pequeños con respecto a g y h , y tomando en cuenta que $F_i(G_0, H_0) = 0$ para $i = 1, 2$, el sistema (4.8) puede ser aproximado por

$$\begin{aligned}\frac{dg}{dt} &= \frac{\partial F_1(G_0, H_0)}{\partial G} g + \frac{\partial F_1(G_0, H_0)}{\partial H} h + J(t) \\ \frac{dh}{dt} &= \frac{\partial F_2(G_0, H_0)}{\partial G} g + \frac{\partial F_2(G_0, H_0)}{\partial H} h + \sigma' dW_t.\end{aligned}\tag{4.9}$$

Siguiendo el mismo análisis que en el trabajo de Ackerman *et al* [2], los signos de las derivadas parciales en la ecuación (4.9) son conocidos y obtenemos el sistema

$$\frac{dg}{dt} = -m_1g - m_2h + J(t) \quad (4.10)$$

$$\frac{dh}{dt} = m_4g - m_3h + \sigma' dW_t, \quad (4.11)$$

donde m_1, m_2, m_3 y m_4 son constantes positivas. Tomando derivadas en (4.10), resulta

$$\frac{d^2g}{dt^2} = -m_1 \frac{dg}{dt} - m_2 \frac{dh}{dt} + \frac{dJ}{dt}. \quad (4.12)$$

Sustituyendo $\frac{dh}{dt}$ de la ecuación (4.11) en la ecuación (4.12) se obtiene

$$\begin{aligned} \frac{d^2g}{dt^2} &= -m_1 \frac{dg}{dt} - m_2(m_4g - m_3h + \sigma' dW_t) + \frac{dJ}{dt} \\ &= -m_1 \frac{dg}{dt} - m_2m_4g + m_2m_3h - m_2\sigma' dW_t + \frac{dJ}{dt}. \end{aligned} \quad (4.13)$$

Ahora bien, de la ecuación (4.10) sabemos que

$$m_2h = -\frac{dg}{dt} - m_1g + J(t).$$

Sustituimos en la ecuación (4.13) y al ordenar los términos obtenemos

$$\frac{d^2g}{dt^2} + (m_1 + m_3) \frac{dg}{dt} + (m_1m_3 + m_2m_4)g = m_3J(t) + \frac{dJ}{dt} - m_2\sigma' dW_t.$$

Denotando

- $\alpha = \frac{m_1+m_3}{2},$
- $\omega_0^2 = m_1m_3 + m_2m_4,$
- $S(t) = m_3J(t) + \frac{dJ}{dt},$
- $\sigma dW_t = -m_2\sigma' dW_t$

resulta el modelo

$$\frac{d^2g}{dt^2} + 2\alpha \frac{dg}{dt} + \omega_0^2g = S(t) + \sigma dW_t. \quad (4.14)$$

Finalmente, como la función $S(t)$ es nula, salvo en el pequeño instante de tiempo en que la glucosa es ingerida, obtenemos el modelo continuo estocástico

$$\frac{d^2g}{dt^2} + 2\alpha\frac{dg}{dt} + \omega_0^2g = \sigma dW_t. \quad (4.15)$$

Este modelo es una versión lineal del oscilador armónico perturbado con ruido blanco estudiado en los precedentes capítulos. Para resolver esta ecuación, la escribiremos en forma vectorial como

$$\dot{z} = Az + CdW_t, \quad (4.16)$$

donde

$$z(t) = \begin{bmatrix} g(t) \\ g'(t) \end{bmatrix} = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -2\alpha \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 0 \\ \sigma \end{bmatrix}.$$

Consideraremos la condición inicial $z(0) = z_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$. Usando el método de variación de parámetros, la solución general de la ecuación (4.16) viene dada por

$$\begin{aligned} z(t) &= e^{At}z_0 + \int_0^t e^{A(t-s)}CdW_s \\ &= \frac{1}{\lambda_1 - \lambda_2} \left(\begin{bmatrix} -\lambda_2x_0e^{\lambda_1t} + \lambda_1x_0e^{\lambda_2t} + y_0e^{\lambda_1t} - y_0e^{\lambda_2t} \\ -\lambda_1\lambda_2x_0e^{\lambda_1t} + \lambda_1\lambda_2x_0e^{\lambda_2t} + \lambda_1y_0e^{\lambda_1t} - \lambda_2y_0e^{\lambda_2t} \end{bmatrix} \right. \\ &\quad \left. + \sigma \int_0^t \begin{bmatrix} e^{\lambda_1(t-s)} - e^{\lambda_2(t-s)} \\ \lambda_1e^{\lambda_1(t-s)} - \lambda_2e^{\lambda_2(t-s)} \end{bmatrix} dW_s \right), \end{aligned} \quad (4.17)$$

donde $\lambda_1 = -\alpha + \sqrt{\alpha^2 - \omega_0^2}$ y $\lambda_2 = -\alpha - \sqrt{\alpha^2 - \omega_0^2}$ son los autovalores de la matriz A . Tomando en cuenta que $x(0) = g(0) = G_0 - G_0 = 0$, entonces $z_0 = \begin{bmatrix} 0 \\ y_0 \end{bmatrix}$.

De esta manera

$$z(t) = \frac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} y_0e^{\lambda_1t} - y_0e^{\lambda_2t} + \sigma \int_0^t (e^{\lambda_1(t-s)} - e^{\lambda_2(t-s)}) dW_s \\ \lambda_1y_0e^{\lambda_1t} - \lambda_2y_0e^{\lambda_2t} + \sigma \int_0^t (\lambda_1e^{\lambda_1(t-s)} - \lambda_2e^{\lambda_2(t-s)}) dW_s \end{bmatrix}. \quad (4.18)$$

Suponiendo el caso en que el sistema es críticamente amortiguado $\alpha^2 - \omega_0^2 < 0$ y denotando $\omega^2 = \omega_0^2 - \alpha^2$, se sigue de (4.18) que la solución, para la concentración de glucosa en la sangre, viene dada por

$$G(t) = G(0) + \frac{G'(0)}{\omega}e^{-\alpha t} \sin(\omega t) + \frac{\sigma}{\omega} \int_0^t (e^{-\alpha(t-s)} \sin[\omega(t-s)]) dW_s. \quad (4.19)$$

Si bien en el caso determinístico se pueden determinar de manera muy aproximada los parámetros del modelo, en el caso estocástico es mucho más complejo estimar los parámetros de la ecuación (4.19). Es por ello que en la próxima sección estableceremos un modelo discreto estocástico que aproxime al modelo estocástico continuo (4.15). También encontraremos un estimador de máxima verosimilitud adecuado para realizar la estimación.

4.2.2. Modelo discreto

En esta sección usaremos el método de linealización local desarrollado en el trabajo de Ozaki [13], con la idea de obtener un esquema discreto que aproxime al modelo continuo (4.15). Además, un estimador basado en la máxima verosimilitud será establecido. Supongamos que $t \in [0, T]$. La ecuación (4.15) escribe como el sistema dinámico estocástico bivariado

$$\dot{z}(t) = f(z(t); \theta) + d\mathbf{W}_t, \tag{4.20}$$

donde

1. $z(t) = \begin{bmatrix} g(t) \\ g'(t) \end{bmatrix}$,

2. $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es una función lineal de g y g' dada por

$$f(z(t); \theta) = \begin{bmatrix} g'(t) \\ -2\alpha g'(t) - \omega_0^2 g(t) \end{bmatrix}. \tag{4.21}$$

donde θ denota un vector cuyas coordenadas son los parámetros.

3. $d\mathbf{W}_t = \begin{bmatrix} 0 \\ \sigma dW_t \end{bmatrix}$.

La matriz de varianza-covarianza para el ruido blanco Gaussiano bivariado continuo viene dada por

$$\mathbb{E} [d\mathbf{W}_t d\mathbf{W}_s^{Tr}] = \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2 \delta(t-s) \end{bmatrix}, \tag{4.22}$$

donde Tr denota la traspuesta de un vector o matriz. Note que la matriz (4.22) es singular.

Supongamos que $s \in [t, t + \Delta t) \subset [0, T]$. Consideremos la parte determinística de la ecuación (4.20). Tomando derivadas en ambos lados obtenemos

$$\ddot{z}(s) = J(\theta)\dot{z}(s),$$

donde

$$J(\theta) = \frac{\partial f}{\partial z} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -2\alpha \end{bmatrix} \quad (4.23)$$

es la matriz Jacobiana de la función f . De esta manera la solución de la ecuación determinística es

$$z(t + \Delta t) = e^{J(\theta)}z(t).$$

Consideremos ahora la ecuación estocástica

$$\dot{z}(s) = e^{J(\theta)}z(t) + d\mathbf{W}_s. \quad (4.24)$$

Usando variación de parámetros se obtiene

$$z(t + \Delta t) = e^{J(\theta)}z(t) + \xi_{t+\Delta t}, \quad (4.25)$$

donde

$$\xi_{t+\Delta t} = \int_t^{t+\Delta t} e^{J(\theta)(t+\Delta t-s)} d\mathbf{W}_s.$$

La matriz de varianza-covarianza del ruido discreto $\xi_{t+\Delta t}$ viene dada por

$$\begin{aligned} \Sigma_{t+\Delta t} &:= \mathbb{E} [\xi_{t+\Delta t} \xi_{s+\Delta t}^{Tr}] \\ &= \frac{\sigma^2}{(\lambda_1 - \lambda_2)^2} \begin{bmatrix} \frac{1}{2\lambda_1} E_1 - \frac{2}{\lambda_1 + \lambda_2} E_{1,2} + \frac{1}{2\lambda_2} E_2 & \frac{1}{2} E_1 - E_{1,2} + \frac{1}{2} E_2 \\ \frac{1}{2} E_1 - E_{1,2} + \frac{1}{2} E_2 & \frac{\lambda_1}{2} E_1 - \frac{2\lambda_1\lambda_2}{\lambda_1 + \lambda_2} E_{1,2} + \frac{\lambda_2}{2} E_2 \end{bmatrix}, \end{aligned} \quad (4.26)$$

donde

- λ_1 and λ_2 son los autovalores de la matriz $J(\theta)$,
- $E_1 = e^{2\lambda_1\Delta t} - 1$,
- $E_2 = e^{2\lambda_2\Delta t} - 1$ y

■ $E_{1,2} = e^{(\lambda_1 + \lambda_2)\Delta t} - 1.$

Note que la matriz de varianza-covarianza del ruido discreto es simétrica. Más aun, para $\Delta t > 0$ fijo, tiene rango 2. Esto nos permite escribir el modelo discreto (4.25) de la siguiente manera:

$$z(t + \Delta t) = e^{J(\theta)}z(t) + B(\Delta t; \theta)\Xi_{t+\Delta t}, \tag{4.27}$$

donde

1. $B(\Delta t; \theta) = \frac{1}{\sigma}\sqrt{\Sigma_{t+\Delta t}}$ es uno de los dos factores de la descomposición de Cholesky de la matriz $\frac{1}{\sigma^2}\Sigma_{t+\Delta t}$ y
2. $\Xi_{t+\Delta t}$ es un ruido Gaussiano bivariado de media cero y matriz de varianza-covarianza dada por $\sigma^2 I$.

Sean $\{t_i, i = 1, \dots, N\} \subset [0, T]$ los instantes de tiempo en que el proceso $z(t)$ es observado. La densidad conjunta del proceso discreto $z(t_i)$ viene dada por

$$\begin{aligned} & p(z(t_1), z(t_2), \dots, z(t_N); \theta, \sigma^2) \\ &= \frac{p(z(t_1))}{(2\pi)^{\frac{N-1}{2}}\sigma^{2(N-1)}|\det B(\theta)|^{N-1}} \prod_{i=2}^{N-1} \exp \left[-\frac{1}{2\sigma^2} (z(t_i) - e^{J(\theta)}z(t_{i-1}))^{Tr} \right. \\ & \quad \left. \times [B^{-1}(\Delta t; \theta)]^{Tr} [B^{-1}(\Delta t; \theta)] (z(t_i) - e^{J(\theta)}z(t_{i-1})) \right]. \end{aligned}$$

De esta manera, la log-verosimilitud del modelo discreto (4.27) es dada por $l(\theta) = \log [p(z(t_1), z(t_2), \dots, z(t_N); \theta, \sigma^2)]$ y por lo tanto, manteniendo sólo los términos que dependen de θ , definimos el contraste de verosimilitud aproximado, para el modelo discreto (4.27) mediante

$$M_{N,\Delta t}(z(t_1), \dots, z(t_N); \theta) = \frac{1}{2N} \sum_{i=0}^{N-1} \|B^{-1}(\Delta t, \theta) [z(t_i) - e^{J(\theta)}z(t_{i-1})]\|^2. \tag{4.28}$$

Observe que las matrices $B^{-1}(\Delta t, \theta)$ y $e^{J(\theta)}$ no dependen del tiempo.

En los teoremas (2.0.1) y (2.0.2), se demuestra la consistencia del contraste (4.28), para el oscilador no lineal, en el caso en que los datos son observados totalmente, es decir; se observa la posición y la velocidad del sistema. Sin embargo, durante

una prueba oral de tolerancia a la glucosa se mide únicamente la concentración de sangre en plasma y no se mide la tasa a la cual decrece (o crece). Es por ello que nos basaremos en el comentario hecho en la observación 3.11, donde se sugiere utilizar un contraste en el cual, la velocidad es aproximada por un esquema de Euler de primer orden estándar. Dicho contraste viene dado por

$$M_{N,h}(x_1, \dots, x_N, \theta) = \frac{1}{2(N-1)} \sum_{i=2}^N \left\| B \left(\begin{bmatrix} x(i) \\ \frac{x(i)-x(i-1)}{h} \end{bmatrix}; \theta \right)^{-1} \begin{bmatrix} x(i+1) \\ \frac{x(i+1)-x(i)}{h} \end{bmatrix} - A \left(\begin{bmatrix} x(i) \\ \frac{x(i)-x(i-1)}{h} \end{bmatrix}; \theta \right) \begin{bmatrix} x(i) \\ \frac{x(i)-x(i-1)}{h} \end{bmatrix} \right\|^2, \quad (4.29)$$

donde $x(i)$ es la posición de la partícula en el tiempo de observación t_i . Es importante señalar que en el modelo que proponemos para la dinámica de la glucosa, el proceso

$$z(t) = \begin{bmatrix} g(t) \\ g'(t) \end{bmatrix}$$

no se observa totalmente. Entonces, recordando que las matrices B^{-1} y A no dependen de $g(t)$, si consideramos la aproximación $g'(t) = \frac{g(t+\Delta t)-g(t)}{\Delta t}$, el contraste que proponemos se define mediante:

$$M_{N,h}(g(t_1), \dots, g(t_N), \theta) = \frac{1}{2(N-1)} \sum_{i=2}^N \left\| B(\theta)^{-1} \begin{bmatrix} g(t_{i+1}) \\ \frac{g(t_{i+1})-g(t_i)}{\Delta t} \end{bmatrix} - A(\theta) \begin{bmatrix} g(t_i) \\ \frac{g(t_i)-g(t_{i-1})}{\Delta t} \end{bmatrix} \right\|^2, \quad (4.30)$$

donde $g(t_i)$ es la concentración de glucosa en el plasma sanguíneo, en el tiempo t_i en que se mide.

La siguiente etapa es el estudio de simulaciones. En este sentido, debemos señalar que pese al buen desempeño del estimador para el caso de datos parcialmente observados que se estudió en el capítulo anterior, en la simulación se usaron varios valores de N (la cantidad de datos): $N = 100, 1000, 5000$ y 10000 . En el caso de la prueba oral de tolerancia únicamente se obtienen 5 datos. Este estudio está en progreso y hemos podido observar dos problemas: el dato inicial del parámetro ω_0 con el que inicia el algoritmo y el paso de tiempo h con el que se aproxima a las matrices B y A son muy sensibles a pequeños cambios. Pensamos que con algunas estrategias con las cuales estamos trabajando, se pudieran tener los resultados y de esta manera probar el modelo con datos verdaderos.

Capítulo 5

CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

Se demostró la consistencia en probabilidad del estimador, basado en el método de linealización local propuesto en el artículo de Ozaki [13], para el oscilador armónico estocástico no-lineal

$$\ddot{x} + g_1(x | a) \dot{x} + g_2(x | b) x = N(t),$$

para el caso en que los datos son observados totalmente. Además, se propuso un estimador, para los parámetros del mismo modelo, en el caso que se observa únicamente la posición. Se elaboraron algoritmos para la generación de datos de ambos modelos y para la estimación de los parámetros en los casos en que los datos son observados total y parcialmente. La estimación de los parámetros en el modelo de van der Pol fue más precisa que la estimación en el modelo de Duffin, en ambos escenarios (observación total o parcial de los datos). Sin embargo, en el modelo de Duffin la estimación de los parámetros en el oscilador de Kramers fue muy buena en ambos escenarios. Si comparamos con los resultados obtenidos por Ozaki en el artículo [13] encontramos: para el modelo de van der Pol $\hat{a} = 0,9878$, $\hat{b} = -0,0112$, $\hat{c} = -0,9742$, $\hat{d} = 0,9878$ y $\hat{\sigma}^2 = 1,0141$ y para el modelo de Duffin $\hat{a} = 3,7954$, $\hat{b} = 0,3551$, $\hat{c} = 14,9385$, $\hat{d} = 0,1589$, $\hat{\sigma}^2 = 4,0270$ (recordemos que

Ozaki utilizó para el modelo de Duffin $\sigma^2 = 4$ y nosotros $\sigma^2 = 0,5$). Podemos inferir entonces que si bien, para el caso de datos totalmente observados ya Ozaki había observado un buen desempeño del estimador, en los casos en que se observa sólo la posición nuestro contraste sugerido muestra un muy buen desempeño si se compara con los resultados obtenidos por otros autores. Finalmente, se propuso un modelo estocástico continuo que simula la dinámica de la glucosa en plasma, durante una prueba oral de tolerancia a la glucosa. A partir del modelo continuo, se construyó un modelo discreto estocástico y un contraste de máxima verosimilitud para estimar los parámetros.

5.2. Recomendaciones

1. El siguiente paso, que surge de manera natural, es la demostración de la normalidad asintótica del estimador estudiado, para el caso en que los datos son observados totalmente.
2. Demostrar la consistencia y la normalidad asintótica para el caso en que los datos son observados parcialmente.
3. Una observación interesante es que, por ejemplo, en la figura 3.3 se observa una relación entre la curva de la posición y el potencial: en el caso del oscilador de Kramers, la partícula se mantiene casi la misma cantidad de tiempo en dos franjas (potencial simétrico), mientras que en la tercera columna la partícula se mantiene poco tiempo en una franja y mucho en otra (potencial asimétrico). Esto sugiere que en estos modelos, el potencial puede sugerir un comportamiento similar al que se observa en algunos modelos estocásticos con régimen de Markov oculto y sería interesante profundizar en ese estudio.
4. Con respecto al modelo de detección de diabetes, será interesante realizar simulaciones que permitan estimar adecuadamente los parámetros del modelo y de allí establecer un criterio que permita discernir si el paciente tiene diabetes tipo II.

Capítulo 6

Apéndice

Demostración del lema 1.2.2 Sea la matriz $B_{ih} := B_{ih}(h; \theta)$ como se define en la ecuación (1.15) y consideremos la transformación

$$\begin{aligned} \Xi_h &= B_h^{-1} (z_{2h} - A_h z_h) \\ \Xi_{2h} &= B_{2h}^{-1} (z_{3h} - A_{2h} z_{2h}) \\ &\vdots \\ \Xi_{(N-1)h} &= B_{Nh}^{-1} (z_{Nh} - A_{(N-1)h} z_{(N-1)h}), \end{aligned}$$

donde $B_{ih}^{-1} = B_{ih}^{-1}(h; \theta)$ y $A_{ih} = A_{ih}(h; \theta)$, para todo $i = 1, \dots, N$. Las variables aleatorias $\Xi_{2h}, \dots, \Xi_{(N-1)h}$ son independientes y tienen distribución Normal, con parámetros 0 y $\sigma^2 I$. La Jacobiana de esta transformación viene dada por

$$J(z, \Xi) = \begin{bmatrix} B_1^{-1} & 0 & 0 & 0 & \cdots & 0 & 0 \\ \frac{\partial \Xi_3}{\partial z_2} & B_2^{-1} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \frac{\partial \Xi_4}{\partial z_3} & B_3^{-1} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & \frac{\partial \Xi_m}{\partial z_{m-1}} & B_{m-1}^{-1} \end{bmatrix}$$

la cual es una matriz triangular y por lo tanto su determinante viene dado por $\prod_{i=1}^{m-1} \frac{1}{\det B_i}$. De la fórmula para la distribución Normal bivariada se tiene

$$p(\Xi_i) = \frac{1}{2\pi \det(\Sigma)^{\frac{1}{2}}} \exp \left[-\frac{1}{2} \Xi_i' \Sigma^{-1} \Xi_i \right], \quad i = 2, \dots, m.$$

En nuestro caso, $\Sigma = \sigma^2 I$, entonces $\det \Sigma = \sigma^4$, $\Sigma^{-1} = \frac{1}{\sigma^2} I$, y se sigue que

$$\begin{aligned} p(\Xi_i) &= \frac{1}{2\pi \sigma^2} \exp \left[-\frac{1}{2\sigma^2} \Xi_i' \Xi_i \right] \\ &= \frac{1}{2\pi \sigma^2} \exp \left[-\frac{1}{2\sigma^2} [z_i - A_{i-1} z_{i-1}]' [B_{i-1}^{-1}]' [B_{i-1}^{-1}] [z_i - A_{i-1} z_{i-1}] \right]. \end{aligned}$$

De esta manera se concluye que:

$$\begin{aligned} p(z_1, z_2, \dots, z_m) &= p(\Xi_2 | z_1) \cdots p(\Xi_m | z_1) p(z_1) = \frac{p(z_1)}{(2\pi)^{m-1} \sigma^{2(m-1)}} \prod_{i=2}^m \frac{1}{|\det B_{i-1}|} \\ &\quad \times \exp \left[-\frac{1}{2\sigma^2} [z_i - A_{i-1} z_{i-1}]' [B_{i-1}^{-1}]' [B_{i-1}^{-1}] [z_i - A_{i-1} z_{i-1}] \right]. \end{aligned}$$

□

Demstración del lema 2.0.3 Considerando la ecuación de observación (1.10), las ecuaciones para $x(s)$ y $y(s)$ vienen dadas por

$$x(s) - x(s') = y(s)(s - s') + \int_{s'}^s y(t) dt = \int_{s'}^s (y(t) - y(s)) dt \quad (6.1)$$

y

$$y(t) - y(s') = \sigma (W_t - W_{s'}) - \int_{s'}^t P(z(u); \theta) du, \quad (6.2)$$

por lo tanto

$$x(s) - x(s') = \sigma \int_{s'}^s (W_t - W_{s'}) dt - \int_{s'}^s \int_{s'}^t P(z(u); \theta) dudt. \quad (6.3)$$

Usando la desigualdad de Jensen obtenemos

$$\begin{aligned} &\mathbb{E} [|x(s) - x(s')|^{2k}] \\ &= \mathbb{E} \left[\left| \sigma \int_{s'}^s (W_t - W_{s'}) dt - \int_{s'}^s \int_{s'}^t P(z(u); \theta) dudt \right|^{2k} \right] \\ &\leq \mathbf{C}_\theta \left\{ \sigma^{2k} \mathbb{E} \left[\left| \int_{s'}^s (W_t - W_{s'}) dt \right|^{2k} \right] + \mathbb{E} \left[\left| \int_{s'}^s \int_{s'}^t P(z(u); \theta) dudt \right|^{2k} \right] \right\}. \end{aligned} \quad (6.4)$$

Utilizando de nuevo la desigualdad de Jensen y haciendo un cambio de variables adecuado, para el primer sumando en (6.4), tenemos

$$\begin{aligned}
 \mathbb{E} \left[\left| \int_{s'}^s (W_t - W_{s'}) dt \right|^{2k} \right] &= \mathbb{E} \left[(s - s')^{2k} \left| \int_{\frac{s'}{s-s'}}^{\frac{s}{s-s'}} (W_{r(s-s')} - W_{s'}) dr \right|^{2k} \right] \\
 &\leq (s - s')^{2k} \int_{\frac{s'}{s-s'}}^{\frac{s}{s-s'}} \mathbb{E} \left[|W_{r(s-s')} - W_{s'}|^{2k} \right] dr \\
 &= \frac{(2k)!(s - s')^{2k}}{2^k k!} \int_{\frac{s'}{s-s'}}^{\frac{s}{s-s'}} [r(s - s') - s']^k dr \\
 &= \frac{(2k)!}{2^k (k + 1)!} (s - s')^{3k} = O([s - s']^{3k}), \tag{6.5}
 \end{aligned}$$

Para el segundo sumando en (6.4) resulta

$$\begin{aligned}
 \mathbb{E} \left[\left| \int_{s'}^s \int_{s'}^t P(z(u); \theta) du dt \right|^{2k} \right] &= \mathbb{E} \left[(s - s')^{2k} \left| \int_{s'}^s \int_{s'}^t P(z(u); \theta) du \frac{dt}{s - s'} \right|^{2k} \right] \\
 &= (s - s')^{2k} \mathbb{E} \left[\left| \int_{s'}^s P((x(u), y(u)) | \theta) \left[\int_u^s \frac{dt}{s - s'} \right] du \right|^{2k} \right] \\
 &= (s - s')^{2k} \mathbb{E} \left[\left| \int_{s'}^s P((x(u), y(u)) | \theta) (s - u) \frac{du}{s - s'} \right|^{2k} \right] \\
 &= (s - s')^{2k} \mathbb{E} \left[\left| \int_{\frac{s'}{s-s'}}^{\frac{s}{s-s'}} P((\dot{x}_{r(s-s')}, x(r(s - s')))) | \theta) [s - r(s - s')] dr \right|^{2k} \right] \\
 &\leq \frac{(s - s')^{4k}}{2k + 1} \mathbb{E} \left[|P(z(r(s - s'))); \theta|^{2k} \right] dr.
 \end{aligned}$$

Usando la desigualdad (1.3) se obtiene

$$|P(z(r(s - s'))); \theta|^{2k} \leq \mathbf{C}_\theta \|z(r(s - s'))\|^{2k\gamma_3},$$

y por lo tanto

$$\mathbb{E} \left[|P(z(r(s - s'))); \theta|^{2k} \right] \leq \mathbf{C}_\theta \mathbb{E} \left[\mathbf{C}_\theta \|z(r(s - s'))\|^{2k\gamma_3} \right] < \infty,$$

donde la última desigualdad se sigue del hecho que los momentos polinomiales del proceso $z(t)$ son finitos. Como el proceso $z(t)$ es estacionario, se sigue de (6.6)

$$\mathbb{E} \left[\left| \int_{s'}^s \int_{s'}^t P(z(u); \theta) du dt \right|^{2k} \right] \leq \mathbf{C}_\theta \mathbb{E} \left[\mathbf{C}_\theta \|z_0\|^{2k\gamma_3} \right] (s - s')^{4k} = O([s - s']^{4k}). \tag{6.7}$$

De (6.5) y (6.7), se deduce de la ecuación (6.4) que

$$\mathbb{E} \left[|x(s) - x(s')|^{2k} \right] \leq \mathbf{C}_\theta \left\{ \sigma^{2k} O([s - s']^{3k}) + O([s - s']^{4k}) \right\} = O([s - s']^{3k}).$$

Realizando cálculos muy similares, se obtiene a partir de la ecuación (6.2)

$$\mathbb{E} \left[|y(s) - y(s')|^{2k} \right] = O([s - s']^k).$$

□

Demostración del lema 1.2.1 Con la idea de simplificar la notación, en esta demostración consideraremos

$$1. J(z_{ih}; \theta) = J_{ih}(\theta) = \begin{bmatrix} 0 & 1 \\ d_{12}(z_{ih}; \theta) & d_{22}(z_{ih}; \theta) \end{bmatrix}.$$

2. Los autovalores de la matriz $J_{ih}(\theta)$ se denotan mediante

$$\begin{aligned} \mu_1(z_{ih}, h; \theta) &= \mu_1 = \frac{-d_{22}(z_{ih}; \theta) + \sqrt{d_{22}^2(z_{ih}; \theta) + 4d_{12}(z_{ih}; \theta)}}{2} \\ \mu_2(z_{ih}, h; \theta) &= \mu_2 = \frac{-d_{22}(z_{ih}; \theta) - \sqrt{d_{22}^2(z_{ih}; \theta) + 4d_{12}(z_{ih}; \theta)}}{2}. \end{aligned}$$

$$3. \Sigma_{ih}(h; \theta) = \Sigma_{ih}(\theta) = \begin{bmatrix} a_{11}(\theta) & a_{12}(\theta) \\ a_{12}(\theta) & a_{22}(\theta) \end{bmatrix}.$$

$$4. \Sigma_{ih}^{-1}(h; \theta) = \Sigma_{ih}^{-1}(\theta) = \begin{bmatrix} b_{11}(\theta) & b_{12}(\theta) \\ b_{12}(\theta) & b_{22}(\theta) \end{bmatrix}.$$

$$5. A_{ih}(h; \theta) = A_{ih}(\theta).$$

6. En todas las expresiones que dependen de h , omitiremos h en sus argumentos a menos que sea necesario recordar la dependencia.

En primer lugar demostraremos que la sucesión $\Sigma_{ih}(\theta)$ converge uniformemente con respecto a θ en probabilidad. De las hipótesis H1–H3 no es difícil verificar que existe una constante positiva \mathbf{C}_θ tal que para $j = 1, 2$ se cumple $|\mu_j| \leq \mathbf{C}_\theta \|z_{ih}\|^\gamma$, para cierto exponente γ . De esta manera se demostraría que los autovalores de la

matriz $J_{ih}(\theta)$ son acotados en probabilidad. La demostración se efectuará para el elemento $a_{11}(\theta)$ ya que los otros elementos pueden ser tratados de manera muy similar. Tenemos

$$\begin{aligned} \frac{a_{11}(\theta)}{h^3} &= \frac{1}{3} + \frac{1}{(\mu_1 - \mu_2)^2} \sum_{k=4}^{\infty} \frac{(2\mu_1)^{k-1} - 2(\mu_1 + \mu_2)^{k-1} + (2\mu_2)^{k-1}}{k!} h^{k-3} \quad (6.8) \\ &= \frac{1}{3} + \frac{1}{(\mu_1 - \mu_2)^2} \sum_{k=4}^{\infty} \frac{(\mu_1)^{k-1} [2^{k-1} - 2(1 + \frac{\mu_2}{\mu_1})^{k-1} + 2^{k-1}(\frac{\mu_2}{\mu_1})^{k-1}]}{k!} h^{k-3}. \end{aligned}$$

Denotando $x = \frac{\mu_2}{\mu_1}$ y definiendo el polinomio de grado $k - 1$

$$Q_{k-1}(x) = 2^{k-1} - 2(1+x)^{k-1} + 2^{k-1}x^{k-1},$$

se puede demostrar fácilmente que 1 es raíz doble Q y por lo tanto

$$Q_{k-1}(x) = (x - 1)^2 \tilde{Q}_{k-3}(x).$$

Se sigue de (6.8)

$$\frac{a_{11}(\theta)}{h^3} = \frac{1}{3} + \sum_{k=4}^{\infty} \frac{(\mu_1)^{k-3} \tilde{Q}_{k-3}(\frac{\mu_2}{\mu_1})}{k!} h^{k-3} = \frac{1}{3} + h \sum_{k=1}^{\infty} \frac{(\mu_1)^k \tilde{Q}_k(\frac{\mu_2}{\mu_1})}{(k+3)!} h^{k-1}.$$

Nos disponemos a acotar esta expresión. Por un lado, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que $||\tilde{Q}_k(x) - \tilde{Q}_k(1)|| < \varepsilon$ para todo $x \in (1 - \delta, 1 + \delta)$. Escogiendo $\varepsilon = |\tilde{Q}_k(1)|$ se obtiene $|\tilde{Q}_k(x)| \leq 2|\tilde{Q}_k(1)|$, para todo $x \in (1 - \delta, 1 + \delta)$. Por otro lado, para x fuera del intervalo $(1 - \delta, 1 + \delta)$ tenemos que $|\tilde{Q}_k(x)| \leq \frac{1}{\delta^2} |Q_k(x)|$. Así, $|\tilde{Q}_k(x)| \leq 2|\tilde{Q}_k(1)| + \frac{1}{\delta^2} |Q_k(x)|$ para todo x . Además $|Q_k(x)| \leq 32^k(1 + |x|^k)$, es decir; $|Q_k(\frac{\mu_2}{\mu_1})| \leq 32^k \frac{\mu_2^k + \mu_1^k}{\mu_1^k}$. Estos cálculos junto con la desigualdad $|\mu_j| \leq \|J_{ih}(\theta)\|$ permiten inferir que

$$\begin{aligned} \left| \sum_{k=1}^{\infty} \frac{(\mu_1)^k \tilde{Q}_k(\frac{\mu_2}{\mu_1})}{(k+3)!} h^{k-1} \right| &\leq \mathbf{C} \left[2 \sum_{k=0}^{\infty} \frac{|\mu_1|^{k+1}}{(k+4)!} h^k + 3 \sum_{k=0}^{\infty} \frac{|\mu_1|^{k+1} + |\mu_2|^{k+1}}{(k+4)!} h^k \right] \\ &\leq \frac{\mathbf{C}}{4} \|J_{ih}(\theta)\| e^{h\|J_{ih}(\theta)\|}. \end{aligned}$$

De la hipótesis H5'' concluimos

$$\left| \frac{1}{h^3} a_{11}(\theta) - \frac{1}{3} \right| = o_{\mathbb{P}}(1),$$

uniformemente con respecto a θ .

Ahora demostraremos la convergencia uniforme en probabilidad de la sucesión $\Sigma_{ih}^{-1}(\theta)$. Tal y como señalamos antes, demostraremos la convergencia uniforme únicamente para un elemento de la matriz ya que los otros se trabajan de manera muy similar. Consideraremos el elemento $b_{22}(\theta)$. Para demostrar la convergencia uniforme en probabilidad con respecto a θ , basta considerar la convergencia uniforme en probabilidad del determinante $\Delta(\theta) = a_{11}(\theta)a_{22}(\theta) - a_{12}^2(\theta)$. En efecto,

$$|hb_{22}(\theta) - 4| = \left| \frac{\frac{a_{11}(\theta)}{h^3}}{\frac{\Delta(\theta)}{h^4}} - 4 \right| \leq \frac{1}{\frac{\Delta(\theta)}{h^4}} \left| \frac{a_{11}(\theta)}{h^3} - \frac{1}{3} \right| + \frac{4}{\frac{\Delta(\theta)}{h^4}} + \left| \frac{\Delta(\theta)}{h^4} - \frac{1}{12} \right|. \quad (6.9)$$

Los dos términos dentro del valor absoluto convergen uniformemente en probabilidad debido a lo que se demostró antes para la matriz $\Sigma_{ih}(\theta)$. Únicamente nos queda estudiar la fracción $(\frac{\Delta(\theta)}{h^4})^{-1}$. Usando la convergencia uniforme en probabilidad de $\Sigma_{ih}(\theta)$, sabemos que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\mathbb{P}\left\{ \sup_{\theta \in \Theta} \left| \frac{\Delta(\theta)}{h^4} - \frac{1}{12} \right| < \delta \right\} \geq 1 - \varepsilon.$$

Dado que, para todo $\omega \in \left\{ \sup_{\theta \in \Theta} \left| \frac{\Delta(\theta)(\omega)}{h^4} - \frac{1}{12} \right| < \delta \right\}$ se cumple que $\frac{1}{12} - \delta < \frac{\Delta(\theta)(\omega)}{h^4} < \frac{1}{12} + \delta$, entonces se infiere

$$\frac{1}{12} - \delta < \inf_{\theta \in \Theta} \frac{\Delta(\theta)(\omega)}{h^4} < \frac{1}{12} + \delta.$$

De esta manera

$$\begin{aligned} \mathbb{P}\left\{ \sup_{\theta \in \Theta} \left(\frac{\Delta(\theta)}{h^4} \right)^{-1} < \frac{1}{\frac{1}{12} - \delta} \right\} &= \mathbb{P}\left\{ \frac{1}{\inf_{\theta \in \Theta} \frac{\Delta(\theta)(\omega)}{h^4}} < \frac{1}{\frac{1}{12} - \delta} \right\} \\ &\geq \mathbb{P}\left\{ \frac{1}{12} - \delta < \inf_{\theta \in \Theta} \frac{\Delta(\theta)(\omega)}{h^4} < \frac{1}{12} + \delta \right\} \\ &\geq \mathbb{P}\left\{ \forall \theta, \frac{1}{12} - \delta < \frac{\Delta(\theta)(\omega)}{h^4} < \frac{1}{12} + \delta \right\} \\ &= \mathbb{P}\left\{ \sup_{\theta \in \Theta} \left| \frac{\Delta(\theta)}{h^4} - \frac{1}{12} \right| < \delta \right\} \geq 1 - \varepsilon, \end{aligned}$$

por lo tanto la sucesión $(\frac{\Delta(\theta)}{h^4})^{-1}$ es uniformemente acotada en probabilidad y esto implica que (6.9) converge uniformemente en probabilidad, es decir;

$$\sup_{\theta \in \Theta} |hb_{22}(\theta) - 4| \rightarrow 0,$$

en probabilidad. De los resultados obtenidos se concluye que

$$\frac{h^3}{2} \|\Sigma_{ih}^{-1}(\theta)\| = O_{\mathbb{P}}(1), \quad (6.10)$$

uniformemente con respecto a θ . □

Esbozo de la demostración del lema 2.0.5 Recordemos de (2.9) que la expresión para el contraste es

$$\begin{aligned} \frac{1}{h} M_{N,h}(z_h^{Nh}; \theta) &= \frac{1}{2hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(\theta) [z_{(i+1)h} - A_{ih}(\theta) z_{ih}], z_{(i+1)h} - A_{ih}(\theta) z_{ih} \rangle \\ &= \frac{1}{h} [\mathcal{D}_1(\theta) + \mathcal{D}_2(\theta) - \mathcal{D}_3(\theta)]. \end{aligned} \quad (6.11)$$

Recordemos también que estamos interesados en estudiar la diferencia

$$\frac{1}{h} [M_{N,h}(z_h^{Nh}; \theta) - M_{N,h}(z_h^{Nh}; \theta_0)] = \frac{1}{h} [\mathcal{D}_1(\theta) - \mathcal{D}_1(\theta_0) + \mathcal{D}_2(\theta) - \mathcal{D}_2(\theta_0) - \mathcal{D}_3(\theta) + \mathcal{D}_3(\theta_0)]. \quad (6.12)$$

Consideremos el término $\mathcal{D}_2(\theta)$ en (6.11). Tenemos que

$$\begin{aligned} \sup_{\theta \in \Theta} |\mathcal{D}_2(\theta)| &\leq \sup_{\theta \in \Theta} \{h^6 \|\Sigma_{ih}^{-1}(\theta)\|\} \frac{1}{N} \sum_{i=0}^{N-1} \sup_{\theta \in \Theta} \left\{ \sum_{k=3}^{\infty} \frac{\|J_{ih}(\theta)\|^{k-1} h^{k-3}}{k!} \|f(z_{ih}; \theta)\| \right\}^2 \\ &\leq \sup_{\theta \in \Theta} \{h^6 \|\Sigma_{ih}^{-1}(\theta)\|\} \frac{1}{N} \sum_{i=0}^{N-1} e^{(2h\|J_{ih}(\theta)\|)} \|J_{ih}(\theta)\|^4 \|f(z_{ih}; \theta)\|^2. \end{aligned}$$

De la hipótesis H5', tomando esperanza y usando el resultado del lema 1.2.1 obtenemos

$$|\mathcal{D}_2(\theta)| = O_{\mathbb{P}}(h^3),$$

uniformemente con respecto a θ . Entonces

$$\frac{|\mathcal{D}_2(\theta)|}{h} = O_{\mathbb{P}}(h^2), \quad (6.13)$$

uniformemente con respecto a θ .

Para estudiar el término $\mathcal{D}_3(\theta)$, supondremos que el análisis del término $\mathcal{D}_1(\theta)$ ya está hecho (se hará más adelante), es decir; supondremos que

$$\frac{1}{2N} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(\theta) (\mathcal{W}_i(\theta) + \mathcal{K}_i(\theta)), \mathcal{W}_i(\theta) + \mathcal{K}_i(\theta) \rangle = O_{\mathbb{P}}(1), \quad (6.14)$$

uniformemente en θ . Ahora bien, usando la desigualdad de Cauchy–Schwarz dos veces, para el término $\mathcal{D}_3(\theta)$ tenemos que

$$|\mathcal{D}_3(\theta)| \leq \left(\frac{1}{N} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(\theta)(\mathcal{W}_i(\theta) + \mathcal{K}_i(\theta)), \mathcal{W}_i(\theta) + \mathcal{K}_i(\theta) \rangle \right)^{\frac{1}{2}} \\ \times \left(\frac{1}{N} \sum_{i=0}^{N-1} \left\langle \Sigma_{ih}^{-1}(\theta) \sum_{k=3}^{\infty} \frac{J_{ih}^{k-1}(\theta)}{k!} h^{k-3} f(z_{ih}; \theta), \sum_{k=3}^{\infty} \frac{J_{ih}^{k-1}(\theta)}{k!} h^{k-3} f(z_{ih}; \theta) \right\rangle \right)^{\frac{1}{2}}.$$

De (6.13) y (6.14) se sigue

$$\frac{|\mathcal{D}_3(\theta)|}{h} = O_{\mathbb{P}}(1)O(h^{\frac{1}{2}}), \quad (6.15)$$

lo cual converge a cero uniformemente en θ . Sólo nos queda demostrar la convergencia uniforme del término $\mathcal{D}_1(\theta)$. De acuerdo a (2.12), el término $\mathcal{D}_1(\theta)$ viene dado por

$$\frac{\mathcal{D}_1(\theta)}{h} = \frac{1}{2hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(\theta)\mathcal{W}_i(\theta), \mathcal{W}_i(\theta) \rangle + \frac{1}{2hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(\theta)\mathcal{K}_i(\theta), \mathcal{K}_i(\theta) \rangle \\ + \frac{1}{hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(\theta)\mathcal{W}_i(\theta), \mathcal{K}_i(\theta) \rangle.$$

Para el segundo sumando tenemos

$$\frac{1}{2hN} \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(\theta)\mathcal{K}_i(\theta), \mathcal{K}_i(\theta) \rangle = \frac{h^3 b_{22}(\theta)}{8N} \sum_{i=0}^{N-1} \ell_1^2(z_{ih}; \theta).$$

Por la hipótesis del lema 2.0.5 y del lema 1.2.1, tomando esperanza resulta

$$\frac{1}{8N} \sum_{i=0}^{N-1} \mathbb{E}[\sup_{\theta \in \Theta} \ell_1^2(z_{ih}; \theta)] = \frac{1}{8} \mathbb{E}[\sup_{\theta \in \Theta} \ell_1^2(z_0; \theta)] < \infty.$$

Así

$$\frac{1}{2hN} \sum_{i=0}^{N-1} \sup_{\theta \in \Theta} \{ \langle \Sigma_{ih}^{-1}(\theta)\mathcal{K}_i(\theta), \mathcal{K}_i(\theta) \rangle \} = o_{\mathbb{P}}(h^2).$$

El tercer término se trabaja de manera similar y se obtiene

$$\frac{1}{2hN} \sum_{i=0}^{N-1} \sup_{\theta \in \Theta} \{ \langle \Sigma_{ih}^{-1}(\theta)\mathcal{W}_i(\theta), \mathcal{K}_i(\theta) \rangle \} \rightarrow 0,$$

en probabilidad. Para terminar la demostración, consideraremos la diferencia en (6.12) y demostraremos que la sucesión

$$\frac{1}{2hN} \left[\sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(\theta) \mathcal{W}_i(\theta), \mathcal{W}_i(\theta) \rangle - \sum_{i=0}^{N-1} \langle \Sigma_{ih}^{-1}(\theta_0) \mathcal{W}_i(\theta_0), \mathcal{W}_i(\theta_0) \rangle \right],$$

converge uniformemente con respecto a θ en probabilidad, a una función continua con respecto a θ . Por ahora supondremos que este resultado ya está demostrado y verificaremos que (6.14) es válido. A la vista de los resultados precedentes, el único término que debemos estudiar es

$$\frac{1}{2N} \sum_{i=0}^{N-1} \langle \Sigma_d^{-1} \mathcal{W}_i(\theta), \mathcal{W}_i(\theta) \rangle = h \sum_{i=1}^{10} \mathcal{G}_{iN}.$$

Comenzaremos con los tres últimos términos de la suma. Dichos términos vienen dados como el producto de una sucesión acotada uniformemente en probabilidad con respecto a θ , por la suma $\frac{1}{N} \sum_{i=0}^{N-1} \mathcal{Z}_i$, para ciertas variables aleatorias i.i.d. \mathcal{Z}_i las cuales no dependen de θ . De la Ley de los Grandes Números se deduce que esas sumas convergen a su esperanza. Esto da como resultado que $h \sum_{i=8}^{10} \mathcal{G}_{iN}$ es acotada en probabilidad uniformemente en θ . Demostraremos ahora que $\sum_{i=1}^7 \mathcal{G}_{iN}$ converge en probabilidad uniformemente en θ a una función continua de θ . Luego, después de multiplicar por h , tiende a cero uniformemente en θ lo cual nos permitirá concluir que (6.14) vale. Consideremos la suma $\sum_{i=1}^7 \mathcal{G}_{iN}$. El primer término viene dado por

$$\begin{aligned} \mathcal{G}_{1N} &= \frac{h^3 b_{11}(\theta)}{8N} \sum_{i=0}^{N-1} [P(z_{ih}; \theta_0) - P(z_{ih}; \theta)]^2 \\ &+ \frac{b_{11}(\theta)}{2hN} \sum_{i=0}^{N-1} \left(\int_{ih}^{(i+1)h} [(i+1)h-u] \{P_x(\tilde{z}(u); \theta_0)[x(u) - x_{ih}] + P_y(\tilde{z}(u); \theta_0)[y(u) - y_{ih}]\} du \right)^2 \\ &+ \frac{h^2 b_{11}(\theta)}{2hN} \sum_{i=0}^{N-1} [P(z_{ih}; \theta_0) - P(z_{ih}; \theta)] \\ &\times \int_{ih}^{(i+1)h} [(i+1)h-u] \{P_x(\tilde{z}(u); \theta_0)[x(u) - x_{ih}] + P_y(\tilde{z}(u); \theta_0)[y(u) - y_{ih}]\} du. \end{aligned}$$

En el segundo sumando, el único término que depende de θ es $b_{11}(\theta)$ y su convergencia en probabilidad, uniforme en θ ya fue demostrada. El tercer sumando converge en probabilidad uniformemente en θ . Esto es consecuencia de:

la convergencia del segundo término, la convergencia en probabilidad, uniforme en θ del primer sumando (que demostraremos a continuación) y la desigualdad de Cauchy–Schwarz. Demostraremos la convergencia en probabilidad del primer sumando. En este caso

$$\begin{aligned} \mathcal{G}_{1N} &= \left(\frac{h^3 b_{11}(\theta)}{8} - \frac{3}{2}\right) \frac{1}{N} \sum_{i=0}^{N-1} [P(z_{ih}; \theta_0) - P(z_{ih}; \theta)]^2 \\ &\quad + \frac{3}{2} \frac{1}{N} \sum_{i=0}^{N-1} [P(z_{ih}; \theta_0) - P(z_{ih}; \theta)]^2. \end{aligned}$$

Consideremos el primer sumando. Usando las hipótesis H3 y H4 tenemos

$$\sup_{\theta \in \Theta} \left(\frac{1}{N} \sum_{i=0}^{N-1} [P(z_{ih}; \theta_0) - P(z_{ih}; \theta)]^2 \right) \leq \sup_{\theta \in \Theta} \left(\frac{1}{N} \sum_{i=0}^{N-1} \|\nabla_{\theta} P(z_{ih}; \lambda_1 \theta_0 + \lambda_2 \theta)\|^2 \|\theta_0 - \theta\|^2 \right).$$

Más aún,

$$\mathbb{E} \left[\sup_{\theta \in \Theta} \frac{1}{N} \sum_{i=0}^{N-1} [P(z_{ih}; \theta_0) - P(z_{ih}; \theta)]^2 \right] \leq \mathbf{C} \mathbb{E}[\|z_0\|^\gamma].$$

Por lo tanto este término es acotado en probabilidad, uniformemente en θ . Este resultado implica la convergencia a cero en probabilidad, uniforme en θ para el primer sumando. Estudiemos ahora el segundo sumando.

Definiendo $G(z; \theta) := [P(z; \theta_0) - P(z; \theta)]^2$, se verifica que

$$\sup_{\theta \in \Theta} |G(z; \theta)| \leq \sup_{\theta \in \Theta} (\|\nabla_{\theta} P(z; \lambda_1 \theta_0 + \lambda_2 \theta)\|^2 \|\theta_0 - \theta\|^2) \leq \mathbf{C} \|z\|^\gamma.$$

Del Teorema de Convergencia Dominada y del hecho que el proceso z_{ih} es mixing, obtenemos una Ley de los Grandes Números, uniforme con respecto a θ . En consecuencia el segundo término converge a $\frac{3}{2} \mathbb{E}[(P(z_0; \theta_0) - P(z_0; \theta))^2]$ uniformemente en θ . Los términos \mathcal{G}_{iN} , $i = 2, 3, 4, 5, 7$ se trabajan de manera muy similar. En efecto, la descomposición en tres términos que hiciéramos de \mathcal{G}_{2N} nos permite verificar que el primer término tiende a cero en probabilidad, uniformemente en θ . Los otros dos términos no dependen de θ . Esto nos permite concluir que el término original \mathcal{G}_{2N} converge en probabilidad a $3\sigma \mathbb{E}[g_1(x_0; \theta_0)]$, uniformemente en θ . Los términos \mathcal{G}_{iN} , $i = 3, 4, 5, 7$ se trabajan de manera muy similar. Finalmente, el término \mathcal{G}_{6N} se estudia de manera similar a \mathcal{G}_{1N} . \square

Demostración de la afirmación en la observación 2.0.4.

- **Oscilador lineal.** En este caso $g_1(x; a) = a$ y $g_2(x; b) = b^2$. La matriz J viene dada por

$$J = \begin{pmatrix} 0 & 1 \\ -b^2 & -a \end{pmatrix}.$$

De esta manera, la matriz exponencial es constante y por lo tanto se cumple la hipótesis.

- **Oscilador de Kramer.** En este caso, nuevamente $g_1(x; a) = a$ y tomando por ejemplo el potencial cuadrático $V(x) = \frac{b_4}{4}x^4 + \frac{b_2}{2}x^2$ tenemos que $g_2(x; \mathbf{b}) = b_4x^2 + b_2$, donde $\mathbf{b} = (b_4, b_2)$. De esta manera

$$J = \begin{pmatrix} 0 & 1 \\ -ay - 3b_4x^2 - b_2 & -a \end{pmatrix}.$$

Se deduce entonces que $\|J(z)\| = 1 + |ay + 2b_4x^3 + (b_4x^2 + b_2)| + |a|$ y por lo tanto

$$\mathbb{E}[e^{\|J(z_0)\|}] \leq \mathbf{C} \int_{\mathbb{R}^2} e^{2ay} e^{2b_4|x|^3 + b_4x^2} e^{-\frac{2a}{\sigma^2}(\frac{1}{2}y^2 + (\frac{b_4}{4}x^4 + \frac{b_2}{2}x^2))} dx dy < \infty.$$

Aquí hemos usado la expresión explícita para la medida invariante de este modelo que aparece en el trabajo de Wu [21]. Para otras formas polinómicas del potencial el resultado también vale.

- **Oscilador de Van der Pol.** En este modelo se tiene que $g_1(x; \mathbf{a}) = a_1x^2 - a_2$, donde ambas constantes son positivas. And again $g_2(x; b) = x^2$. Así,

$$J = \begin{pmatrix} 0 & 1 \\ -2a_1xy - b^2 & -a_1x^2 - a^2 \end{pmatrix}.$$

Ahora bien, en la fórmula (5,16) del trabajo de Wu citado antes, se demostró que existen dos constantes arbitrarias positivas $a > 0$ y $0 < \varepsilon < g_1(x; \mathbf{a}) \frac{b^2}{2\sigma^2}$ tales que

$$\int e^{(aH(x,y) + (g_1(x; \mathbf{a}) \frac{b^2}{2\sigma^2} - \varepsilon)x^4)} d\mu(z) < \infty, \quad (6.16)$$

donde $H(x, y) = \frac{1}{2}y^2 + \frac{b^2}{2}x^2$ y μ es la medida invariante del sistema. Este último resultado implica, con un simple cálculo, que en este caso también vale $\mathbb{E}[e^{2h\|J(z_0)\|}] < \infty$. En efecto, la relación (6.16) implica que $\mathbb{E}[e^{\alpha x_0^4 + \frac{\alpha}{2}y_0^2}] < \infty$, para cierta constante β . Entonces

$$\begin{aligned}\mathbb{E}[e^{2h\|J(z_0)\|}] &\leq \mathbf{C} \int_{\mathbb{R}^2} e^{2h(2a_1|xy|)} e^{2ha_1x^2} d\mu(z) \\ &\leq \mathbf{C} \int_{\mathbb{R}^2} e^{2ha_1(x^2+y^2)} e^{2ha_1x^2} d\mu(z) < \infty,\end{aligned}$$

para $h < h_0$, con h_0 suficientemente pequeño.

- **Modelo Hipoelíptico de FitzHugh-Nagumo.** En el trabajo de León-Samson [9], se considera una ligera modificación del modelo de FitzHugh-Nagumo dada por el sistema

$$\begin{aligned}dX_t &= Y_t dt \\ dY_t &= \frac{1}{\varepsilon}(Y_t(1 - \varepsilon - 3X_t^2) - X_t(\gamma - 1) - X_t^3 - (s + \beta))dt + \sigma dW_t\end{aligned}$$

En este caso, $g_1(x) = \frac{1}{\varepsilon}(\varepsilon + 3x^2 - 1)$ y $g_2(x)x = \frac{1}{\varepsilon}[x(\gamma - 1) + x^3 + (s + \beta)]$. Además

$$f(z) = \begin{bmatrix} y \\ -\frac{1}{\varepsilon}(\varepsilon + 3x^2 - 1)y - \frac{x(\gamma-1)}{\varepsilon} - \frac{x^3}{\varepsilon} - \frac{(s+\beta)}{\varepsilon} \end{bmatrix}.$$

De esta manera se obtiene

$$J(z) = \begin{pmatrix} 0 & 1 \\ -\frac{1}{\varepsilon}[6xy + (\gamma - 1) + 3x^2] & -\frac{1}{\varepsilon}(\varepsilon + 3x^2 - 1) \end{pmatrix}.$$

La norma de esta última matriz viene dada por

$$\|J(z)\| \leq 1 + \frac{1}{\varepsilon} [6|xy| + (\gamma - 1) + 3x^2 + (1 + \varepsilon) + 3x^2].$$

Por lo tanto

$$e^{2h\|J(z)\|} \leq \mathbf{C}e^{12h(|xy|+x^2)} \leq \mathbf{C}e^{2h(3y^2+9x^2)}.$$

Por otro lado, en el trabajo antes citado de León-Samson, se encontró una función de Liapunov $\Psi(z)$ la cual tiene el siguiente comportamiento asintótico:

$$\Psi(z) \sim e^{aH(z)} e^{by} e^{d|x|^3}, \text{ cuando } |x| \rightarrow \infty,$$

para ciertas constantes a, b, d y donde $H(z) = \frac{y^2}{2} + \frac{1}{\varepsilon}(\frac{x^4}{4} + \frac{\gamma-1}{2}x^2 + (s + \beta)x)$ es el Hamiltoniano del sistema. Sabemos de Wu [21], que denotando por μ la medida invariante del sistema, se cumple $\int \Psi(z)d\mu(z) < \infty$. En consecuencia tenemos que

$$\mathbb{E}[e^{2h\|J(z_0)\|}] \leq \mathbf{C} \int e^{2h(9x^2+3y^2)} d\mu(z) < \infty,$$

para $h < h_0$, con h_0 suficientemente pequeño.

□

Remark. Con un procedimiento análogo se demuestra que

$$\mathbb{E}[e^{(2h\|J(z_0)\|)}\|J(z_0)\|^4\|f(z_0)\|^2] < \infty$$

vale para todos los precedentes modelos.

Demostración de la afirmación en la observación 2.0.6.

- **Oscilador lineal.** En este caso $P(z_0; \theta) = -ay_0 - b^2x_0$ y $\theta = (a, b)$. Así,

$$\nabla_{\theta}P(z_0; \theta) = (-y_0, -2bx_0).$$

- **Oscilador de Kramer.** En este caso $g_1(x; a) = a$ y tomando por ejemplo el potencial $V(x) = \frac{b_4}{4}x^4 + \frac{b_2}{2}x^2$ se tiene que $P(z_0; \theta) = -ay_0 - \frac{b_2}{2}x_0 - \frac{b_4}{4}x_0^3$ y $\theta = (a, b_2, b_4)$. Por lo tanto

$$\nabla_{\theta}P(z_0; \theta) = \left(-y_0, \frac{-x_0}{2}, \frac{-3b_4x_0^2}{2}\right).$$

- **Oscilador de Van der Pol.** Aquí se tiene que $g_1(x; \mathbf{a}) = a_1x^2 - a_2$, con $a_1, a_2 > 0$ y $g_2(x; b) = b^2$. En este caso $P(z_0; \theta) = -(a_1x_0^2 - a_2)y_0 - b^2x_0$ y $\theta = (a_1, a_2, b)$. Entonces

$$\nabla_{\theta}P(z_0; \theta) = (-x_0^2y_0, y_0, -2bx_0).$$

- **Modelo Hipoelíptico de FitzHugh-Nagumo.** En este caso

$$g_1(x)y = \frac{1}{\varepsilon}(\varepsilon + 3x^2 - 1)y \quad \text{y} \quad g_2(x)x = \frac{1}{\varepsilon} [x(\gamma - 1) + x^3 + (s + \beta)].$$

De esta manera

$$P(z_0; \theta) = -\frac{1}{\varepsilon}(\varepsilon + 3x_0^2 - 1)y_0 - \frac{1}{\varepsilon} [x_0(\gamma - 1) + x_0^3 + (s + \beta)],$$

y

$$\theta = (\varepsilon, \gamma, s, \beta).$$

Entonces se obtiene

$$\nabla_{\theta}P(z_0; \theta) = \frac{1}{\varepsilon} \left(\frac{(3x_0^2y_0 - 1)y_0 + x_0(\gamma - 1) + x_0^3 + s + \beta}{\varepsilon}, -x_0, -1, -1 \right).$$

□

Bibliografía

- [1] L. Abbaoui and A. Bendjeddou. Exact stationary solutions for a class of nonlinear oscillators. *Advances and Applications in Statistics*, 1(1), 2001.
- [2] E. Ackerman, L. Gatewood, J. Rosevear, and G. Molnar. Blood glucose regulation and diabetes. In *Concepts and Models of Biomathematics.*, chapter 4, pages 131–156. Marcel Dekker, 1969.
- [3] L. Schimansky-Geier B. Lindner. Analytical approach to the stochastic fitzhugh-nagumo system and coherence resonance. *Physical Review*, E(60):7270–7276, 1999.
- [4] Marco Bramanti. *An Invitation to Hypoelliptic Operators and Hörmander’s Vector Fields*. Springer, 1st edition, 2014.
- [5] P. Cattiaux, J. León, and C. Prieur. Estimation for stochastic damping hamiltonian systems under partial observations–i.invariant density. *Stochastic Processes and their Applications*, 124:1236–1260, 2014.
- [6] V. Genon-Catalot. Maximum contrast estimation for diffusion processes from discrete observations. *Statistics*, (21):99–116, 1990.
- [7] V. Genon-Catalot, T. Jeantheau, and C. Laredo. Parameter estimation for discretely observed stochastic volatility models. *Bernoulli* 5, 5:855–872, 1999.
- [8] P.D. Spanos J.B. Roberts. *Random Vibration and Statistical Linearization*. Dover, 2003.
- [9] A. Samson J.R. León. Hypoelliptic stochastic fitzhugh-nagumo neuronal model: mixing, up-crossing and estimation of the spike rate. *hal-01492590*, 2017.

BIBLIOGRAFÍA

- [10] Paul Malliavin. Stochastic calculus of variations and hypoelliptic operators. *Proceedings of the International Conference on Stochastic Differential Equations*, pages 195–263, 1976.
- [11] Gitterman Moshe. *The Noisy Oscillator. The First Hundred Years, From Einstein Until Now*. World Scientific Publishing Co. Pte. Ltd., 1st edition, 2005.
- [12] J. Nicolau. Nonparametric estimation of second-order stochastic differential equations. *Econometric Theory*, 23(5):880–898, 2007.
- [13] T. Ozaki. Statistical identification of nonlinear random vibration systems. *Journal of applied mechanics*, 111:186–191, 1989.
- [14] T. Ozaki. A bridge between nonlinear time series models and nonlinear stochastic dynamical systems: a local linearization approach. *Statistica Sinica*, 2:113–135, 1992.
- [15] C. Prieur P. Cattiaux, J. León. Estimation for stochastic damping hamiltonian systems under partial observations–ii. drift term. *ALEA*, 11(1):1236–1260, 2014.
- [16] Y. Pokern, A. Stuart, and P. Wiberg. Parameter estimation for partially observed hypoelliptic diffusions. *Journal of the Royal Statistical Society*, 71, Part 1:49–73, 2009.
- [17] A. Samson and M. Thieullen. Contrast estimator for completely or partially observed hypoelliptic diffusion. *Stochastic Processes and their Applications*, 122:2521–2552, 2012.
- [18] G. Stoltz T. Lelièvre, M. Rousset. *Free Energy Computations A Mathematical Perspective*. World Scientific, 2010.
- [19] D. Talay. Stochastic hamiltonian systems:exponential convergence to the invariant measure, and discretization by the implicit euler scheme. *Markov Processes and Related Fields*, (8):1–36, 2002.
- [20] A. Wald. Note on the consistency of the maximum likelihood estimate. *Annals of Mathematical Statistic*, 20:595–601, 1949.
- [21] L. Wu. Large and moderate deviations and exponential convergence for stochastic damping hamiltonian systems. *Stochastic Processes and their applications*, 91:205–238, 2001.