

# Búsqueda de diseños cuasi-óptimos eficientes a partir de un diseño D-óptimo para observaciones correlacionadas espacialmente

Cristian David Correa Álvarez

Universidad Nacional de Colombia

Sede Medellín

Facultad de Ciencias

Escuela de Estadística

Medellín, Colombia

2014



# Búsqueda de diseños cuasi-óptimos eficientes a partir de un diseño D-óptimo para observaciones correlacionadas espacialmente

Cristian David Correa Álvarez

Trabajo presentado como requisito parcial para optar al título de:  
**Magister en Estadística**

Director:  
Víctor Ignacio López Ríos, Ph.D. en Estadística

Universidad Nacional de Colombia  
Sede Medellín  
Facultad de Ciencias  
Escuela de Estadística  
Medellín, Colombia  
2014



A Dios por la fortaleza que me dio para alcanzar esta meta, a mis padres y hermanos por el apoyo brindado en el transcurso de mis estudios.



# Agradecimientos

Agradezco, a Dios por haberme permitido llegar hasta este punto y poner en mi camino a personas que fortalecieron mi conocimiento.

A mi director de trabajo de grado, profesor Víctor Ignacio López Ríos por su acompañamiento, dedicación y enseñanzas durante la realización de este trabajo, a mis jurados, profesora Maria Eugenia Castañeda López y profesor Rene Iral Palomino por sus valiosas aportes y sugerencias para el perfeccionamiento de este trabajo, a mis Maestros por su tiempo, por su apoyo, así como por la sabiduría y contribuciones que me transmitieron en el desarrollo de mi formación profesional como Magister. Además doy un agradecimiento muy grande a mi familia, que me brindaron todo su apoyo en el transcurso de mis estudios. A la Universidad Nacional sede Medellín y en especial a la Escuela de Estadística que me dieron la oportunidad de formar parte de ellas.

A todas y cada una de las personas que de alguna u otra manera, contribuyeron a que lograra esta meta que me propuse en la vida y que me ha permitido crecer intelectualmente como profesional y como ser humano.





## Resumen

En este trabajo se construyen diseños cuasi-óptimos a partir de un diseño D-óptimo para modelos con estructura de correlación de los errores; ésta se presenta en algunos problemas cuando las observaciones son tomadas de la misma unidad experimental y se modela con la función de covarianza Matérn.

Se analizan tres casos de la función de Matérn que dependen de un parámetro desconocido  $\rho$ ; el cual es el rango o parámetro de distancia que mide que tan rápido decaen las correlaciones. Se consideran los casos cuando el parámetro  $\rho$  es conocido y desconocido. Los diseños hallados se evalúan a partir del cálculo de la eficiencia de éstos para diferentes valores de  $\rho$  y se comparan con la del diseño cuasi-óptimo obtenido bajo observaciones incorrelacionadas. Se observa que cuando no se tiene en cuenta la estructura de correlación de los errores se pierde eficiencia al utilizar un diseño para observaciones incorrelacionadas. Para seleccionar la función de covarianza Matérn más apropiada para modelar los datos, se utilizan como criterios de selección de modelos: el criterio de información bayesiano (BIC) y el criterio de información de Akaike (AIC). Se realiza un estudio de simulación con los modelos seleccionados, con el fin de ver si el diseño es adecuado para estimar los parámetros del modelo. Finalmente, se comparan los dos diseños: el de observaciones incorrelacionadas y el de observaciones correlacionadas. Se observa que este último, además de tener mejor eficiencia, estima mejor los parámetros del modelo.

**Palabras clave:** Diseños D-óptimos, D-eficiencia, Función de Matérn, Criterio de información bayesiano, Criterio de información de Akaike, Observaciones correlacionadas.

## Abstract

In this work, we built quasi-optimal designs for models with correlated structure in the errors; this kind of structure appears in some problems when the observations are taken from the same experimental unit and it is modeled using the so called Matern covariance function.

We analyze three cases of Matern function depending on an unknown parameter  $\rho$ ; which is the range or distance parameter that measures how quickly they decay the correlations. The case where  $\rho$  is known and unknown are considered. The designs are evaluated by calculating the efficiency for different values of  $\rho$  and compared with the quasi-optimal design obtained under uncorrelated observations. It was observed that when not taken into account the correlation structure in the errors is lost efficiency when using a design for uncorrelated observations. To select the most appropriate Matern covariance function to model the data, are used as model selection criteria: the bayesian information criterion (BIC) and the Akaike information criterion (AIC). A simulation study was performed using the selected models, in order to see whether the design is suitable to estimate the model

parameters. Finally, the two optimal designs are compared, optimal design with correlated observations and optimal design with uncorrelated observations. We observe that optimal design with correlated observations has good efficiency and better parameter estimation too.

**Keywords:** D-optimal designs, D-efficiency, Matern function, Bayesian information criterion, Akaike information criterion, Correlated observations.

# Contenido

<b>Agradecimientos</b>	<b>vii</b>
<b>Resumen</b>	<b>ix</b>
<b>1. Introducción</b>	<b>1</b>
<b>2. Marco Teórico</b>	<b>5</b>
2.1. Introducción a los diseños óptimos . . . . .	5
2.1.1. Criterio D-optimalidad . . . . .	6
2.1.2. Eficiencia de un diseño . . . . .	7
2.2. Estimación de los parámetros del modelo . . . . .	8
2.2.1. Regresión lineal . . . . .	8
2.2.2. Regresión no lineal . . . . .	8
2.2.3. Ejemplos regresión no lineal . . . . .	10
2.2.4. Diseños óptimos . . . . .	17
2.2.5. Regresión no lineal con observaciones correlacionadas . . . . .	18
2.3. Modelos Compartimentales . . . . .	19
<b>3. Diseños óptimos para observaciones correlacionadas</b>	<b>23</b>
3.1. Descripción del modelo . . . . .	23
3.1.1. Estimación por máxima verosimilitud . . . . .	24
3.1.2. Matriz de información . . . . .	25
3.2. Función de covarianza Matérn . . . . .	29
3.2.1. Estructuras de correlación . . . . .	30
3.3. Metodología . . . . .	31
3.4. Aplicación 1: Concentración de la Tetraciclina . . . . .	31
3.4.1. Estimación de valores iniciales (tetraciclina) . . . . .	32
3.4.2. Búsqueda y comparación de diseños cuasi-óptimos para observaciones correlacionadas (tetraciclina) . . . . .	33
3.5. Aplicación 2: Concentración de la sulfisoxazole . . . . .	37
3.5.1. Estimación de valores iniciales (sulfisoxazole) . . . . .	39
3.5.2. Búsqueda y comparación de diseños cuasi-óptimos para observaciones correlacionadas (sulfisoxazole) . . . . .	39

---

<b>4. Estudio de simulación</b>	<b>43</b>
4.1. Simulación usando el caso 2 de la función de covarianza de Matérn . . . . .	43
4.2. Metodología propuesta para las comparaciones entre los diseños considerados	44
<b>5. Conclusiones y recomendaciones</b>	<b>47</b>
5.1. Conclusiones . . . . .	47
5.2. Recomendaciones . . . . .	47
<b>A. Anexo</b>	<b>49</b>
<b>Bibliografía</b>	<b>59</b>

# Lista de Tablas

2-1.	Datos de la producción de pasto en función del tiempo . . . . .	11
2-2.	Estimación, error estándar (es) e intervalos de confianza aproximados del 95 % ( $L_{inf}, L_{sup}$ ) para los parámetros . . . . .	12
2-3.	Concentración del ácido ferúlico en función del tamaño de la raíz . . . . .	13
2-4.	Estimación e intervalos de confianza aproximados del 95 % para los parámetros .	17
2-5.	Diseños óptimos de 4, 5, 6, 7, 8 y 9 puntos para el rendimiento del pasto . . . .	18
2-6.	Diseños óptimos de 4, 5, 6, 7, 8 y 9 puntos para el ácido ferúlico . . . . .	18
3-1.	Valores de las concentraciones de tetraciclina para cada tiempo . . . . .	31
3-2.	Parámetros estimados para cada caso de la función de covarianza de Matérn . .	35
3-3.	Diseños cuasi-óptimos $\xi_{OC}$ y eficiencias para cada caso de la función de covarianza de Matérn . . . . .	36
3-4.	Criterios de información de los diseños cuasi-óptimos para los tres casos de la función de covarianza de Matérn (tetraciclina) . . . . .	37
3-5.	Comparación de diseños exactos con el diseño cuasi-óptimo eficiente del caso 2 de la función de covarianza de Matérn . . . . .	37
3-6.	Concentraciones de sulfisoxazole en (mcg/ml) vs el tiempo (en minutos) . . . .	38
3-7.	Parámetros estimados para cada caso de la función de covarianza de Matérn . .	40
3-8.	Diseños cuasi-óptimos $\xi_{OC}$ para cada caso de la función de covarianza de Matérn . . .	40
3-9.	Criterios de información de los diseños cuasi-óptimos para los tres casos de la función de covarianza de Matérn (sulfisoxazole) . . . . .	40
3-10.	Comparación de diseños exactos con el diseño cuasi-óptimo eficiente del caso uno de la función de covarianza de Matérn . . . . .	41
4-1.	Resultado de los parámetros estimados del estudio de simulación . . . . .	44
4-2.	Parámetros estimados del estudio de simulación para el diseño con observaciones independientes . . . . .	45
4-3.	Parámetros estimados del estudio de simulación para el diseño con observaciones correlacionadas . . . . .	46



# Lista de Figuras

2-1.	Rendimiento del pasto vs tiempo después de la pastura . . . . .	11
2-2.	Curva estimada del rendimiento de pasto vs tiempo después de la pastura .	13
2-3.	Curva estimada de la longitud de la raíz vs la concentración del ácido ferúlico en mM . . . . .	14
2-4.	Gráfico de los cuantiles teóricos vs los muestrales . . . . .	15
2-5.	Gráfica de residuales vs residuales previos . . . . .	16
2-6.	Gráficas de residuales ( $\hat{Y}_i$ vs $e_i$ , concentración vs $e_i$ y $\hat{Y}_i$ vs $ e_i $ ) . . . . .	16
2-7.	Modelo de tres compartimientos . . . . .	20
3-1.	Gráfica de concentración de tetraciclina vs tiempo . . . . .	32
3-2.	Gráfica de parámetros estimados de la concentración de tetraciclina vs tiempo	33
3-3.	Determinante de la matriz $A$ en función de $\rho$ . . . . .	34
3-4.	Valores de $\rho$ vs eficiencia del diseño $\xi_{OI}$ con respecto al diseño $\xi_{OC}$ . . . . .	35
3-5.	Diseños óptimos vs $\rho$ . . . . .	36
3-6.	Gráfica de la concentración de sulfisoxazole vs tiempo . . . . .	38
3-7.	Gráfica de parámetros estimados de la concentración de sulfisoxazole vs tiempo	39





# 1. Introducción

En experimentos prácticos pueden surgir complejidades en cuanto a que algunas variables explicativas no pueden estar bajo control del experimentador y las respuestas pueden ser correlacionadas, debido a que las observaciones en algunos casos son tomadas en el mismo individuo o unidad experimental (Gabrielsson & Weiner 2000). Este tipo de problemas surge de manera natural en el área de farmacocinética, allí es de interés estudiar el vector de parámetros en un modelo de compartimientos, estimar una función no lineal de éste; que es de utilidad para entender la cinética de un medicamento bajo estudio (López-Ríos 2008). En estos casos es necesario formular diseños o aún mejor diseños óptimos donde la estructura del error no cumple con el supuesto de incorrelación.

La teoría de los diseños óptimos se utiliza para obtener los mejores tiempos para la toma de las observaciones (Pazman 1986). Un diseño se define como un conjunto de posibles tiempos donde se deben tomar las observaciones. Cuando las observaciones están correlacionadas se opta por hallar diseños exactos, donde el interés es encontrar un conjunto de tiempos que optimicen un funcional de la matriz de información de Fisher (Amo et al. 2012). En este trabajo se estudian diseños exactos sin considerar réplicas.

Se supone que el modelo que explica una variable respuesta  $Y$ , en términos de los puntos del diseño  $t_i$ , es una función no lineal en los parámetros definida por:

$$Y(t_i) = \eta(t_i, \theta) + \varepsilon_i, \quad i = 1, \dots, N, \quad (1-1)$$

donde  $\eta$  es una función no lineal,  $\theta \in \mathbb{R}^p$ ,  $t_i$  representa el  $i$ -ésimo tiempo o la  $i$ -ésima ubicación de los datos en el espacio ( $i = 1, \dots, N$ ) y toma valores en el espacio de diseño  $\chi \subseteq \mathbb{R}$  y  $\varepsilon^T = (\varepsilon_1, \dots, \varepsilon_N)$  es un vector de errores aleatorios, donde  $\varepsilon \sim N_N(0, \Sigma)$ , con  $Cov(\varepsilon) = [E(\varepsilon(t_i)\varepsilon(t_j))]_{N \times N} = \sigma^2 A$ . Las diferentes estructuras de la matriz  $A$  se muestran en el capítulo 3.

Algunos autores han utilizado estructuras de covarianza particulares, Dette et al. (2008) muestran que para observaciones altamente correlacionadas el diseño D-óptimo está cerca del diseño igualmente espaciado y que éste es por lo general muy eficiente, incluso para tamaños moderados de la correlación, mientras que el diseño D-óptimo obtenido bajo los supuestos de observaciones independientes produce una pérdida sustancial de la eficiencia. Dette et al. (2013) modelan la estructura del error para modelos lineales usando funciones de covarianza tipo kernels, muestran que la distribución uniforme es óptima para una clase de modelos de regresión trigonométricos y que la distribución arco seno es óptima

para el modelo de regresión polinómica con estructura de correlación definido por la potencia logarítmica. Patan & Bogacka (2007), proponen un procedimiento numérico para determinar las aproximaciones de la matriz de información del diseño óptimo cuando los errores están correlacionados. Bogacka & Wright (2005) analizan diseños no lineales localmente óptimos para un modelo de cinética química con errores correlacionados, teniendo en cuenta dos clases de diseños determinados por los parámetros del modelo y la función de varianza del error. Amo et al. (2012), modelan la estructura de covarianza de los errores asumiendo una función de covarianza exponencial, calculan diseños óptimos exactos bajo diferentes criterios, para la discriminación entre dos modelos y la estimación de los parámetros. Por otra parte, Pazman (2007), utiliza criterios de optimalidad para la estimación de parámetros basados en los errores cuadrados medios (MSE) de los estimadores, presenta brevemente expresiones obtenidas para muestras muy pequeñas a través de las densidades de probabilidad de los estimadores.

En este trabajo se calculan diseños cuasi-óptimos, cuando la matriz de covarianza del vector de errores tiene una estructura particular, modelada usando la función de Matérn (Matern 1960), ampliamente conocida en el área de estadística espacial. Esta función se considera apropiada para el tipo de modelos que se estudian en esta tesis. Se analizan tres casos particulares de ésta y se selecciona la que mejor se ajuste a los datos utilizando los criterios de información: criterio de información de Akaike (AIC) (Akaike 1974) y el criterio de información bayesiano (BIC) (Schwarz 1978). Con la función seleccionada, se hace un estudio de simulación para ver que tan eficientes son las estimaciones de los parámetros ( $\theta$ ), la varianza ( $\sigma^2$ ) y la rapidez con la que decaen las correlaciones ( $\rho$ ); ésto implica que los tiempos  $t_i$ ,  $i = 1, \dots, N$  son seleccionados de forma óptima utilizando el criterio D-optimalidad. Por último, se compara el diseño cuasi-óptimo obtenido cuando los errores están correlacionados con el diseño cuasi-óptimo obtenido con errores incorrelacionados, con el fin de cuantificar la magnitud de la eficiencia perdida en la estimación de los parámetros del modelo.

Este trabajo considera adicionalmente los siguientes capítulos. En el segundo capítulo se dan conceptos fundamentales en la teoría de los diseños óptimos como son: la definición de diseños óptimos, matriz de información, criterio D-optimalidad y eficiencia. También, se mencionan los métodos de estimación para modelos de regresión lineal y no lineal, algunos ejemplos donde se visualiza la estimación de los parámetros, diseños óptimos para éstos y se definen modelos de compartimientos que son utilizados para construir los modelos utilizados. En el tercer capítulo se hace un desarrollo teórico de los diseños óptimos para observaciones correlacionadas, la función de covarianza de Matérn considerada para modelar la estructura de covarianza del término de error, algunos casos particulares de ésta, la función de verosimilitud de observaciones correlacionadas y se muestran aplicaciones, donde se estudian modelos de dos compartimientos para aplicar la metodología propuesta; se calculan diseños cuasi-óptimos modelando la estructura de covarianza del error con los tres casos propuestos de la función de covarianza de Matérn y se selecciona la que mejor se

ajuste a los datos. En el cuarto capítulo se hace un estudio de simulación, para observar que tan eficiente es el diseño seleccionado en el capítulo anterior para estimar los parámetros del modelo. También se compara el diseño hallado con el de observaciones incorrelacionadas, con relación a la estimación de los parámetros del modelo. En el quinto capítulo se muestran conclusiones generales y recomendaciones para trabajos futuros.



## 2. Marco Teórico

El interés de los diseños óptimos es determinar las condiciones experimentales, asociadas a las combinaciones de los tratamientos, que minimizan algún funcional de la matriz de información (López-Ríos & Ramos-Quiroga 2007); lo que significa que un diseño dado contiene las condiciones experimentales de los niveles de uno o varios factores y las repeticiones en cada uno de éstos. En este capítulo se ilustran algunos conceptos básicos de la teoría de diseños óptimos para modelos no lineales con observaciones incorrelacionadas y correlacionadas.

### 2.1. Introducción a los diseños óptimos

Hay situaciones en algunas áreas de investigación tales como: medicina, agricultura, industria, entre otros, donde interesa explicar una variable respuesta  $Y$ , a través de una función no lineal en los parámetros usando una variable explicativa  $t$ :

$$Y(t_i) = \eta(t_i, \theta) + \varepsilon_i, \quad i = 1, \dots, N, \quad (2-1)$$

donde  $\theta = (\theta_1, \dots, \theta_p)^T \in \mathbb{R}^p$ ,  $t_i$  es el valor de la variable explicativa en el  $i$ -ésimo individuo o unidad experimental ( $i = 1, \dots, N$ ) y toma valores en el espacio de diseño  $\chi \subseteq \mathbb{R}$ ,  $\eta(t_i, \theta)$  es una función no lineal del vector de parámetros desconocidos  $\theta$  y  $\varepsilon^T = (\varepsilon_1, \dots, \varepsilon_N)$  es el vector de errores, donde  $\varepsilon \sim N_N(0, \sigma^2 I)$ . Un diseño  $\xi$  se define como una medida de probabilidad con soporte finito sobre el espacio de diseño  $\chi$  y se representa por:

$$\xi = \left\{ \begin{array}{ccc} t_1 & \cdots & t_k \\ w_1 & \cdots & w_k \end{array} \right\}, \quad (2-2)$$

donde los  $\{t_1, \dots, t_k\}$  son conocidos como  $k$  puntos de soporte del diseño, con ( $k \leq N$ ) o puntos a muestrear, son elementos del espacio de diseño  $\chi$  con pesos asociados  $w_i = \xi(t_i)$ , para  $i = 1, 2, \dots, k$  tal que  $\sum_i w_i = 1$ . Los  $w_i$  indican la proporción de corridas experimentales que se deben hacer en cada punto de soporte. De la expresión 2-2 se desprenden diseños definidos así: diseño exacto y diseño aproximado o continuo. El diseño es exacto si los pesos  $w_i = r_i/N$ , donde los  $r_i$  son enteros y  $\sum_i^k r_i = N$ ; mientras que en el diseño continuo los pesos  $w_i$  son cualquier número entre 0 y 1 ( $0 < w_i \leq 1$ ), es decir, no se hace ninguna restricción a los pesos de ser números racionales (Atkinson et al. 2007).

Para cada diseño  $\xi$  se define la matriz de información:

$$M(\xi; \theta) = \sum_{i=1}^k f(t_i, \theta) f^T(t_i, \theta) w_i, \quad (2-3)$$

donde  $f(t_i; \theta) = \left[ \frac{\partial \eta(t_i, \theta)}{\partial \theta_1}, \dots, \frac{\partial \eta(t_i, \theta)}{\partial \theta_p} \right]^T$ ; para  $i = 1, 2, \dots, k$ . Es de resaltar que la matriz de información es definida no negativa, es simétrica, es singular si el soporte de  $\xi$  contiene menos puntos que el número de parámetros y es aditiva (Fedorov & Hackl 1997). En este trabajo los diseños a considerar serán diseños exactos, denotados por:  $\xi = \{t_1, t_2, \dots, t_N\}$  es decir, es una sucesión de  $N$ -puntos.

La teoría de los diseños óptimos está relacionada con la selección de un diseño  $\xi$  que maximiza algún funcional de la matriz de información (Kiefer 1959). De esta manera los criterios de optimalidad son definidos como juicios que maximizan o minimizan algún funcional real (con un significado estadístico) de la matriz de información sobre la clase de todos los diseños aproximados definidos en  $\chi$  (López-Ríos & Ramos-Quiroga 2007). En este trabajo se utiliza el criterio D-optimalidad, con el fin de hallar diseños que maximicen el logaritmo del determinante de la matriz de información de Fisher, existen otros criterios; los cuales son citados por Atkinson et al. (2007).

### 2.1.1. Criterio D-optimalidad

El criterio D-optimalidad se define como el siguiente funcional:

$$\psi(\xi; \theta) := \psi_D(M(\xi; \theta)) = |M^{-1}(\xi; \theta)|^{1/p}, \quad (2-4)$$

donde  $p$ : número de parámetros del modelo.

Un diseño  $\xi^*$  es llamado D-óptimo si minimiza la varianza generalizada asociada a los estimadores de máxima verosimilitud del vector de parámetros, la cual es proporcional a:

$$|M^{-1}(\xi; \theta)|, \quad (2-5)$$

equivalente a hallar el diseño  $\xi^*$  que minimiza  $\psi_D(M(\xi; \theta))$ .

Según lo expuesto en Atkinson et al. (2007), es más conveniente el uso de una función convexa dada por el logaritmo natural, puesto que facilita los cálculos y además la función logaritmo natural del determinante es convexa; así, un diseño  $\xi$  se dice que es D-óptimo si minimiza  $-\log(|M(\xi)|)$  o equivalentemente si maximiza  $\log(|M(\xi)|)$ , y se denota por  $\xi^*$ .

Existen teoremas de equivalencia para verificar que un diseño dado es D-óptimo. En la literatura aparece el teorema de equivalencia de Kiefer & Wolfowitz (1960). Este teorema, para un diseño D-óptimo indica que un diseño  $\xi^*$  con matriz de información  $M(\xi^*, \theta)$ , definida positiva, es D-óptimo si, y sólo sí, minimiza el  $\max_{t \in \chi} d(t; \xi^*, \theta)$  si, y sólo sí, en los puntos de soporte del diseño se cumple:

$$d(t; \xi^*, \theta) = p, \quad (2-6)$$

donde  $d(t; \xi^*, \theta) = f^T(t; \theta)M^{-1}(\xi^*; \theta)f(t; \theta)$ . El resultado anterior muestra que el máximo de la función de varianza de la respuesta predicha,  $d(t; \xi^*, \theta)$ , se obtiene en los puntos de soporte del diseño D-óptimo y es igual al número de parámetros del modelo, en caso de D-optimalidad, el peso de cada punto de soporte es inversamente proporcional a  $p$ ,  $w_i = \frac{1}{p}$ .

### 2.1.2. Eficiencia de un diseño

Una forma de medir la eficiencia de un diseño  $\xi$ , con respecto al diseño  $\psi$ -óptimo  $\xi^*$ , encontrado al maximizar el criterio  $\psi$ ; es a partir del cociente del criterio de optimalidad evaluado tanto en  $\xi$  como en  $\xi^*$ ; es decir,

$$\Psi_{ef}(\xi) = \frac{\Psi(\xi; \theta)}{\Psi(\xi^*; \theta)}, \quad (2-7)$$

de donde  $0 < \Psi_{ef}(\xi) \leq 1$ . Un diseño  $\xi$  es comparable con un diseño  $\psi$ -óptimo en la medida que su eficiencia sea cercana a uno, López-Ríos (2008).

Para el criterio D-optimalidad se define la D-eficiencia ( $D_{ef}(\xi)$ ) como:

$$D_{ef}(\xi) = \left[ \frac{|M(\xi; \theta)|}{|M(\xi^*; \theta)|} \right]^{1/p}. \quad (2-8)$$

Al tomar la  $p$ -ésima raíz cuadrada se obtiene una medida de eficiencia que tiene las dimensiones de una razón de varianza. La  $D_{ef}(\xi)$  permite determinar la capacidad del diseño  $\xi$  para estimar de manera eficiente los parámetros del modelo en comparación con el diseño D-óptimo,  $\xi^*$ .

Según Atkinson et al. (2007),  $D_{ef}(\xi)$  es proporcional al número de puntos de soporte del diseño, lo cual permite interpretar la D-eficiencia como el número de réplicas requeridas del diseño para que sea tan eficiente como el diseño D-óptimo.

Algunos autores han construido diseños óptimos para observaciones correlacionadas, entre otros: Muller & Pazman (2003) construyen un algoritmo para determinar el diseño óptimo aproximado para la regresión lineal con errores correlacionados. Pazman (2007), utiliza criterios de optimalidad para la estimación de parámetros basados en el cuadrado medio del error (MSE) de los estimadores. Menciona brevemente expresiones obtenidas para muestras muy pequeñas a través de las densidades de probabilidad de los estimadores. Además, Bogacka & Wright (2005) analizan diseños no lineales localmente óptimos, para un modelo de cinética química con errores correlacionados; teniendo en cuenta dos clases de diseños determinados por los parámetros del modelo y la función de varianza del error.

En la siguiente sección, se presenta el desarrollo teórico de la estimación de los parámetros del modelo, para la regresión lineal, no lineal y no lineal con errores correlacionados.

## 2.2. Estimación de los parámetros del modelo

### 2.2.1. Regresión lineal

Para estimar los parámetros del modelo 2-1, de un diseño  $\xi = \{t_1, \dots, t_N\}$  cuando la función  $\eta(\xi, \theta) := [\eta(t_1, \theta), \dots, \eta(t_N, \theta)]^T = F\theta$  es lineal en  $\theta$  y  $\varepsilon \sim N(0, \sigma^2 I)$  se utiliza el método de mínimos cuadrados. En este método, se desea encontrar el vector  $\hat{\theta}$  de estimadores de mínimos cuadrados que minimice

$$S(\theta) = \varepsilon^T \varepsilon = (Y - F\theta)^T (Y - F\theta) \quad (2-9)$$

El estimador de  $\theta$  que minimiza 2-9 está dado por:

$$\hat{\theta} = (F^T F)^{-1} F^T Y, \quad (2-10)$$

siempre y cuando exista la matriz inversa  $(F^T F)^{-1}$ . La matriz  $(F^T F)^{-1}$  siempre existe si los regresores son linealmente independientes, ésto es, si ninguna columna de la matriz  $F$  es una combinación lineal de las demás columnas. El desarrollo del cálculo de los estimadores se puede encontrar en Montgomery (2006).

### 2.2.2. Regresión no lineal

En el caso donde la función  $\eta(\xi, \theta)$  es no lineal en  $\theta$ , donde  $\xi = \{t_1, \dots, t_N\}$  y  $\varepsilon \sim N(0, \sigma^2 I)$ , es decir,  $\varepsilon$  se distribuye normal multivariado con vector de medias cero y matriz de varianzas y covarianzas  $\sigma^2 I$ , los parámetros del modelo 2-1 se estiman utilizando el método de mínimos cuadrados no lineales generalizados. Este método se utiliza cuando las varianzas del término de error son distintas, es decir, cuando se presenta heterocedasticidad o cuando existe un cierto grado de correlación entre las observaciones (Buse 1973).

Sea  $\theta^0$  el valor verdadero de  $\theta$ , haciendo expansión de Taylor de primer orden alrededor de  $\theta^0$  para  $\eta(\xi, \theta)$  se obtiene:

$$\eta(\xi, \theta) \cong \eta(\xi, \theta^0) + F.(\xi, \theta^0)(\theta - \theta^0). \quad (2-11)$$

Reemplazando 2-11 en 2-1 se tiene:

$$Y \cong \eta(\xi, \theta^0) + F.(\xi, \theta^0)(\theta - \theta^0) + \varepsilon$$

donde  $F.(\xi, \theta^0)$  es una matriz cuya componente  $(i, j)$ , está dada por:  $\frac{\partial \eta(t_i, \theta)}{\partial \theta_j}$ , y  $\eta(\xi, \theta^0)$  es un vector columna,  $N \times 1$  con  $i$ -ésima componente dada por:  $\eta(t_i, \theta^0)$ . Luego la ecuación anterior es equivalente a:

$$Y - \eta(\xi, \theta^0) \cong F.(\xi, \theta^0)(\theta - \theta^0) + \varepsilon$$



Ahora, si se hace  $Y^* = Y - \eta(\xi, \theta^0)$ ,  $\delta = \theta - \theta^0$  y  $F := F(\xi, \theta^0)$  se tiene la estructura del modelo lineal

$$Y^* \cong F\delta + \varepsilon \quad (2-12)$$

La solución de 2-12, realizando un proceso similar al del caso lineal, es la siguiente:

$$\begin{aligned} \hat{\delta} &= (F^T F)^{-1} F^T Y^* \\ \widehat{\theta - \theta^0} &= [F^T(\xi, \theta^0) F(\xi, \theta^0)]^{-1} F^T(\xi, \theta^0) [Y - \eta(\xi, \theta^0)] \end{aligned}$$

Por tanto,

$$\hat{\theta} = \theta^0 + [F^T(\xi, \theta^0) F(\xi, \theta^0)]^{-1} F^T(\xi, \theta^0) [Y - \eta(\xi, \theta^0)] \quad (2-13)$$

Los estimadores de mínimos cuadrados para  $\theta$  se hallan mediante métodos iterativos que depende de un valor inicial  $\theta^0$  (Seber & Wild 2003). Algunas de las estrategias utilizadas para escoger el valor de  $\theta^0$  son: (a) El método de tanteo: consiste en seleccionar distintos valores de los parámetros, graficar las curvas correspondientes, observar que tan cerca están de los puntos reales y escoger los valores de los parámetros que ajusten la mejor curva a los datos; (b) Método de la rejilla: consiste en dar una rejilla de valores para los parámetros como espacio de búsqueda de mejores valores iniciales para la estimación vía mínimos cuadrados no lineales; (c) Definición de funciones inicializadoras: consiste en construir una función que permita dar valores iniciales para la optimización de la función de interés, en algunos casos, se puede hallar a partir de la linealización de la función  $\eta(t, \theta)$ , entre otros (Ritz & Streibig 2008).

Por otra parte, cuando las observaciones son independientes, pero el vector de errores  $\varepsilon$  no tiene varianza constante, se debe modelar la estructura de varianza, porque el ignorarla puede afectar considerablemente tanto los intervalos de confianza como los de predicción (Carroll & Ruppert 1988).

Modelar la heterogeneidad en la varianza explícitamente implica formular un modelo de regresión donde la varianza del término de error no sea constante y hacer la estimación del nuevo modelo, en este caso se considera en el modelo 2-1 que la función  $\eta(t, \theta)$  es no lineal y  $\varepsilon \sim N(0, \Sigma)$ , donde  $\Sigma$  es una matriz diagonal con la componente  $i$  dada por  $\sigma^2 g(t_i, \theta, \rho)$ . La elección de la función de varianza es muy importante para realizar la modelación de los datos y se puede hacer a partir de un soporte teórico. Sin embargo, en la mayoría de los casos se hace a partir de indicaciones cualitativas. Por ejemplo, a partir del análisis de residuos se observa que la varianza de las observaciones crece con su esperanza o la relación entre la varianza de las observaciones y su esperanza se puede describir por medio de una parábola. La elección del modelo se basa en los residuales dentro de una colección de posibles modelos.

Cuando la varianza es creciente se pueden considerar los siguientes casos:

- La varianza es una función de una potencia de la respuesta media:

$$\sigma^2 g(t_i, \theta, \rho) = \sigma^2 \eta(t_i, \theta)^\rho$$

- La varianza es una función lineal de la respuesta, es decir:

$$\sigma^2 g(t_i, \theta, \rho) = \sigma^2 (1 + \rho \eta(t_i, \theta))$$

- La varianza se puede escribir a partir de una parábola, como una generalización del caso anterior. Sea  $y_{\min}$  el menor valor de la función de regresión y  $y_{\max}$  el más grande, estos valores dependen de  $\theta$ . Entonces el modelo es:

$$\sigma^2 g(t_i, \theta, \rho) = \sigma^2 + \sigma^2 \rho_1 (y_{\max} + \rho_2 - \eta(t_i, \theta)) (\eta(t_i, \theta) - y_{\min})$$

Como una varianza es positiva por definición se debe tener cuidado al usar algunos de estos modelos, porque las funciones mencionadas pueden ser en realidad negativas para algunos valores de la variable independiente  $t_i$  y los parámetros estimados del modelo.

Para estimar los parámetros del modelo de regresión heterocedástico se utiliza el logaritmo de la verosimilitud, donde:

$$l(\theta, \sigma^2, \rho) = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^N \left[ \log \sigma^2 g(t_i, \theta, \rho) + \frac{(Y_i - f(t_i, \theta))^2}{\sigma^2 g(t_i, \theta, \rho)} \right].$$

Los estimadores de máxima verosimilitud  $\hat{\theta}$ ,  $\hat{\sigma}^2$ ,  $\hat{\rho}$  que maximizan la función log-verosimilitud no pueden ser obtenidos explícitamente; éstos deben ser obtenidos mediante métodos numéricos (Huet et al. 2004). En este trabajo, se utilizan funciones del  $R$  que permite estimar de manera eficiente la función de log-verosimilitud y de esta manera obtener los parámetros del modelo, tales como: *DEoptim*, *nls2*, *nls*, *nlminb*, *nls*, (R Development Core Team 2013).

En los siguientes ejemplos se ilustra la estimación de los parámetros de un modelo de regresión no lineal y regresión no lineal cuando los errores tienen una estructura de correlación específica, respectivamente.

### 2.2.3. Ejemplos regresión no lineal

#### Ejemplo 1

En este ejemplo, tratado por Ratkowsky (1983), se observa el rendimiento o producción del pasto en función del tiempo después del último pastoreo; los resultados se muestran en la tabla 2-1:

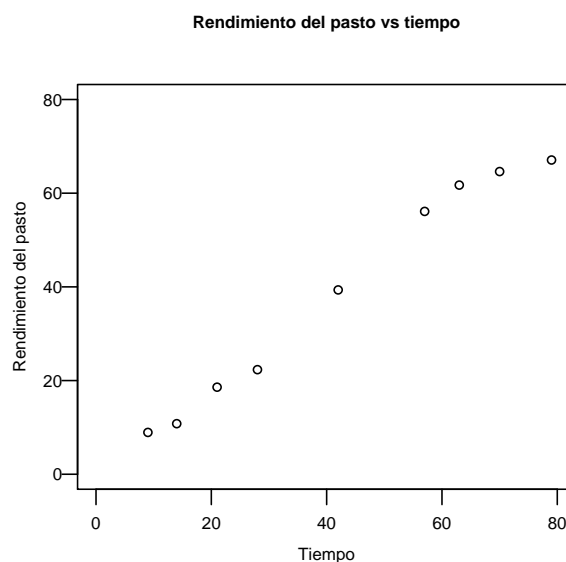
El modelo ajustado a estos datos fue el siguiente:

Tiempo	9	14	21	28	42	57	63	70	79
Rendimiento	8.93	10.8	18.59	22.33	39.35	56.11	61.73	64.62	67.08

**Tabla 2-1.:** Datos de la producción de pasto en función del tiempo

$$Y(t_i) = \theta_1 - \theta_2 \exp[-\exp(\theta_3 + \theta_4 \log(t_i))] + \varepsilon_i, \quad i = 1, \dots, 9. \quad (2-14)$$

Donde,  $\theta_1 \in \mathbb{R}^+$  y  $\theta_1 > \theta_2 \exp[-\exp(\theta_3 + \theta_4 \log(t_i))]$ ,  $\theta_i \in \mathbb{R}$ , para  $i = 2, 3, 4$ ,  $\eta(t_i, \theta)$  es una función no lineal que describe la relación entre el rendimiento y el tiempo,  $t_i$  es el  $i$ -ésimo tiempo después del último pastoreo, los  $\varepsilon_i$  tienen media 0, varianza constante  $\sigma^2$  y son independientes porque las medidas son tomadas en diferentes unidades experimentales.



**Figura 2-1.:** Rendimiento del pasto vs tiempo después de la pastura

En la figura 2-1 se muestra en el eje  $x$  los tiempos después del último pastoreo y en el eje  $y$  el rendimiento del pasto, esta gráfica indica que la curva de respuesta puede presentar un punto de inflexión, por lo que el modelo de Weibull representa una buena elección (Huet et al. 2004). Generalmente, la elección de la función  $\eta$  depende del conocimiento de los fenómenos biológicos observados.

En este trabajo, no se explora la selección de la función  $\eta$  apropiada del modelo. En este ejemplo se muestra cómo estimar los parámetros del modelo suponiendo que la función  $\eta$  es correcta.

Para estimar los parámetros, se buscan valores iniciales de éstos, utilizando el método de la

rejilla. Después de hacer un análisis exploratorio a la información, se optó por las siguientes rejillas para hallar los valores iniciales de los parámetros:  $\theta_1 = \theta_2 = 10:100:10$  (indica que la búsqueda de estos parámetros inicia en 10 y termina en 100 con incrementos de 10),  $\theta_3 = -10:10:1$  y  $\theta_4 = 1:10:1$ ; con estos valores se ajusta el modelo no lineal utilizando la función *nls2* del *R* (ver anexo), se obtienen los siguientes valores iniciales:  $\theta_1 = 80, \theta_2 = 70, \theta_3 = -8$  y  $\theta_4 = 2$ ; con estos valores se ajusta el modelo utilizando la función *nls* del *R*, dando los parámetros estimados mostrados en la tabla **2-2**.

	$\hat{\theta}$	<i>es</i>	$L_{inf}$	$L_{sup}$
$\theta_1$	69.955	2.362	63.8834	76.0269
$\theta_2$	61.681	3.193	53.4743	69.8885
$\theta_3$	-9.209	0.817	-11.3098	-7.1080
$\theta_4$	2.378	0.221	1.8097	2.94586

**Tabla 2-2.:** Estimación, error estándar (*es*) e intervalos de confianza aproximados del 95 % ( $L_{inf}, L_{sup}$ ) para los parámetros

En la tabla **2-2**, se muestra para cada parámetro el estimador  $\hat{\theta}$ , el error estándar (*es*) y un intervalo de confianza aproximado del 95 %. Se observa con una confianza aproximada del 95 % que los valores de los parámetros se encuentran en los siguientes intervalos: [63.8834, 76.0269], [53.4743, 69.8885], [-11.3098, -7.1080] y [1.8097, 2.94586] para  $\theta_1, \theta_2, \theta_3$  y  $\theta_4$ , respectivamente, el estimador del parámetro con menor variabilidad es  $\hat{\theta}_4$ . También se puede decir, que estos parámetros son significativos, porque ninguno de los intervalos contiene al cero, lo que indica que el modelo utilizado para modelar la producción de pasto con respecto al tiempo de crecimiento es adecuado.

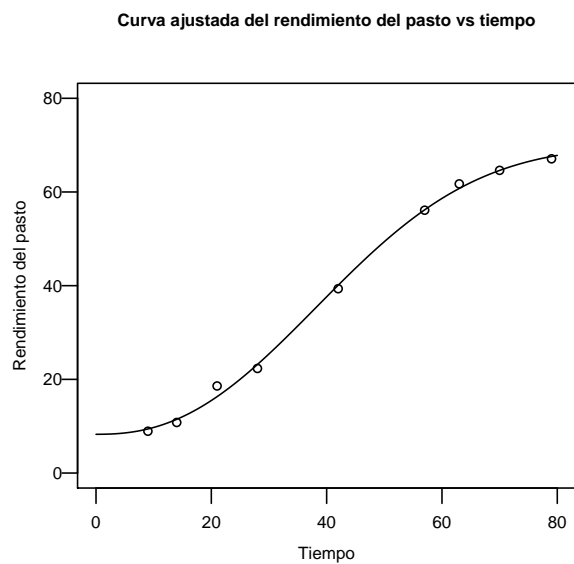
En la figura **2-2** se observa que la curva de la función se ajusta bien a los datos, mostrando la adecuación del modelo a los datos.

## Ejemplo 2

Los datos de la tabla **2-3** se pueden encontrar en la librería *ryegression* "datos *drc*" del *R*, provienen de un estudio que investiga la acción conjunta de mezclas de ácidos fenólicos en la inhibición del crecimiento de *perennial ryegrass* (tipo de semilla de hierba) y su significado en investigación alelopática. Se estudia la toxicidad del ácido ferúlico antioxidante de origen vegetal, que se puede encontrar naturalmente en las aceitunas y en las moras cuando se aplica al *perennial ryegrass* para un rango de concentraciones. La respuesta es la longitud de la raíz en cm y la variable predictora es la concentración del ácido ferúlico en mM.

El modelo para los datos es el siguiente:

$$Y(t_i) = \eta(t_i, \theta) + \varepsilon_i, \quad i = 1, \dots, 24, \quad (2-15)$$



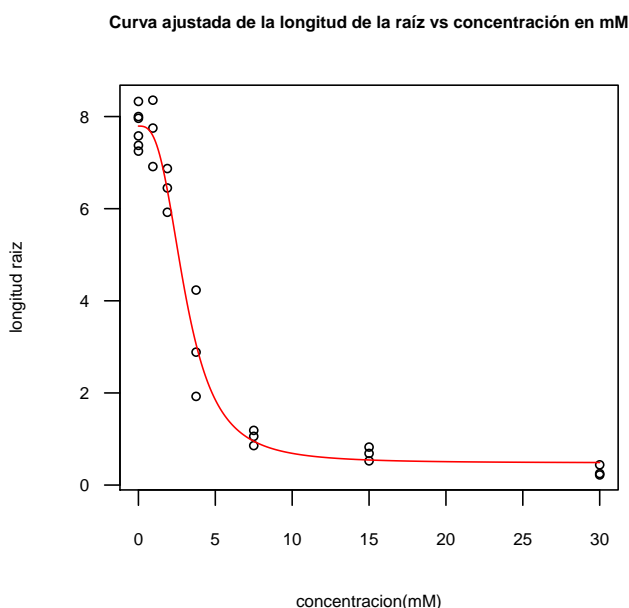
**Figura 2-2.:** Curva estimada del rendimiento de pasto vs tiempo después de la pastura

Concentración	0.00	0.00	0.00	0.00	0.00	0.00	0.94	0.94
Tamaño Raíz	7.58	8.0	8.328571	7.250	7.375	7.9625	8.355555	6.914285
Concentración	0.94	1.88	1.88	1.88	3.75	3.75	3.75	7.50
Tamaño Raíz	7.75	6.871428	6.45	5.922222	1.925	2.885714	4.233333	1.1875
Concentración	7.50	7.50	15.00	15.00	15.00	30.00	30.00	30.00
Tamaño Raíz	0.857142	1.057142	0.6875	0.525	0.825	0.25	0.22	0.44

**Tabla 2-3.:** Concentración del ácido ferúlico en función del tamaño de la raíz

Donde,  $\eta(t_i, \theta) = \theta_3 + (\theta_4 - \theta_3)/[1 + (t_i/\theta_5)^{\theta_2}]$  es una función no lineal que describe la relación entre la concentración del ácido ferúlico y la longitud de la raíz,  $t_i$  es la  $i$ -ésima longitud de la raíz, los  $\varepsilon_i$  tienen media 0, varianza no constante.

Para analizar los datos, primero se hace el ajuste del modelo sin tener en cuenta la estructura de correlación en los errores, se observará si se cumple el supuesto de homogeneidad de varianza, utilizando gráficos y test de hipótesis. Luego, se ajusta el modelo teniendo en cuenta la estructura de correlación de los errores.

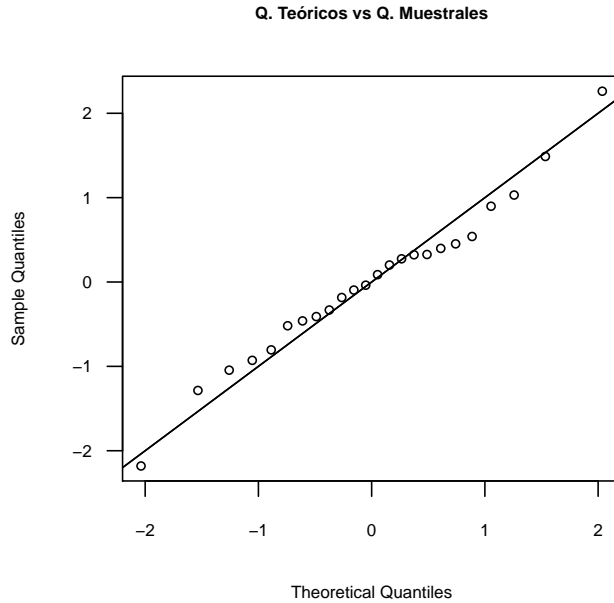


**Figura 2-3.:** Curva estimada de la longitud de la raíz vs la concentración del ácido ferúlico en mM

Los parámetros se estiman, utilizando valores iniciales hallados por el método de tanteo; los valores iniciales son:  $\theta_2 = 1, \theta_3 = 0.6, \theta_4 = 8$  y  $\theta_5 = 3$ ; con estos valores se ajusta el modelo utilizando la función *nls* del *R*. La curva ajustada a los datos con los parámetros estimados se muestra en la figura 2-3, se observa un buen ajuste a los datos. Ahora, con los residuales del modelo se verifica si se cumple el supuesto de normalidad, independencia y homogeneidad de varianza.

Para validar el supuesto de normalidad se realiza el gráfico de los cuantiles teóricos vs los muestrales. Éste se muestra en la figura 2-4, donde junto con el test de normalidad de Shapiro Wilk (Shapiro & Wilk 1965), dando un  $p - valor = 0.9345$ ; como  $p - valor = 0.9345 > 0.05 = \alpha$  no se rechaza la hipótesis de normalidad en los errores del modelo. Ahora, el supuesto de independencia se puede garantizar a partir de la formulación del experimento al usar unidades experimentales distintas en cada medición.

También, a partir del gráfico de rezagos; se grafican los residuales contra los residuales previos. En la figura **2-5** se muestra el gráfico, donde se observa que no hay una tendencia lineal, indicando que los errores son incorrelacionados.



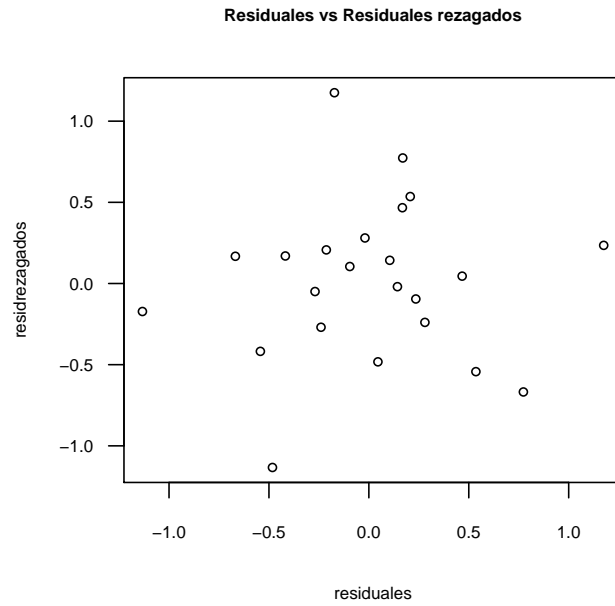
**Figura 2-4.:** Gráfico de los cuantiles teóricos vs los muestrales

El supuesto de homogeneidad se valida teniendo en cuenta la figura **2-6**, en la que se muestran las gráficas de  $\hat{Y}_i$  vs  $e_i$ , concentración vs  $e_i$  y  $\hat{Y}_i$  vs  $|e_i|$ . Se observa en los gráficos que posiblemente el supuesto de homogeneidad de varianza no se cumple, de hecho se puede observar que la varianza de los residuales disminuye conforme la concentración aumenta. Para comprobar lo anterior se utiliza el test de Bartlett para la homogeneidad de varianza de los errores (Bartlett 1937). Este test es sensible al supuesto de normalidad, la función en  $R$  es *bartlett.test*. El test arrojó un  $p$ -valor = 0.03973 < 0.05 =  $\alpha$ , indicando que evidentemente el supuesto de homogeneidad de varianza no se cumple; como esto ocurre se debe usar una función que modele la variabilidad de los datos.

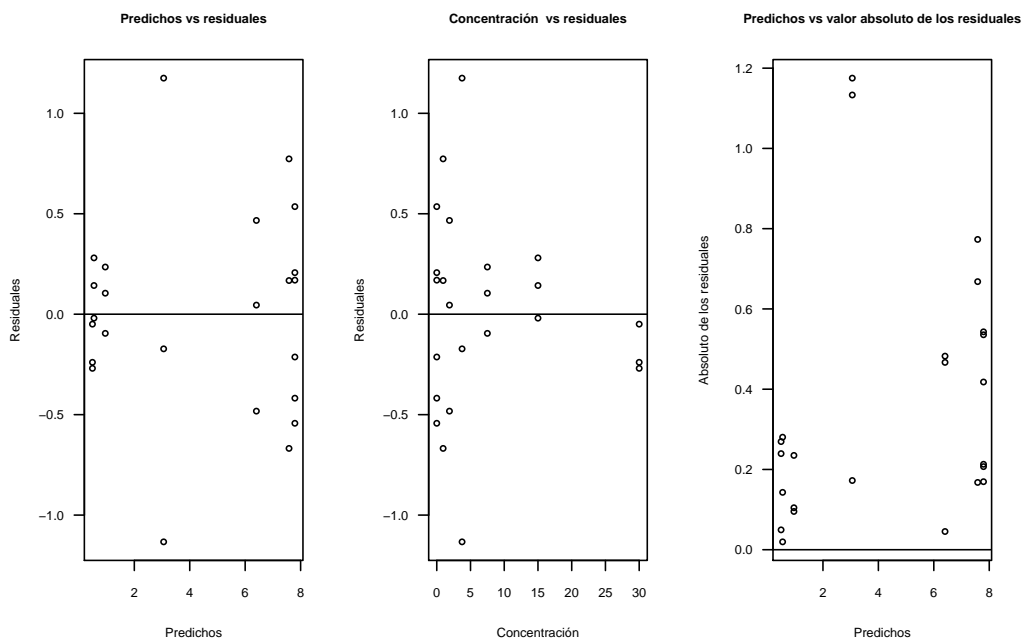
Para modelar la varianza del error se utiliza la función  $V(\varepsilon_i) = \sigma^2 f(t_i, \theta)^{2\rho}$  (Huet et al. 2004). Luego, teniendo en cuenta el modelo y la función de varianza, se puede usar mínimos cuadrados no lineales generalizados para estimar el vector de parámetros  $\theta$  y  $\rho$  de tal forma que minimicen:

$$Q(\theta, \rho) = \sum_{i=1}^{24} V(Y_i)^{-1} (Y_i - f(t_i, \theta))^2, \quad (2-16)$$

donde,  $V(Y_i) = V(\varepsilon_i) = \sigma^2 f(t_i, \theta)^{2\rho}$ . Estos valores se calculan utilizando la función del



**Figura 2-5.:** Gráfica de residuales vs residuales previos



**Figura 2-6.:** Gráficas de residuales ( $\hat{Y}_i$  vs  $e_i$ , concentración vs  $e_i$  y  $\hat{Y}_i$  vs  $|e_i|$ )



$R$  *gnls* de la librería *nlme*. Los resultados obtenidos para  $\hat{\rho}$  y  $\hat{\sigma}$  son: 0.4185 y 0.5196, respectivamente y los de  $\hat{\theta}$ , junto con el error estándar (*es*) y un intervalo de confianza aproximado del 95% se muestran en la tabla **2-4**. Reemplazando los valores estimados de  $\rho$  y  $\sigma$  en la función de varianza se tiene  $\hat{V}(\varepsilon_i) = (0.5196)^2 f(t_i, \theta)^{2(0.4185)}$ . Utilizando la prueba de razón de verosimilitudes, se muestra que el parámetro  $\rho \neq 0$ , puesto que el  $p$ -valor = 0.0228 < 0.05, indicando que la función de varianza es apropiada para modelar la estructura de heteroscedasticidad del error.

	$\hat{\theta}$	<i>es</i>	$L_{inf}$	$L_{sup}$
$\theta_2$	2.712	0.383	2.0259	3.9385
$\theta_3$	0.422	0.107	0.0444	0.9184
$\theta_4$	7.844	0.267	7.3972	8.1887
$\theta_5$	3.055	0.205	2.6703	3.4456

**Tabla 2-4.**: Estimación e intervalos de confianza aproximados del 95% para los parámetros

En la tabla **2-4** se observa que el parámetro estimado con menor variabilidad es  $\hat{\theta}_3$  y los parámetros son significativos, puesto que ninguno de los intervalos de confianza contienen a cero. Esto indica que el modelo utilizado es apropiado para modelar la concentración del ácido ferúlico contra la longitud de la raíz.

En la siguiente sección se muestran los diseños óptimos para cada uno de los ejemplos estudiados en ésta.

#### 2.2.4. Diseños óptimos

Los diseños óptimos para los ejemplos de la sección anterior se presentan a continuación:

- Para el ejemplo 1, se construyen diseños óptimos de 4, 5, 6, 7, 8 y 9 tiempos experimentales después del último pastoreo, luego se calcula la eficiencia de estos con respecto al diseño D-óptimo de 4 puntos. Los diseños se muestran en la tabla **2-5**

Es de resaltar que los diseños de la tabla **2-5** tienen pesos asociados  $w_i = 1/N$ , respectivamente y que el tiempo máximo de experimentación es 79, porque es el mayor de los tiempos iniciales considerados en la investigación. Se observa que los diseños de 5, 6, 7, 8 y 9 tiempos experimentales pierden eficiencia con respecto al diseño D-óptimo, donde el diseño que tiene 6 tiempos se observa la mayor pérdida y la menor se presenta es en el diseño de 8 tiempos. Además, el diseño de 8 puntos se puede reducir a uno de 4, tomando dos réplicas en cada tiempo. Se recomiendan los diseños  $\xi_1, \xi_3, \xi_4$  o  $\xi_5$  dependiendo del número de tiempos que se requieran, puesto que sus eficiencias son superiores al 95%.

- Para el ejemplo 2, se construyen diseños óptimos de 4, 5, 6, 7, 8 y 9 concentraciones experimentales del ácido ferúlico, luego se calcula la eficiencia de estos con respecto

$\xi$	$D_{ef}(\xi)$
$\xi^* = \left\{ \begin{array}{cccc} 0.00012 & 28.56721 & 54.06998 & 79.00000 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \right\}$	1.00000
$\xi_1 = \left\{ \begin{array}{ccccc} 0.00063 & 28.56705 & 28.56726 & 54.06994 & 79.00000 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \end{array} \right\}$	0.95136
$\xi_2 = \left\{ \begin{array}{cccccc} 0.00003 & 0.00016 & 28.56721 & 54.06996 & 54.06999 & 79.00000 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{array} \right\}$	0.94281
$\xi_3 = \left\{ \begin{array}{cccccc} 0.00001 & 0.00032 & 28.56721 & 28.56722 & 54.07001 & 54.06997 & 79.00000 \\ 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \end{array} \right\}$	0.96102
$\xi_4 = \left\{ \begin{array}{cccccc} 0.00214 & 0.00318 & 28.56707 & 28.56677 & 54.06908 & 54.06908 & 79.00000 & 79.00000 \\ 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \end{array} \right\}$	0.99999
$\xi_5 = \left\{ \begin{array}{cccccc} 0.00102 & 0.00106 & 28.56825 & 28.56561 & 28.56912 & 54.06999 & 54.07064 & 79.00000 & 79.00000 \\ 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 \end{array} \right\}$	0.98371

**Tabla 2-5.:** Diseños óptimos de 4, 5, 6, 7, 8 y 9 puntos para el rendimiento del pasto

al diseño D-óptimo de 4 puntos. Los diseños se muestran en la tabla **2-6**.

$\xi$	$D_{ef}(\xi)$
$\xi^* = \left\{ \begin{array}{cccc} 0.00026 & 2.46560 & 5.30382 & 30.00000 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \right\}$	1.00000
$\xi_1 = \left\{ \begin{array}{ccccc} 0.00029 & 2.20642 & 2.85889 & 5.42156 & 30.00000 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \end{array} \right\}$	0.95254
$\xi_2 = \left\{ \begin{array}{cccccc} 0.00095 & 2.23252 & 2.78598 & 5.37985 & 30.00000 & 30.00000 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{array} \right\}$	0.94347
$\xi_3 = \left\{ \begin{array}{cccccc} 0.00027 & 2.46559 & 2.46560 & 5.30380 & 5.30381 & 30.00000 & 30.00000 \\ 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \end{array} \right\}$	0.96103
$\xi_4 = \left\{ \begin{array}{cccccc} 0.00011 & 0.00138 & 2.46557 & 2.46561 & 5.30376 & 5.30382 & 30.00000 & 30.00000 \\ 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \end{array} \right\}$	0.99999
$\xi_5 = \left\{ \begin{array}{cccccc} 0.00002 & 0.00016 & 2.46560 & 2.46560 & 5.30378 & 5.30389 & 5.30382 & 30.00000 & 30.00000 \\ 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 & 1/9 \end{array} \right\}$	0.98372

**Tabla 2-6.:** Diseños óptimos de 4, 5, 6, 7, 8 y 9 puntos para el ácido ferúlico

En la tabla **2-6** los diseños tienen pesos asociados  $w_i = 1/N$ , respectivamente, el máximo tiempo de experimentación es 30, porque es la mayor de las concentraciones iniciales consideradas en el experimento. Se observa que los diseños de 5, 6, 7, 8 y 9 concentraciones experimentales pierden eficiencia con respecto al diseño D-óptimo, se observa una mayor reducción en el diseño que tiene 6 tiempos, mientras que en el diseño con 8 tiempos de experimentación se presenta la menor reducción. Se recomiendan los diseños  $\xi_1, \xi_3, \xi_4$  o  $\xi_5$  dependiendo del número de concentraciones que se requieran, puesto que sus eficiencias son superiores al 95%.

### 2.2.5. Regresión no lineal con observaciones correlacionadas

En experimentos prácticos donde las observaciones son tomadas en el mismo individuo se presenta un grado de correlación entre las observaciones, el cual debe ser modelado para hacer un buen análisis del experimento. Es decir, se debe buscar una función que modele la correlación de los datos. En este caso se considera en el modelo 2-1 que la

función  $\eta(t_i, \theta) = f(t_i, \theta)$  es no lineal y  $\varepsilon \sim N(0, \Sigma)$ , donde  $\Sigma$  es la matriz de varianzas y covarianzas del término de error. Para estimar los parámetros del modelo de regresión con estructura de varianzas-covarianzas  $\Sigma$  se utiliza el logaritmo de la verosimilitud, dada por:

$$\log(L(\theta, \Sigma)) = l(\theta, \Sigma) = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log(|\Sigma|) - \frac{1}{2} [(Y - \eta(t, \theta))^T \Sigma^{-1} (Y - \eta(t, \theta))] \quad (2-17)$$

Para un conjunto de observaciones dadas  $y_1, \dots, y_N$ , la log-verosimilitud es una función del vector de parámetros  $\theta$  y de la matriz de varianzas y covarianzas  $\Sigma$ . Los estimadores de máxima verosimilitud de los parámetros, que maximizan la función de verosimilitud, no pueden ser obtenidos explícitamente; éstos deben ser calculados mediante métodos numéricos. Ésto se tratará con mayor detalles en la sección 3.1.1.

A continuación se definen los modelos de compartimientos, con los cuales se van a desarrollar algunos ejemplos presentados en este trabajo.

## 2.3. Modelos Compartimentales

En esta sección se estudian algunos conceptos de farmacocinética y modelos de compartimiento, los cuales son de gran importancia para entender los ejemplos que se desarrollan en este trabajo.

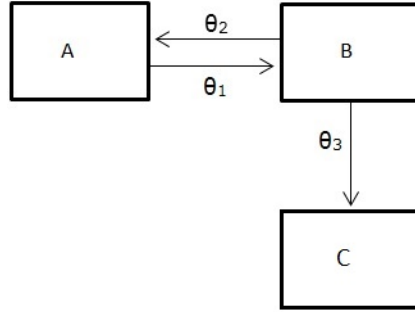
La farmacocinética estudia el modo en que actúan y se comportan los medicamentos en el organismo (Clark & Smith 1989). Hay dos fases asociadas a la cinética de un medicamento, la fase de absorción y la de distribución en el organismo.

La fase de absorción ocurre cuando los medicamentos son transportados por vía oral o intravenosa a la sangre. Por otra parte, la fase de distribución ocurre cuando el medicamento se encuentra en la sangre y penetra a los órganos y tejidos del organismo con el fin de poder actuar. Comprender el comportamiento de un medicamento es muy importante para entender los modelos compartimentales.

Una clase de modelos no lineales de especial importancia, resulta al considerar la respuesta de un fenómeno descrito por un sistema lineal de ecuaciones diferenciales ordinarias como resultado de un análisis compartimental. Un sistema compartimental, se entiende como aquel que está formado por un número finito de subsistemas macroscópicos, llamados compartimientos, cada uno de los cuales es homogéneo y está bien mezclado, además hay intercambio de materiales entre los compartimientos y el medio. Allí pueden haber entradas del medio en uno o más compartimientos y también pueden existir salidas de uno o más compartimientos al medio.

Los modelos de compartimientos lineales son aquellos donde las tasas de transferencia de sustancias del compartimiento  $j$  al compartimiento  $k$  son proporcionales a la cantidad de sustancia que está presente en el compartimiento  $j$ . Supuesto conocido como cinética de primer orden. Estos modelos conducen a sistemas lineales de ecuaciones diferenciales. Los

compartimientos son descritos por cuadrados y el cambio de sustancia de un compartimiento a otro por una flecha, (Bates & Watts 1988). Como se muestra en la figura 2-7.



**Figura 2-7.:** Modelo de tres compartimientos

La figura 2-7 es un modelo de tres compartimientos, con tasa reversible y tasa de transferencia  $\theta_i$ ,  $i = 1, 2, 3$ .

El sistema de ecuaciones diferenciales del modelo es:

$$\frac{dA(t)}{dt} = -\theta_1 A(t) + \theta_2 B(t) \quad (2-18)$$

$$\frac{dB(t)}{dt} = \theta_1 A(t) - (\theta_2 + \theta_3) B(t) \quad (2-19)$$

$$\frac{dC(t)}{dt} = \theta_3 B(t) \quad (2-20)$$

Este sistema de ecuaciones se resuelve por matrices de la forma  $X' = AX$  dando el siguiente resultado:

$$A(t) = \frac{\theta_2}{\theta_1} a_1 + a_2 \frac{\theta_2 \lambda_2}{(\theta_1 + \lambda_2) \theta_3} e^{\lambda_2 t} + a_3 \frac{\theta_2 \lambda_3}{(\theta_1 + \lambda_3) \theta_3} e^{\lambda_3 t} \quad (2-21)$$

$$B(t) = a_1 + a_2 \frac{\lambda_2}{\theta_3} e^{\lambda_2 t} + a_3 \frac{\lambda_3}{\theta_3} e^{\lambda_3 t} \quad (2-22)$$

$$C(t) = a_2 e^{\lambda_2 t} + a_3 e^{\lambda_3 t} \quad (2-23)$$

Donde,  $\lambda_1 = 0$ ,  $\lambda_2 = \frac{-(\theta_1 + \theta_2 + \theta_3) + \sqrt{(\theta_1 + \theta_2 + \theta_3)^2 - 4\theta_1 \theta_3}}{2}$  y  $\lambda_3 = \frac{-(\theta_1 + \theta_2 + \theta_3) - \sqrt{(\theta_1 + \theta_2 + \theta_3)^2 - 4\theta_1 \theta_3}}{2}$  son los valores propios que resultan de la ecuación  $\lambda[\lambda^2 + \lambda(\theta_1 + \theta_2 + \theta_3) + \theta_1 \theta_3] = 0$ , y  $a_1, a_2$  y  $a_3$  están dadas por:

$$a_1 = \frac{(\lambda_2 - \lambda_3)}{\theta_2 \left[ \frac{(\lambda_2 - \lambda_3)}{\theta_1} - \frac{\lambda_2}{(\theta_1 + \lambda_2)} + \frac{\lambda_3}{(\theta_1 + \lambda_3)} \right]}$$
$$a_2 = -\frac{\theta_3}{\theta_2 \left[ \frac{(\lambda_2 - \lambda_3)}{\theta_1} - \frac{\lambda_2}{(\theta_1 + \lambda_2)} + \frac{\lambda_3}{(\theta_1 + \lambda_3)} \right]}$$
$$a_3 = \frac{\theta_3}{\theta_2 \left[ \frac{(\lambda_2 - \lambda_3)}{\theta_1} - \frac{\lambda_2}{(\theta_1 + \lambda_2)} + \frac{\lambda_3}{(\theta_1 + \lambda_3)} \right]}$$

En este trabajo la función  $\eta$  de 2-1 se considera no lineal y proviene de un modelo compartimental, con dos componentes y tasas reversible de transferencia (Allen 1983).



# 3. Diseños óptimos para observaciones correlacionadas

En este capítulo se presenta el desarrollo teórico-práctico de la estimación de los parámetros del modelo, para la regresión no lineal con observaciones correlacionadas, la matriz de información para observaciones correlacionadas y la función de covarianza de Matérn que se utiliza para modelar la estructura de correlación del error. También se muestran las respectivas estructuras de correlación que se utilizan para desarrollar el modelo propuesto, se hallan los respectivos diseños óptimos y se muestran algunas aplicaciones.

## 3.1. Descripción del modelo

En el capítulo anterior se estudió de forma general la teoría de los diseños óptimos, en esta sección se muestra la terminología cuando las observaciones están correlacionadas.

Cuando las observaciones están correlacionadas los errores aleatorios del modelo 2-1 se distribuyen normal multivariado con vector de medias cero y matriz de varianzas y covarianzas  $\Sigma$ , esto es:  $\varepsilon \sim N_N(0, \Sigma)$ , donde  $\Sigma = Cov(\varepsilon) = [E(\varepsilon(t_i)\varepsilon(t_j))]_{N \times N} = \sigma^2 A$ , y la matriz  $A$  es de la forma:

$$A = \begin{pmatrix} a(t_1, t_1) & a(t_1, t_2) & \cdots & a(t_1, t_N) \\ a(t_2, t_1) & a(t_2, t_2) & \cdots & a(t_2, t_N) \\ \vdots & \vdots & \ddots & \vdots \\ a(t_N, t_1) & a(t_N, t_2) & \cdots & a(t_N, t_N) \end{pmatrix}. \quad (3-1)$$

En este trabajo se consideran diseños exactos, sin considerar réplicas, lo que indica que cada medición se debe tomar en una corrida del modelo 2-1, es decir, que se requiere  $t_i \neq t_j$  para  $i, j = 1, 2, \dots, N$  y  $i \neq j$ , se escribe  $\xi = \{t_1, \dots, t_N\}$

Algunos autores han utilizado estructuras de correlación en los errores, tales como: Dette et al. (2008) consideran que los errores están centrados y provienen de un proceso autorregresivo estacionario, donde la correlación entre dos mediciones depende de la distancia entre ellas. Los autores modelaron la estructura de covarianza con  $Cov(\varepsilon(t_i), \varepsilon(t_j)) = \sigma^2 \lambda^{|t_i - t_j|}$  ( $\lambda$  es una constante conocida), mostraron que el diseño igualmente espaciado suele ser muy eficiente, incluso para tamaños moderados de la correlación, mientras que el diseño D-óptimo obtenido bajo los supuestos de observaciones independientes produce una pérdida

sustancial de la eficiencia. Dette et al. (2013) modelaron la estructura del error para modelos lineales con la función de covarianza tipo kernel  $K(t_i, t_j) = \sigma^2 h(t_i - t_j)$ , mostraron que la distribución uniforme es óptima para una clase de modelos de regresión trigonométricos con una amplia clase de covarianza tipo kernel y que la distribución arco seno es óptimo para el modelo de regresión polinómica con estructura de correlación definida por la potencia logarítmica. Patan & Bogacka (2007), proponen un procedimiento numérico para determinar las aproximaciones de la matriz de información del diseño óptimo cuando los errores están correlacionados. Por otra parte, Amo et al. (2012), estudiaron modelos de dos y tres compartimientos, los cuales describen la administración oral de un medicamento. Luego, modelaron la correlación del error asumiendo una función de covarianza exponencial  $\sigma^2 \exp(-\rho|t_i - t_j|)$ ; con esta estructura, calcularon diseños óptimos exactos bajo diferentes criterios, tanto para hacer discriminación entre dos modelos como para la estimación de los parámetros de los modelos considerados. Además, utilizaron criterios compuestos para estimar los parámetros y funciones no lineales de éstos.

### 3.1.1. Estimación por máxima verosimilitud

Para el cálculo de la función de verosimilitud se tiene en cuenta que los errores se distribuyen normal multivariado, esto es  $\varepsilon \sim N_N(0, \Sigma)$ , donde  $\Sigma = Cov(\varepsilon) = [E(\varepsilon(t_i)\varepsilon(t_j))]_{N \times N}$ , donde la función de densidad de probabilidad conjunta está dada por:

$$f_Y(y_1, \dots, y_N; \theta; \Sigma) = \frac{1}{(2\pi)^{N/2} |\Sigma|^{1/2}} \exp \left[ -\frac{1}{2} (Y - \eta(\xi, \theta))^T \Sigma^{-1} (Y - \eta(\xi, \theta)) \right], \quad (3-2)$$

donde  $\eta(\xi, \theta)^T = (\eta(t_1, \theta) \dots \eta(t_N, \theta))$ . Ahora, con 3-2 se calcula la verosimilitud, de la siguiente manera:

$$L(\theta, \Sigma) = (2\pi)^{-N/2} |\Sigma|^{-1/2} \exp \left[ -\frac{1}{2} (Y - \eta(\xi, \theta))^T \Sigma^{-1} (Y - \eta(\xi, \theta)) \right]$$

Luego, el logaritmo natural de la verosimilitud es:

$$l(\theta, \Sigma) = \log[L(\theta, \Sigma)] = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log(|\Sigma|) - \frac{1}{2} [(Y - \eta(\xi, \theta))^T \Sigma^{-1} (Y - \eta(\xi, \theta))] \quad (3-3)$$

Para un conjunto de observaciones dadas  $y_1, \dots, y_N$ , la log-verosimilitud es una función de los parámetros. Si  $p$  es la dimensión de  $\theta$  y  $q$  es la dimensión de  $\Sigma$ , el modelo depende de  $p + q$  parámetros. Los estimadores de máxima verosimilitud  $\hat{\theta}, \hat{\Sigma}$  que maximizan la función de verosimilitud no se pueden obtener explícitamente; estos deben ser obtenidos mediante métodos numéricos. En este trabajo se utilizan funciones de  $R$  tales como: *DEoptim*, *nls2*,



*nlminb*, *nls*, que permitan estimar de manera eficiente los parámetros del modelo, que garantizan maximizar la función log-verosimilitud.

Cuando la matriz de varianzas y covarianzas  $\Sigma$  es de la forma  $\Sigma = \sigma^2 A_\rho$ , donde  $A_\rho$  es una matriz que depende de  $t_1, \dots, t_N$  y de un parámetro  $\rho$ , se presenta el siguiente caso particular de la función de log-verosimilitud:

$$l(\theta, \sigma^2, \rho) = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2 |A_\rho|) - \frac{1}{2\sigma^2} [(Y - \eta(\xi, \theta))^T A_\rho^{-1} (Y - \eta(\xi, \theta))] \quad (3-4)$$

Note que los parámetros estimados  $\hat{\theta}, \hat{\sigma}^2, \hat{\rho}$  que maximizan 3-4 no se pueden obtener hallando los puntos críticos de la función dada en 3-4. Es decir, que la estimación de los parámetros se obtiene por métodos numéricos, en particular mediante el uso de las funciones del  $R$  mencionadas anteriormente.

La matriz de información para el diseño  $\xi$  se muestra en el siguiente teorema.

### 3.1.2. Matriz de información

La matriz de información de Fisher para cada diseño  $\xi$  se define como:

$$M(\xi, \Theta) = E \left( -\frac{\partial^2 \log[L(\Theta, \xi)]}{\partial \Theta \partial \Theta^T} \right), \quad (3-5)$$

y es posible mostrar que 3-5 es una matriz de bloques, ésto se ilustra en el teorema 3.1.

**Teorema 3.1.** *Considere el modelo*

$$Y = \eta(\xi, \theta) + \varepsilon, \quad (3-6)$$

donde  $\varepsilon \sim N_N(0, \Sigma_\beta)$ ,  $\beta$  es un vector con  $q$  componentes. Sea  $\xi = \{t_1, t_2, \dots, t_N\}$  un diseño exacto de  $N$ -puntos, entonces la matriz de información de Fisher asociada a  $\Theta^T = (\theta^T, \beta^T)$  está dada por:

$$M(\xi, \theta, \beta) = \begin{pmatrix} \left[ E \left( -\frac{\partial^2 l(\theta, \beta; \xi)}{\partial \theta_i \partial \theta_j} \right) \right]_{p \times p} & 0 \\ 0 & \left[ E \left( -\frac{\partial^2 l(\theta, \beta; \xi)}{\partial \beta_k \partial \beta_l} \right) \right]_{q \times q} \end{pmatrix} = \begin{pmatrix} M_{11} & \mathbf{0} \\ \mathbf{0} & M_{22} \end{pmatrix} \quad (3-7)$$

donde

$$M_{11} = F^T(\theta) \Sigma_\beta^{-1} F(\theta)$$

$$M_{22} = [m_{kl}]_{q \times q} = \left[ \frac{1}{2} \text{tr} \left( \Sigma_\beta^{-1} \frac{\partial \Sigma_\beta}{\partial \beta_k} \Sigma_\beta^{-1} \frac{\partial \Sigma_\beta}{\partial \beta_l} \right) \right]_{q \times q},$$

$$y F(\theta) = \left[ \frac{\partial \eta(\xi, \theta)}{\partial \theta} \right]_{N \times p} = \left[ \frac{\partial \eta(t_i, \theta)}{\partial \theta_j} \right]_{N \times p}$$

### Prueba

Note que para  $i, j = 1, \dots, p$ :

$$-\frac{\partial^2 l(\theta, \beta; \xi)}{\partial \theta_i \partial \theta_j} = \frac{\partial \eta(\xi, \theta)}{\partial \theta_i} \Sigma_\beta^{-1} \frac{\partial \eta(\xi, \theta)}{\partial \theta_j} - [Y - \eta(\xi, \theta)]^T \Sigma_\beta^{-1} \frac{\partial \eta^2(\xi, \theta)}{\partial \theta_i \theta_j} \quad (3-8)$$

y  $k, l = 1, \dots, q$ :

$$\begin{aligned} -\frac{\partial^2 l(\theta, \beta; \xi)}{\partial \beta_k \partial \beta_l} &= [Y - \eta(\xi, \theta)]^T \Sigma_\beta^{-1} \frac{\partial \Sigma_\beta}{\partial \beta_k} \Sigma_\beta^{-1} \frac{\partial \Sigma_\beta}{\partial \beta_l} \Sigma_\beta^{-1} [Y - \eta(\xi, \theta)] \\ &\quad - \frac{1}{2} [Y - \eta(\xi, \theta)]^T \Sigma_\beta^{-1} \frac{\partial^2 \Sigma_\beta}{\partial \beta_k \partial \beta_l} \Sigma_\beta^{-1} [Y - \eta(\xi, \theta)] \\ &\quad + \frac{1}{2} \text{tr} \left( \Sigma_\beta^{-1} \frac{\partial^2 \Sigma_\beta}{\partial \beta_k \partial \beta_l} \right) - \frac{1}{2} \text{tr} \left( \Sigma_\beta^{-1} \frac{\partial \Sigma_\beta}{\partial \beta_k} \Sigma_\beta^{-1} \frac{\partial \Sigma_\beta}{\partial \beta_l} \right) \end{aligned} \quad (3-9)$$

Calculando el valor esperado de 3-8 y 3-9 se obtienen las componentes de  $M_{11}$  y  $M_{22}$ .

**Corolario 3.1.1.** *Si en el modelo 3-6,  $\Sigma_\beta = \sigma^2 A_\rho$ , entonces  $M_{11}$  y  $M_{22}$  están dadas por:*

$$\begin{aligned} M_{11} &= \frac{1}{\sigma^2} F^T(\theta) A_\rho^{-1} F(\theta) \\ M_{22} &= \begin{pmatrix} \frac{1}{2\sigma^4} N & \frac{1}{\sigma^2} \text{tr} \left( A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right) \\ \frac{1}{\sigma^2} \text{tr} \left( A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right) & \frac{1}{2} \text{tr} \left( A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right) \end{pmatrix} \end{aligned}$$

**Definición:** Criterio de optimalidad

Si se cumplen las condiciones del teorema 3.1, se define el siguiente criterio para estimar de forma óptima  $\theta$  y  $\beta$

$$\psi(\xi, \theta, \beta) = |M(\xi, \theta, \beta)|^{1/(p+q)} = \{|M_{11}(\xi, \theta)| |M_{22}(\xi, \beta)|\}^{1/(p+q)} \quad (3-10)$$

Un diseño  $\xi^*$  es D-óptimo si maximiza 3-10, ésto es equivalente a obtener un diseño que minimice:

$$-\log |M(\xi, \theta, \beta)| = -\log |M_{11}(\xi, \theta)| - \log |M_{22}(\xi, \beta)| \quad (3-11)$$

En este trabajo  $\Sigma_\beta = \sigma^2 A_\rho$ . Entonces, 3-11 puede escribirse como:

$$-\log |M(\xi; \theta, \sigma^2, \rho)| = -\log |M_{11}(\xi, \theta)| - \log \left| \frac{1}{4\sigma^4} N^2 S_\lambda^2 \right|, \quad (3-12)$$

donde  $S_\lambda^2 = \frac{1}{N} \sum_{j=1}^N (\lambda_j - \bar{\lambda})^2$ ,  $\lambda_j$  es el j-ésimo valor propio de la matriz  $\left( A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right)$ . Para probar 3-12, basta con mostrar que  $\log |M_{22}| = \log \left| \frac{1}{4\sigma^4} N^2 S_\lambda^2 \right|$ . la fórmula general de  $M_{22}$  se puede ver en Pazman (2007), donde:

$$M_{22} = E \left( -\frac{\partial^2 \log[L(\theta, \beta)]}{\partial \beta \partial \beta^T} \right) = E \left[ - \left( \begin{array}{cc} \frac{\partial^2 \log[L(\theta, \sigma^2, \rho)]}{\partial \sigma^2 \partial \sigma^2} & \frac{\partial^2 \log[L(\theta, \sigma^2, \rho)]}{\partial \sigma^2 \partial \rho} \\ \frac{\partial^2 \log[L(\theta, \sigma^2, \rho)]}{\partial \rho \partial \sigma^2} & \frac{\partial^2 \log[L(\theta, \sigma^2, \rho)]}{\partial \rho \partial \rho} \end{array} \right) \right]$$

Entonces,

i.

$$\begin{aligned} E \left[ -\frac{\partial^2 \log[L(\theta, \sigma^2, \rho)]}{\partial \sigma^2 \partial \sigma^2} \right] &= E \left[ (Y - \eta(\xi, \theta))^T (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \sigma^2} (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \sigma^2} (\sigma^2 A_\rho)^{-1} (Y - \eta(\xi, \theta)) \right] \\ &\quad - \frac{1}{2} \text{tr} \left[ (\sigma^2 A_\rho)^{-1} \frac{\partial^2 \sigma^2 A_\rho}{\partial \sigma^2} (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \sigma^2} \right] \\ &= E \left[ (Y - \eta(\xi, \theta))^T (\sigma^2)^{-1} A_\rho^{-1} A_\rho (\sigma^2)^{-1} A_\rho^{-1} A_\rho (\sigma^2)^{-1} A_\rho^{-1} (Y - \eta(\xi, \theta)) \right] \\ &\quad - \frac{1}{2} \text{tr} \left[ (\sigma^2)^{-1} A_\rho^{-1} A_\rho (\sigma^2)^{-1} A_\rho^{-1} A_\rho \right] \\ &= \frac{1}{(\sigma^2)^3} E \left[ (Y - \eta(\xi, \theta))^T A_\rho^{-1} (Y - \eta(\xi, \theta)) \right] - \frac{1}{2(\sigma^2)^2} N \\ &= \frac{1}{\sigma^6} \text{tr} (A_\rho^{-1} \sigma^2 A_\rho) - \frac{1}{2\sigma^4} N \\ &= \frac{1}{\sigma^4} N - \frac{1}{2\sigma^4} N \\ &= \frac{1}{2\sigma^4} N \end{aligned}$$

ii.

$$\begin{aligned} E \left[ -\frac{\partial^2 \log[L(\theta, \sigma^2, \rho)]}{\partial \rho \partial \sigma^2} \right] &= E \left[ (Y - \eta(\xi, \theta))^T (\sigma^2 A_\rho)^{-1} \sigma^2 \frac{\partial A_\rho}{\partial \rho} (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \sigma^2} (\sigma^2 A_\rho)^{-1} (Y - \eta(\xi, \theta)) \right] \\ &\quad - \frac{1}{2} E \left[ (Y - \eta(\xi, \theta))^T (\sigma^2 A_\rho)^{-1} \frac{\partial^2 \sigma^2 A_\rho}{\partial \rho \partial \sigma^2} (\sigma^2 A_\rho)^{-1} (Y - \eta(\xi, \theta)) \right] \\ &\quad + \frac{1}{2} \text{tr} \left[ (\sigma^2 A_\rho)^{-1} \frac{\partial^2 \sigma^2 A_\rho}{\partial \sigma^2 \partial \rho} \right] - \frac{1}{2} \text{tr} \left[ (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \rho} (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \sigma^2} \right] \\ &= E \left[ (Y - \eta(\xi, \theta))^T (\sigma^2)^{-3} A_\rho^{-1} \sigma^2 \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} A_\rho A_\rho^{-1} (Y - \eta(\xi, \theta)) \right] \\ &\quad - \frac{1}{2} E \left[ (Y - \eta(\xi, \theta))^T (\sigma^2 A_\rho)^{-1} \frac{\partial^2 \sigma^2 A_\rho}{\partial \rho \partial \sigma^2} (\sigma^2 A_\rho)^{-1} (Y - \eta(\xi, \theta)) \right] \\ &\quad + \frac{1}{2} \text{tr} \left[ (\sigma^2 A_\rho)^{-1} \frac{\partial^2 \sigma^2 A_\rho}{\partial \sigma^2 \partial \rho} \right] - \frac{1}{2} \text{tr} \left[ (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \rho} (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \sigma^2} \right] \\ &= (\sigma^2)^{-2} E \left[ (Y - \eta(\xi, \theta))^T A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} (Y - \eta(\xi, \theta)) \right] \\ &\quad - \frac{1}{2} (\sigma^2)^{-2} E \left[ (Y - \eta(\xi, \theta))^T A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} (Y - \eta(\xi, \theta)) \right] \\ &\quad + \frac{1}{2} (\sigma^2)^{-1} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] - \frac{1}{2} (\sigma^2)^{-1} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} A_\rho \right] \\ &= \frac{1}{\sigma^4} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \sigma^2 A_\rho \right] - \frac{1}{2\sigma^4} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \sigma^2 A_\rho \right] \\ &= \frac{1}{\sigma^2} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] - \frac{1}{2\sigma^2} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] \\ &= \frac{1}{\sigma^2} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] \end{aligned}$$

iii.

$$\begin{aligned}
E \left[ -\frac{\partial^2 \log[L(\theta, \sigma^2, \rho)]}{\partial \rho \partial \rho} \right] &= E \left[ (Y - \eta(\xi, \theta))^T (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \rho} (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \rho} (\sigma^2 A_\rho)^{-1} (Y - \eta(\xi, \theta)) \right] \\
&\quad - \frac{1}{2} E \left[ (Y - \eta(\xi, \theta))^T (\sigma^2 A_\rho)^{-1} \frac{\partial^2 \sigma^2 A_\rho}{\partial \rho \partial \rho} (\sigma^2 A_\rho)^{-1} (Y - \eta(\xi, \theta)) \right] \\
&\quad + \frac{1}{2} \text{tr} \left[ (\sigma^2 A_\rho)^{-1} \frac{\partial^2 \sigma^2 A_\rho}{\partial \rho \partial \rho} \right] - \frac{1}{2} \text{tr} \left[ (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \rho} (\sigma^2 A_\rho)^{-1} \frac{\partial \sigma^2 A_\rho}{\partial \rho} \right] \\
&= \frac{1}{\sigma^2} E \left[ (Y - \eta(\xi, \theta))^T A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} (Y - \eta(\xi, \theta)) \right] \\
&\quad - \frac{1}{2\sigma^2} E \left[ (Y - \eta(\xi, \theta))^T A_\rho^{-1} \frac{\partial^2 A_\rho}{\partial \rho \partial \rho} A_\rho^{-1} (Y - \eta(\xi, \theta)) \right] \\
&\quad + \frac{1}{2} \text{tr} \left[ A_\rho^{-1} \frac{\partial^2 A_\rho}{\partial \rho \partial \rho} \right] - \frac{1}{2} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] \\
&= \frac{1}{\sigma^2} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \sigma^2 A_\rho \right] - \frac{1}{2\sigma^2} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \sigma^2 A_\rho \right] \\
&\quad + \frac{1}{2} \text{tr} \left[ A_\rho^{-1} \frac{\partial^2 A_\rho}{\partial \rho \partial \rho} \right] - \frac{1}{2} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] \\
&= \frac{1}{2} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right]
\end{aligned}$$

Luego,

$$\begin{aligned}
|M_{22}| &= \left| \begin{pmatrix} \frac{1}{2\sigma^4} N & \frac{1}{\sigma^2} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] \\ \frac{1}{\sigma^2} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] & \frac{1}{2} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] \end{pmatrix} \right| \\
&= \frac{1}{4\sigma^4} N \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] - \frac{1}{4\sigma^4} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] \\
&= \frac{1}{4\sigma^4} \left[ N \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] - \frac{1}{4\sigma^4} \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] \text{tr} \left[ A_\rho^{-1} \frac{\partial A_\rho}{\partial \rho} \right] \right] \\
&= \frac{1}{4\sigma^4} \left[ N \sum_{i=1}^N \lambda_i^2 - \left( \sum_{i=1}^N \lambda_i \right)^2 \right] \\
&= \frac{1}{4\sigma^4} \left[ N \sum_{i=1}^N \lambda_i^2 - (N\bar{\lambda})^2 \right] \\
&= \frac{1}{4\sigma^4} N \left[ \sum_{i=1}^N \lambda_i^2 - N(\bar{\lambda})^2 \right] \\
&= \frac{1}{4\sigma^4} N \sum_{i=1}^N (\lambda_i - \bar{\lambda})^2 \\
&= \frac{1}{4\sigma^4} N^2 S_\lambda^2
\end{aligned}$$

Por tanto,  $\log |M_{22}| = \log \left| \frac{1}{4\sigma^4} N^2 S_\lambda^2 \right|$ .

A continuación se referencia la función de covarianza de Matérn, la cual será de utilidad para modelar las estructuras de correlación del error que se usarán en el desarrollo de este trabajo, en particular, problemas donde la variable regresora es el tiempo, y las mediciones son tomadas en el mismo individuo o unidad experimental.

## 3.2. Función de covarianza Matérn

La función de covarianza Matérn (Matern 1960) se utiliza por lo general en las áreas de estadística espacial, geoestadística, aprendizaje automático, análisis de imágenes y otras aplicaciones de análisis estadístico multivariable en espacios métricos. Se utiliza comúnmente para definir la covarianza estadística entre las mediciones realizadas en dos puntos que son  $d$  unidades distantes el uno del otro. Si la distancia es la distancia euclidiana, la función de covarianza de Matérn también es isotrópica. Esta propiedad se refiere al hecho de que ciertas magnitudes vectoriales conmensurables, es decir, la razón entre estas magnitudes es un número racional, dan resultados idénticos independientes de la dirección escogida para dicha medida.

La función de covarianza de Matérn entre dos puntos separados por  $d$  unidades de distancia está dada por:

$$F(d; \sigma^2, \rho, \nu) = \sigma^2 \frac{1}{\Gamma(\nu) 2^{\nu-1}} \left( \sqrt{2\nu} \frac{d}{\rho} \right)^\nu K_\nu \left( \sqrt{2\nu} \frac{d}{\rho} \right), \quad (3-13)$$

donde  $\Gamma$  es la función gamma,  $K_\nu$  es la función de Bessel modificada de segunda clase,  $\rho$  es el parámetro de rango o distancia que mide la rapidez con la que decaen las correlaciones,  $\nu$  controla el grado de suavidad y  $d$  la distancia entre un tiempo y otro. Además,  $\rho$  y  $\nu$  son parámetros no negativos de la covarianza (Minasny & McBratney 2005). En este trabajo se construyen diseños cuando la estructura de covarianza del error se puede modelar por la función de covarianza de Matérn para los siguientes tres casos particulares:

1.  $\nu = 1/2$ ,  $F(d; \sigma^2, \rho) = \sigma^2 \exp\left(-\frac{d}{\rho}\right)$
2.  $\nu \rightarrow \infty$ ,  $F(d; \sigma^2, \rho) = \sigma^2 \exp(-d^2/2\rho^2)$
3.  $\nu = \frac{3}{2}$ ,  $F(d; \sigma^2, \rho) = \sigma^2 \left(1 + \frac{\sqrt{3}d}{\rho}\right) \exp\left(-\frac{\sqrt{3}d}{\rho}\right)$

La derivada para cada caso con respecto a  $\rho$  es:

1.  $\frac{\partial F(d; \sigma^2, \rho)}{\partial \rho} = \sigma^2 \exp\left(-\frac{d}{\rho}\right) \left(\frac{d}{\rho^2}\right)$
2.  $\frac{\partial F(d; \sigma^2, \rho)}{\partial \rho} = \sigma^2 \exp(-d^2/2\rho^2) \left(\frac{d^2}{\rho^3}\right)$
3.  $\frac{\partial F(d; \sigma^2, \rho)}{\partial \rho} = \sigma^2 \exp\left(-\frac{\sqrt{3}d}{\rho}\right) \left(\frac{3d^2}{\rho^3}\right)$

A continuación se definen las estructuras de correlación en forma matricial, con las cuales se van a desarrollar algunos ejemplos presentados en este trabajo.

### 3.2.1. Estructuras de correlación

Para calcular los puntos del diseño óptimo del modelo, se necesita hallar los valores de los parámetros que maximizan la función de verosimilitud. Para ésto, 2-1 se puede escribir en forma matricial como:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}_{N \times 1} = \begin{bmatrix} \eta(t_1, \theta) \\ \eta(t_2, \theta) \\ \vdots \\ \eta(t_N, \theta) \end{bmatrix}_{N \times 1} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}_{N \times 1} = \eta(\xi, \theta) + \varepsilon \quad (3-14)$$

donde  $\varepsilon \sim N_N(0, \Sigma)$  y  $Y \sim N_N(\eta(\xi, \theta), \Sigma)$ . Para cada estructura de correlación de 3-13 se calcula la matriz  $\Sigma$ , la cual depende de los parámetros  $\sigma^2$  y  $\rho$ .

La matriz de varianzas y covarianzas para cada caso de la función de covarianza de Matérn considerados en este trabajo son:

1. Caso 1:

$$\Sigma_{\sigma^2, \rho} = \sigma^2 A_\rho = \sigma^2 \begin{pmatrix} 1 & \exp(-|t_1 - t_2|/\rho) & \cdots & \exp(-|t_1 - t_N|/\rho) \\ \exp(-|t_2 - t_1|/\rho) & 1 & \cdots & \exp(-|t_2 - t_N|/\rho) \\ \vdots & \vdots & \ddots & \vdots \\ \exp(-|t_N - t_1|/\rho) & \exp(-|t_N - t_2|/\rho) & \cdots & 1 \end{pmatrix}$$

2. Caso 2:

$$\Sigma_{\sigma^2, \rho} = \sigma^2 A_\rho = \sigma^2 \begin{pmatrix} 1 & \exp(-|t_1 - t_2|^2/2\rho^2) & \cdots & \exp(-|t_1 - t_N|^2/2\rho^2) \\ \exp(-|t_2 - t_1|^2/2\rho^2) & 1 & \cdots & \exp(-|t_2 - t_N|^2/2\rho^2) \\ \vdots & \vdots & \ddots & \vdots \\ \exp(-|t_N - t_1|^2/2\rho^2) & \exp(-|t_N - t_2|^2/2\rho^2) & \cdots & 1 \end{pmatrix}$$

3. Caso 3:

$$\Sigma_{\sigma^2, \rho} = \sigma^2 A_\rho = \sigma^2 \begin{pmatrix} 1 & \cdots & \left(1 + \frac{\sqrt{3}|t_1 - t_N|}{\rho}\right) \exp\left(-\frac{\sqrt{3}|t_1 - t_N|}{\rho}\right) \\ \left(1 + \frac{\sqrt{3}|t_2 - t_1|}{\rho}\right) \exp\left(-\frac{\sqrt{3}|t_2 - t_1|}{\rho}\right) & \cdots & \left(1 + \frac{\sqrt{3}|t_2 - t_N|}{\rho}\right) \exp\left(-\frac{\sqrt{3}|t_2 - t_N|}{\rho}\right) \\ \vdots & \ddots & \vdots \\ \left(1 + \frac{\sqrt{3}|t_N - t_1|}{\rho}\right) \exp\left(-\frac{\sqrt{3}|t_N - t_1|}{\rho}\right) & \cdots & 1 \end{pmatrix}$$

A cada una de estas matrices se les calcula la inversa, para obtener a  $\Sigma^{-1}(\xi)$ , la cual se va a utilizar posteriormente para calcular la matriz de información de Fisher y seguidamente su determinante.

### 3.3. Metodología

La metodología propuesta para calcular los diseños cuasi-óptimos, cuando la matriz de covarianza del vector de errores tiene una estructura particular, modelada con la función de covarianza Matérn es la siguiente: Inicialmente se realiza la estimación de los parámetros del modelo para observaciones incorrelacionadas y los valores del parámetro  $\rho$ ; graficando el determinante de la matriz  $A$  contra cada valor de  $\rho$ , usados posteriormente para calcular el diseño  $\xi_{OI}$  cuasi-óptimo para observaciones incorrelacionadas y como valores iniciales para calcular los parámetros del diseños  $\xi_{OC}$  cuasi-óptimos con observaciones correlacionadas utilizando los tres casos de la función de Matérn mencionados en la sección anterior y los diseños respectivos. Seguidamente, se calcula la eficiencia del diseño  $\xi_{OI}$  con respecto al diseño  $\xi_{OC}$ , para cada valor de  $\rho$ , ésta es graficada para los tres casos de la función de covarianza de Matérn. Luego, se calcula el  $AIC$  y  $BIC$  para los diseños  $\xi_{OC}$  con cada caso y se selecciona el diseño con menor criterio de información. Por último, se realiza un estudio de simulación con el diseño con menor criterio de información; esta comparación se hace para verificar vía simulación si el diseño  $\xi_{OC}$  hallado con la respectiva estructura de correlación, estima mejor los parámetros del modelo, que el diseño  $\xi_{OI}$ .

### 3.4. Aplicación 1: Concentración de la Tetraciclina

La tetraciclina es usada para tratar las infecciones provocadas por bacterias, incluyendo la neumonía y otras infecciones en las vías respiratorias; el acné; infecciones en la piel, los genitales y el sistema urinario; y la infección que causa úlceras estomacales. Los datos fueron tomados de Bates & Watts (1988) p. 281.

Un compuesto de tetraciclina fue administrado oralmente a un individuo y se registró la concentración de hidrocortido de tetraciclina en suero sanguíneo (medida en microgramos por mililitro (mcg/ml)) durante un periodo de 16 horas. Se usó un modelo para 2 compartimientos. Los datos se presentan en la tabla **3-1**:

Tiempo (horas)	1	2	3	4	6	8	10	12	16
Concentración (mcg/ml)	0.7	1.2	1.4	1.4	1.1	0.8	0.6	0.5	0.3

**Tabla 3-1.:** Valores de las concentraciones de tetraciclina para cada tiempo

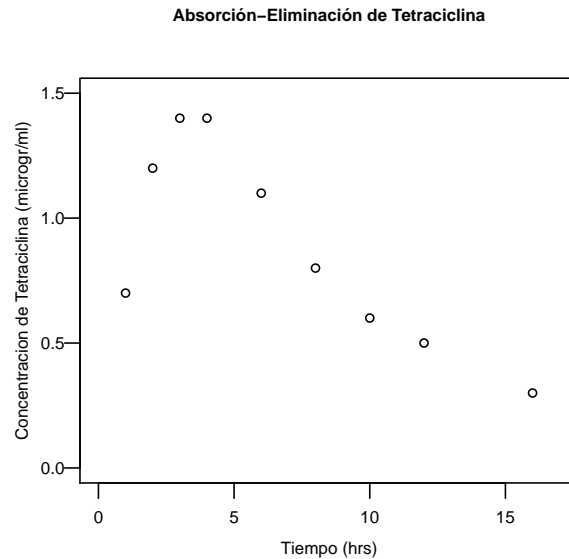
El modelo para estos datos se muestra en 2-1 y la función  $\eta$  está dada por:

$$\eta(t_i, \theta) = \theta_3[\exp(-\theta_1 t_i) - \exp(-\theta_2 t_i)] \quad (3-15)$$

Donde,  $\theta_i > 0$  para  $i = 1, 2, 3$  y  $\theta_1 < \theta_2$ ,  $\eta(t_i, \theta)$  es una función no lineal que proviene de la solución del sistema de ecuaciones diferenciales ordinarias y describe la relación entre la concentración de tetraciclina y el tiempo  $t_i$ , donde se registró la concentración y los  $\varepsilon_i$

tienen media 0, matriz de covarianza  $\Sigma$  y son dependientes porque las medidas son tomadas en el mismo individuo.

En la figura **3-1** se grafican las concentraciones de tetraciclina contra el tiempo, en ésta se observa que la mayor concentración de tetraciclina se presenta en el intervalo [3-5].



**Figura 3-1.:** Gráfica de concentración de tetraciclina vs tiempo

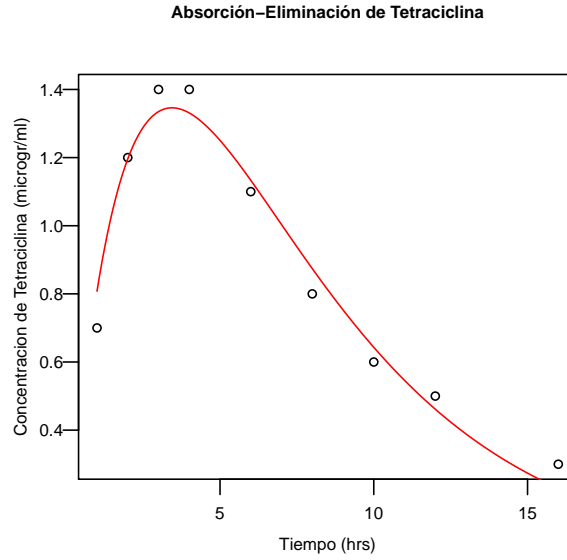
Para hacer el ajuste del modelo con observaciones correlacionadas, primero se estiman los valores iniciales de los parámetros utilizando los métodos para observaciones independientes mencionados en el capítulo 2, esta estimación de los valores iniciales se muestra en la siguiente sección.

### 3.4.1. Estimación de valores iniciales (tetraciclina)

Para estimar los parámetros del modelo 2-1, se realizó un análisis exploratorio donde se obtuvieron los siguientes valores iniciales:  $\theta_1 = 0.2$ ,  $\theta_2 = 1.5$  y  $\theta_3 = 2$ . Con estos valores se estiman los parámetros del modelo para observaciones incorrelacionadas utilizando la función *nls* del *R*, la cual arrojó el siguiente vector de parámetros estimado  $\hat{\theta}^T = (0.183 \ 0.435 \ 4.364)^T$ . Estos valores se grafican en la figura **3-2**; donde se observa que la curva con estos valores pasa cerca de los datos, indicando que la función  $\eta$  es adecuada para modelar los datos.

Los valores estimados para observaciones incorrelacionadas se utilizan como valores iniciales para la estimación de los parámetros del modelo con observaciones correlacionadas y encontrar los diseños cuasi-óptimos respectivos.





**Figura 3-2.:** Gráfica de parámetros estimados de la concentración de tetraciclina vs tiempo

### 3.4.2. Búsqueda y comparación de diseños cuasi-óptimos para observaciones correlacionadas (tetraciclina)

En esta sección se buscan los diseños cuasi-óptimos para cada caso de la función de covarianza de Matérn, estos son comparados con el diseño cuasi-óptimo para observaciones incorrelacionadas por medio de la eficiencia para cada valor de  $\rho$ . También se selecciona el diseño que mejor se ajusta a los datos utilizando los criterios de información *AIC* y *BIC*. Para hallar los diseños, primero se calcula el determinante de la matriz de información de Fisher, pero para esto se deben calcular  $F(\xi; \theta) = \left[ \frac{\partial \eta(t_i, \theta)}{\partial \theta_j} \right]_{N \times p}$ ; las derivadas de la función  $\eta(t_i, \theta)$  con respecto a cada parámetro, esto es:

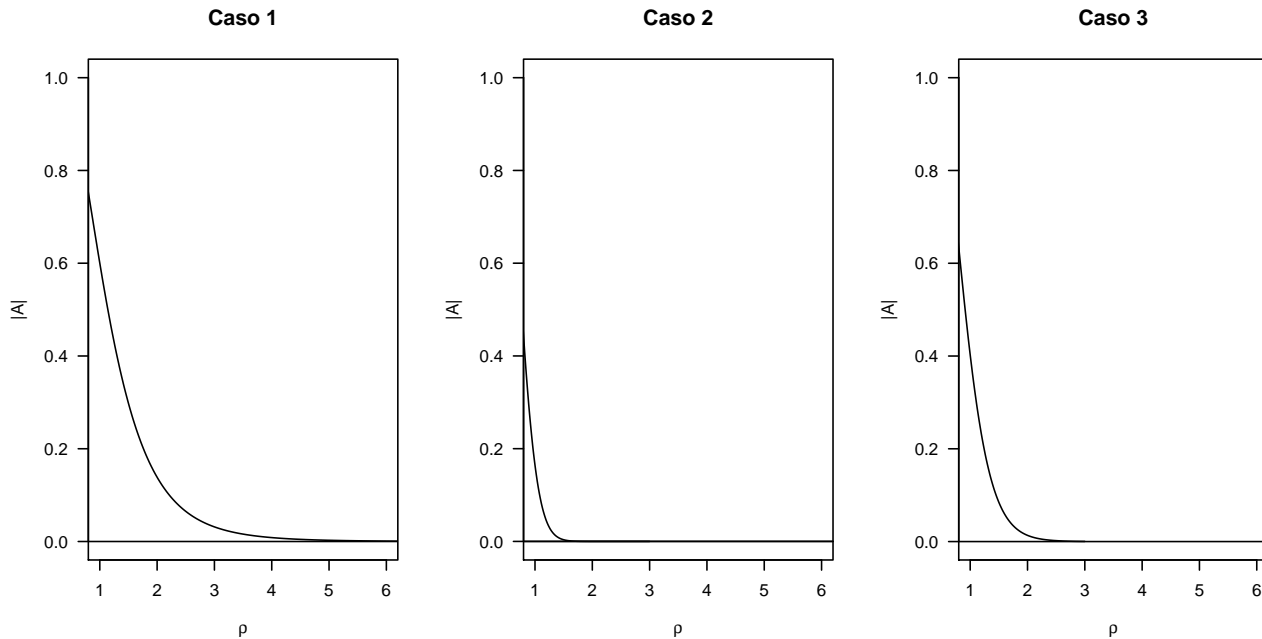
- $\frac{\partial \eta(t_i, \theta)}{\partial \theta_1} = -\theta_3 t_i \exp(-\theta_1 t_i)$
- $\frac{\partial \eta(t_i, \theta)}{\partial \theta_2} = \theta_3 t_i \exp(-\theta_2 t_i)$
- $\frac{\partial \eta(t_i, \theta)}{\partial \theta_3} = \exp(-\theta_1 t_i) - \exp(-\theta_2 t_i)$

Es de resaltar que  $M(\xi; \theta)$  también depende de  $\Sigma^{-1}(\xi)$  y  $\Sigma$  toma las diferentes formas mencionadas en la sección 3.2.1.

Ahora, para calcular la eficiencia del diseño con observaciones correlacionadas con respecto al de observaciones incorrelacionadas, se deben hallar los posibles valores de  $\rho$ ; debido a

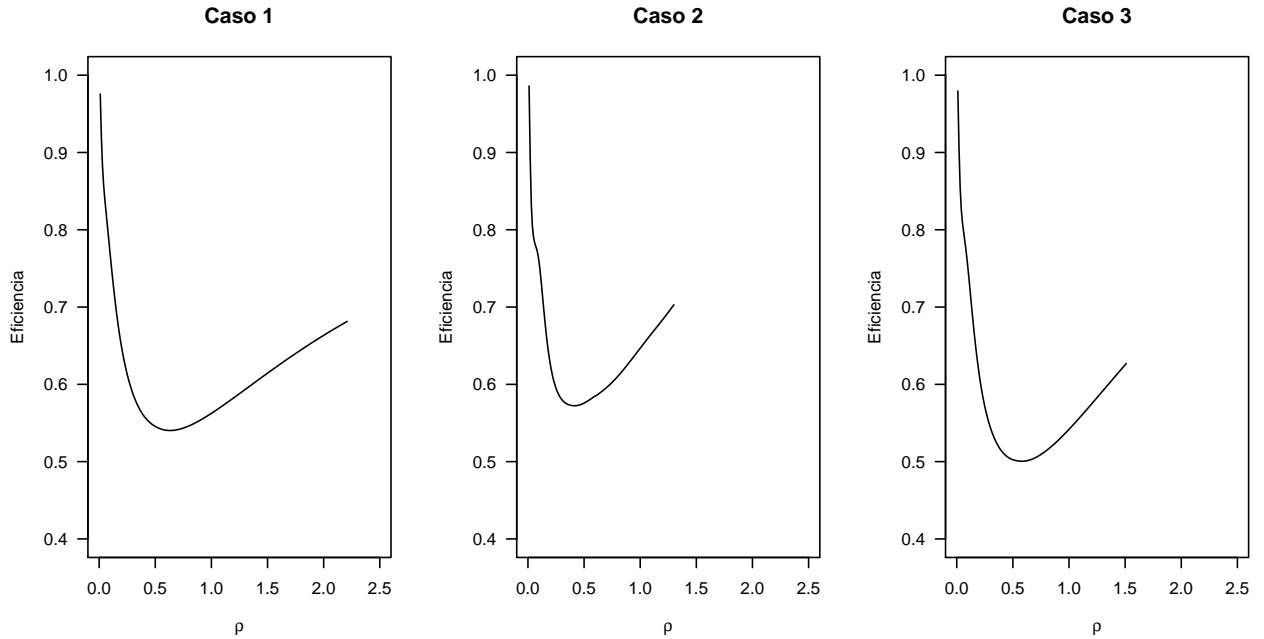
que la matriz  $A$  depende de éstos y son fundamentales para calcular el determinante de la matriz de información de Fisher.

La figura **3-3** muestra los valores del determinante de la matriz  $A$  en cada uno de los tres casos de la función de Matérn para una rejilla de valores de  $\rho$ . Se nota que cuando  $\rho$  es pequeño el determinante de  $A$  es cercano a uno y cuando aumenta el determinante se acerca a cero.



**Figura 3-3.:** Determinante de la matriz  $A$  en función de  $\rho$

Mediante un procedimiento numérico y la figura **3-3**, se obtiene que para el caso 1 de la función de covarianza de Matérn el intervalo de posibles valores para  $\rho$  donde  $|A| > 0$  es  $[0.01, 5]$ , para el caso 2 es  $[0.01, 1.5]$  y para el caso 3 es  $[0.01, 2.3]$ . Con estos posibles valores para  $\rho$  se calcula la eficiencia del diseño  $\xi_{OI}$  con respecto al diseño  $\xi_{OC}$ , mostradas en la figura **3-4**. En esta figura se observa que en el caso 1 de la función de covarianza de Matérn cuando  $\rho = 0.01$  la eficiencia del diseño  $\xi_{OI}$  con respecto al diseño  $\xi_{OC}$  alcanza su máximo en 0.9769 disminuyendo hasta 0.5470, luego aumenta hasta 0.6854, para el caso 2 cuando  $\rho = 0.01$  la eficiencia del diseño  $\xi_{OI}$  con respecto al diseño  $\xi_{OC}$  alcanza su máximo en 0.9830, disminuye hasta 0.5792 y luego aumenta hasta 0.7022 y para el caso 3 se observa que cuando  $\rho = 0.01$  la eficiencia del diseño  $\xi_{OI}$  con respecto al diseño  $\xi_{OC}$  alcanza su máximo en 0.9778, disminuye hasta 0.5062, luego aumenta hasta 0.6287. En general, cuando la correlación entre las observaciones no se tiene en cuenta el diseño utilizado pierde eficiencia y con el caso 2 de la función de covarianza de Matérn es donde se pierde menos eficiencia.



**Figura 3-4.:** Valores de  $\rho$  vs eficiencia del diseño  $\xi_{OI}$  con respecto al diseño  $\xi_{OC}$

La figura 3-5 muestra los puntos óptimos del diseño con estructura de correlación de los errores para algunos valores de  $\rho$ . Se observa que cuando  $\rho = 0.01$  el diseño contiene 3 puntos y cuando  $\rho$  comienza a aumentar el diseño pasa a tener 9 puntos, este comportamiento se presenta en los tres casos considerados de la función de covarianza de Matérn.

Los parámetros estimados que maximizaron la función de verosimilitud se muestran en la tabla 3-2, estos son utilizados para calcular los criterios de información que se muestran en la tabla 3-4 y los diseños cuasi-óptimos  $\xi_{OC}$  en cada caso de la función de covarianza de Matérn que se muestran en la tabla 3-3. Es de resaltar que los pesos  $w_i$  para cada punto de estos diseños son  $w_i = 1/9$ .

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\sigma}^2$	$\hat{\rho}$
Caso 1	0.181085	0.432933	4.312362	0.004015	1.112888
Caso 2	0.183871	0.419717	4.546761	0.005467	1.387299
Caso 3	0.190721	0.404182	5.016977	0.005496	2.007067

**Tabla 3-2.:** Parámetros estimados para cada caso de la función de covarianza de Matérn

En la tabla 3-3, se observa que el diseño  $\xi_{OI}$  pierde eficiencia al ser comparado con cada diseño  $\xi_{OC}$ , para los valores estimados de los parámetros dados en la tabla 3-2. Donde más eficiencia pierde es cuando se compara con el caso 2 y donde menos pierde es cuando se compara con el caso 3.

Para la selección del mejor modelo, teniendo en cuenta la estructura de correlación de las

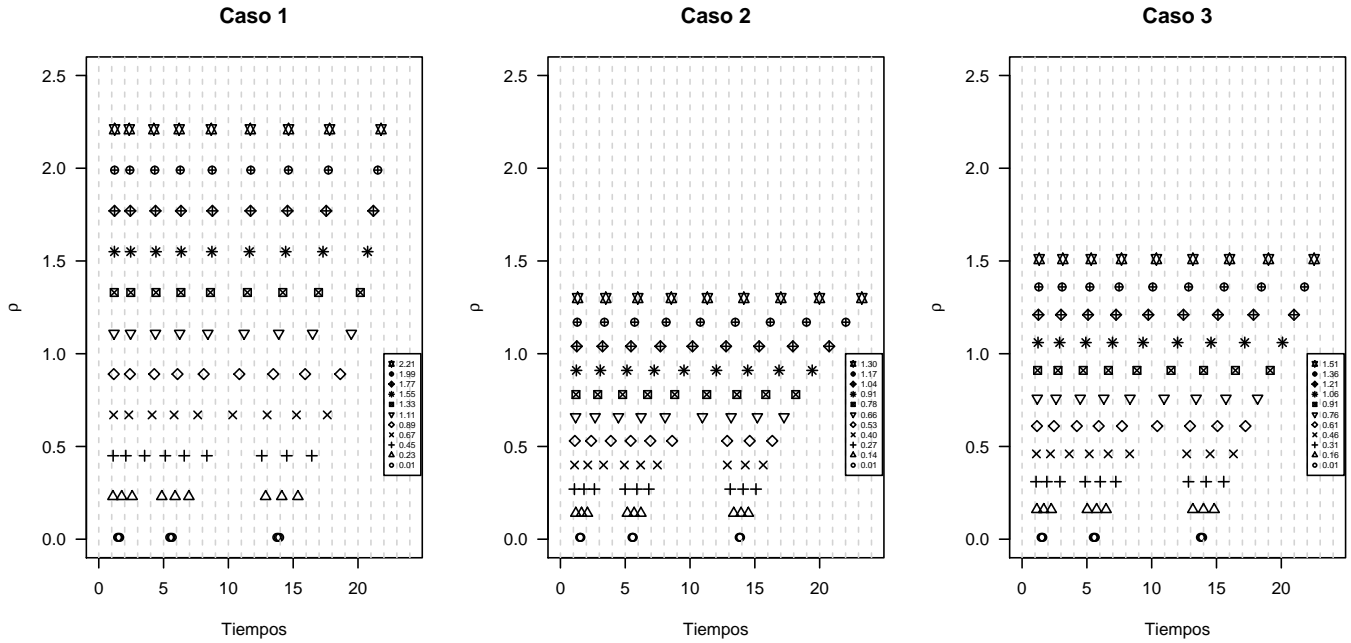


Figura 3-5.: Diseños óptimos vs  $\rho$

Tiempos	1	2	3	4	5	6	7	8	9	$D_{ef}$
$\xi_{OI}$	1.44	1.53	1.59	5.56	5.60	5.73	13.62	13.76	14.20	
Caso 1	1.18	1.97	3.25	4.72	6.14	7.74	12.36	14.30	16.25	0.7465
Caso 2	1.11	1.98	2.16	2.17	2.46	6.17	10.30	14.45	18.57	0.7380
Caso 3	1.54	4.43	7.93	12.32	16.34	16.35	16.36	16.37	19.99	0.7855

Tabla 3-3.: Diseños cuasi-óptimos  $\xi_{OC}$  y eficiencias para cada caso de la función de covarianza de Matérn

observaciones, se utiliza el criterio de información de Akaike y el criterio de información bayesiano, los resultados obtenidos se muestran en la tabla **3-4**.

	Caso 1 ( $\rho = 1.11$ )	Caso 2 ( $\rho = 1.38$ )	Caso 3 ( $\rho = 2.01$ )
AIC	-14.7747	-15.9536	-15.6320
BIC	-22.5775	-23.7564	-23.4348

**Tabla 3-4.:** Criterios de información de los diseños cuasi-óptimos para los tres casos de la función de covarianza de Matérn (tetraciclina)

En la tabla **3-4**, se observa que el modelo con menor criterio de información es el que corresponde al caso 2 de la función de covarianza de Matérn, por lo cual será utilizado posteriormente para hacer el estudio de simulaciones, y a su vez observar que tan eficiente es el diseño  $\xi_{OC}$  para estimar los parámetros del modelo. Ahora, este diseño se compara con el inicial y dos diseños exactos  $x$  espaciados con incrementos de 1.5 y 2 por medio de la eficiencia, los resultados se muestran en la tabla **3-5**.

Tiempos	1	2	3	4	5	6	7	8	9	$D_{ef}$
$\xi_{OC}$	1.11	1.98	2.16	2.17	2.46	6.17	10.30	14.45	18.57	
$\xi_{OC1}$	1.00	2.00	3.00	4.00	6.00	8.00	10.00	12.00	16.00	0.7964
$\xi_{OC2}$	1.00	2.50	4.00	5.50	7.00	8.50	10.00	11.50	13.00	0.7401
$\xi_{OC3}$	1.00	3.00	5.00	7.00	9.00	11.00	13.00	16.00	19.00	0.6818

**Tabla 3-5.:** Comparación de diseños exactos con el diseño cuasi-óptimo eficiente del caso 2 de la función de covarianza de Matérn

En la tabla **3-5**, se observa que los diseños  $\xi_{OC1}$ ,  $\xi_{OC2}$  y  $\xi_{OC3}$  pierden eficiencia, con respecto al diseño cuasi-óptimo  $\xi_{OC}$ , el diseño que más pierde es  $\xi_{OC3}$  y el que menos pierde es  $\xi_{OC1}$ .

### 3.5. Aplicación 2: Concentración de la sulfisoxazole

Los datos sobre el metabolismo de sulfisoxazole fueron obtenidos por Kaplan et al. (1972) y se reproducen en la tabla **3-6**. En este experimento, el sulfisoxazole se administra a un sujeto por vía intravenosa, se tomaron muestras de sangre pasado cierto tiempo (en minutos) después de ser administrado el medicamento y se midió la concentración de sulfisoxazole en el plasma en microgramos por mililitro ( $mcg/ml$ ).

Para los datos intravenosos, el modelo se muestra en 2-1 y la función  $\eta$  proviene de un modelo de 2 compartimientos, que se escribe como una suma de dos exponenciales, dado por:

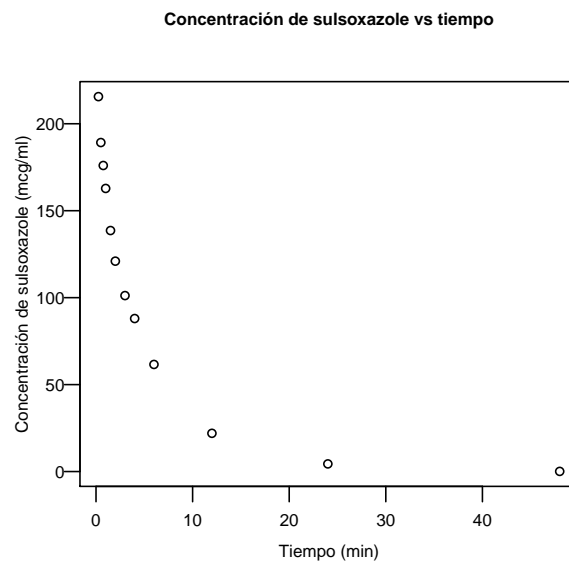
$$\eta(t_i, \theta) = \theta_1 \exp(-\theta_2 t_i) + \theta_3 \exp(-\theta_4 t_i), \quad (3-16)$$

Tiempo (minutos)	0.25	0.50	0.75	1.00	1.50	2.00
Concentración (mcg/ml)	215.6	189.2	176.0	162.8	138.6	121.0
Tiempo (minutos)	3.00	4.00	6.00	12.00	24.00	48.00
Concentración (mcg/ml)	101.2	88.0	61.6	22.0	4.4	0.1

**Tabla 3-6.:** Concentraciones de sulfisoxazole en (mcg/ml) vs el tiempo (en minutos)

donde  $\theta_i \in \mathbb{R}^+$ , para  $i = 1, 2, 3, 4$ ; puesto que son funciones de las tasas de eliminación y absorción del medicamento,  $\eta$  es una función no lineal que describe la relación entre la concentración de sulfisoxazole y el tiempo  $t$  en minutos.

En la figura 3-6 se grafican las concentraciones de sulfisoxazole contra el tiempo, en ésta se observa que a medida que pasan los minutos la concentración decae rápidamente, es decir, se hace mínima.

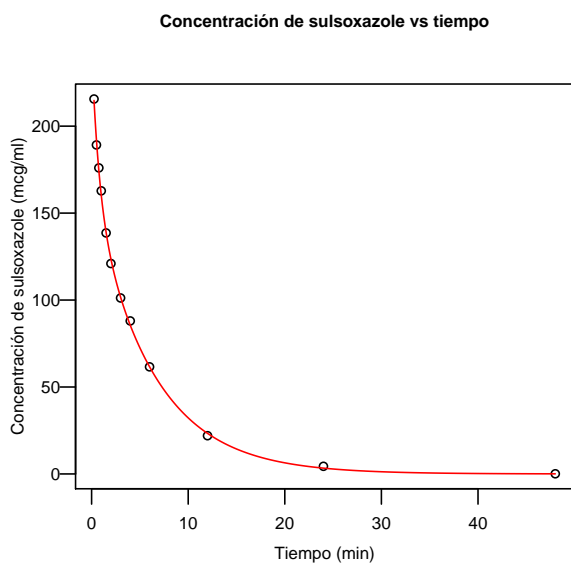


**Figura 3-6.:** Gráfica de la concentración de sulfisoxazole vs tiempo

Para hacer el ajuste del modelo con observaciones correlacionadas, primero se estiman los valores iniciales de los parámetros utilizando los métodos para observaciones independientes mencionados en el capítulo 2, esta estimación de los valores iniciales se muestra a continuación.

### 3.5.1. Estimación de valores iniciales (sulfisoxazole)

Para estimar los parámetros del modelo, se utilizan valores iniciales al rededor del parámetro; éstos se pueden hallar haciendo un análisis exploratorio, de donde se obtiene que para  $\theta_1 = 82$ ,  $\theta_2 = 1.3$ ,  $\theta_3 = 160$  y  $\theta_4 = 0.15$ . Con estos valores se estiman los parámetros del modelo para observaciones incorrelacionadas utilizando la función *nls* del *R*, la cual arrojó el siguiente vector de parámetros estimados  $\hat{\theta} = (81.24, 1.31, 162.59, 0.16)$ . Estos se grafican en la figura 3-7; donde se observa que la curva con estos valores pasa por todos los datos, indicando que la función  $\eta$  es adecuada para modelar las concentraciones de sulfisoxazole con respecto al tiempo.



**Figura 3-7.:** Gráfica de parámetros estimados de la concentración de sulfisoxazole vs tiempo

Los valores estimados para observaciones incorrelacionadas se utilizan como valores iniciales para la estimación de los parámetros del modelo con observaciones correlacionadas y encontrar los diseños cuasi-óptimos respectivos.

### 3.5.2. Búsqueda y comparación de diseños cuasi-óptimos para observaciones correlacionadas (sulfisoxazole)

Para hallar los diseños, primero se calcula el determinante de la matriz de información de Fisher, pero para esto es necesario calcular  $F(\xi; \theta) = \left[ \frac{\partial \eta(t_i, \theta)}{\partial \theta_j} \right]_{N \times p}$ ; las derivadas de la función  $\eta(t_i, \theta)$  con respecto a cada parámetro son:

- $\frac{\partial \eta(t_i, \theta)}{\partial \theta_1} = \exp(-\theta_2 t_i)$
- $\frac{\partial \eta(t_i, \theta)}{\partial \theta_2} = -\theta_1 t_i \exp(-\theta_2 t_i)$
- $\frac{\partial \eta(t_i, \theta)}{\partial \theta_3} = \exp(-\theta_4 t_i)$
- $\frac{\partial \eta(t_i, \theta)}{\partial \theta_4} = -\theta_3 t_i \exp(-\theta_4 t_i)$

Ahora, los parámetros estimados que maximizaron la función de verosimilitud se muestran en la tabla **3-7**, éstos son utilizados para calcular los criterios de información que se muestran en la tabla **3-9** y los diseños cuasi-óptimos  $\xi_{OC}$  en cada caso de la función de covarianza de Matérn que se muestran en la tabla **3-8**. Se observa que el diseño  $\xi_{OI}$  pierde eficiencia al ser comparado con cada diseño  $\xi_{OC}$ , para los valores estimados de los parámetros por cada diseño dados en la tabla **3-7**.

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\sigma}^2$	$\hat{\rho}$
Caso 1	216.099177	0.236927	163.897079	0.010000	69.591309	0.714554
Caso 2	211.762023	0.236364	162.594412	0.011000	62.095269	0.040872
Caso 3	212.938277	0.235426	162.594411	0.012000	67.672404	0.333780

**Tabla 3-7.:** Parámetros estimados para cada caso de la función de covarianza de Matérn

Tiempos	1	2	3	4	5	6	7	8	9	10	11	12	$D_{ef}$
$\xi_{OI}$	0.008	0.014	0.021	0.603	0.701	0.775	2.624	2.726	3.177	8.822	9.409	10.935	
Caso 1	0.002	0.217	0.518	1.023	1.879	2.721	3.662	4.870	7.792	9.315	10.783	12.371	0.8428
Caso 2	0.003	0.937	1.326	2.799	3.222	9.553	10.978	10.982	10.990	10.993	11.007	12.668	0.8608
Caso 3	0.008	0.014	0.021	0.603	1.265	2.207	3.076	4.041	7.878	9.188	10.429	11.745	0.8965

**Tabla 3-8.:** Diseños cuasi-óptimos  $\xi_{OC}$  para cada caso de la función de covarianza de Matérn

Para la selección del mejor modelo teniendo en cuenta la estructura de correlación de las observaciones, se utiliza el criterio de información de Akaike y el criterio de información bayesiano, los resultados obtenidos se muestran en la tabla **3-9**.

	Caso 1 ( $\rho = 0.7145$ )	Caso 2 ( $\rho = 0.0409$ )	Caso 3 ( $\rho = 0.3338$ )
AIC	94.2104	95.5985	94.8748
BIC	84.6953	86.0834	85.3597

**Tabla 3-9.:** Criterios de información de los diseños cuasi-óptimos para los tres casos de la función de covarianza de Matérn (sulfisoxazole)

En la tabla **3-9**, se observa que el modelo con menor criterio de información es el que corresponde al caso 1 de la función de covarianza de Matérn. Este es el adecuado para



modelar la concentración de sulfisoxazole a medida que transcurre el tiempo y los tiempos que se muestran en la tabla **3-8** para este diseño, son los adecuados para tomar las observaciones.

Ahora, este diseño se compara con el inicial y dos diseños exactos  $x$  espaciados con incrementos de 1 y 1.5 por medio de la eficiencia, los resultados se muestran en la tabla **3-10**.

Tiempos	1	2	3	4	5	6	7	8	9	10	11	12	$D_{ef}$
$\xi_{OC}$	0.002	0.217	0.518	1.023	1.879	2.721	3.662	4.870	7.792	9.315	10.783	12.371	
$\xi_{OC1}$	0.250	0.500	0.750	1.000	1.500	2.000	3.000	4.000	6.000	12.000	24.000	48.000	0.6895
$\xi_{OC2}$	0.250	1.750	3.250	4.750	6.250	7.750	9.250	10.750	12.750	13.750	15.250	16.750	0.5723
$\xi_{OC3}$	0.002	1.002	2.002	3.002	4.002	5.002	6.002	7.002	8.002	9.002	10.002	11.002	0.8947

**Tabla 3-10.:** Comparación de diseños exactos con el diseño cuasi-óptimo eficiente del caso uno de la función de covarianza de Matérn

En la tabla **3-10**, se observa que los diseños  $\xi_{OC1}$ ,  $\xi_{OC2}$  y  $\xi_{OC3}$  pierden eficiencia, con respecto al diseño cuasi-óptimo  $\xi_{OC}$ , el diseño que más pierde es  $\xi_{OC2}$  y el que menos pierde es  $\xi_{OC3}$ .



## 4. Estudio de simulación

En este capítulo, se expone y aplica la metodología utilizada para hacer el estudio de simulación, con el fin de analizar si el diseño cuasi-óptimo  $\xi_{OC} = \{1.11, 1.98, 2.16, 2.17, 2.46, 6.17, 10.30, 14.45, 18.57\}$  calculado en el ejemplo 1 para el caso 2 de la función de covarianza de Matérn del capítulo anterior estima de manera eficiente los parámetros del modelo y si además estas estimaciones son mejores que las del diseño  $\xi_{OI} = \{1.44, 1.53, 1.59, 5.56, 5.60, 5.73, 13.62, 13.76, 14.20\}$ .

El *script R* utilizado para calcular los resultados del estudio de simulación, se pueden ver en el anexo.

### 4.1. Simulación usando el caso 2 de la función de covarianza de Matérn

En esta sección, se observa que tan eficiente es el diseño  $\xi_{OC}$  para estimar los parámetros del modelo, teniendo en cuenta el caso 2 de la función de covarianza de Matérn del ejemplo 1 y se realiza de la siguiente forma:

1. Se construye la matriz  $A$ , con los tiempos del diseño  $\xi_{OC}$  y  $\hat{\rho} = 1.38$ , y se generan 1000 errores aleatorios de una normal multivariada con vector de media 0 y matriz de varianzas y covarianzas  $\Sigma = \hat{\sigma}^2 A$ .
2. Se generan 1000 concentraciones, evaluando en la función del modelo  $\eta$  los parámetros estimados  $\Theta_0^T = (\hat{\theta} \hat{\sigma}^2 \hat{\rho})^T = (0.183871 \ 0.419717 \ 4.546761 \ 0.005467 \ 1.387299)^T$  y los tiempos del diseño cuasi-óptimos  $\xi_{OC}$ , y se le adiciona la matriz de errores calculada utilizando la distribución normal multivariada.
3. Por último, con las concentraciones simuladas y los tiempos asociados al diseño óptimo se estiman los parámetros del modelo, cuyos resultados se muestran en la tabla 4-1, donde  $\Theta_0$ : valor real del vector  $\Theta$ ;  $\hat{\Theta}$ : media de los parámetros estimados;  $de$ : desviación estándar de  $\hat{\Theta}$ ;  $ECM$ : error cuadrático medio de cada parámetro;  $ECMR$ : error cuadrático medio relativo.

Los resultados  $de$ ,  $ECM$  y  $ECMR$  se calcularon con las fórmulas:

$$de = \sqrt{\frac{1}{1000} \sum_{i=1}^{1000} (\hat{\gamma}_i - \bar{\gamma})^2} \quad (4-1)$$

$$ECM = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{\gamma}_i - \gamma_j)^2 \quad (4-2)$$

$$ECMR = \frac{1}{1000\gamma_j^2} \sum_{i=1}^{1000} (\hat{\gamma}_i - \gamma_j)^2 \quad (4-3)$$

donde  $\gamma \in \{\theta_1, \theta_2, \theta_3, \sigma^2, \rho\}$ . En la tabla 4-1 se muestran los resultados obtenidos de la estimación de los parámetros.

	$\theta_1$	$\theta_2$	$\theta_3$	$\sigma^2$	$\rho$
$\Theta_0$	0.183871	0.419717	4.546761	0.005467	1.387299
$\bar{\Theta}$	0.180953	0.428796	4.362238	0.003514	1.370431
$de$	0.008357	0.025849	0.021716	0.001840	0.074595
$ECM$	0.000078	0.000749	0.034526	0.000007	0.005843
$ECMR$	0.002315	0.004257	0.001670	0.240686	0.003036

**Tabla 4-1.:** Resultado de los parámetros estimados del estudio de simulación

En la tabla 4-1, se observa que el promedio de los parámetros estimados  $\bar{\Theta}$  es muy parecido a los parámetros  $\Theta_0$ , además, los parámetros estimados tienen un error cuadrático medio relativo con respecto a  $\Theta_0$  del 0.23 %, 0.42 %, 0.16 %, 24.06 % y 0.30 % para  $\hat{\theta}_1$ ,  $\hat{\theta}_2$ ,  $\hat{\theta}_3$ ,  $\hat{\sigma}^2$  y  $\hat{\rho}$ , respectivamente. También se nota que la mayor variabilidad es la que está asociada al parámetro  $\hat{\sigma}^2$  y la de menor es la de  $\hat{\theta}_3$ . Por último, se puede decir, que los estimadores tienen sesgo pequeño; pues se alejan en cantidades pequeñas del verdadero valor del parámetro.

## 4.2. Metodología propuesta para las comparaciones entre los diseños considerados

Para hacer las comparaciones se utiliza el diseño cuasi-óptimo para observaciones incorrelacionadas  $\xi_{OI}$  y el diseño cuasi-óptimo del caso 2 de la función de covarianza de Matérn, para observaciones correlacionadas  $\xi_{OC}$ , seleccionado utilizando el AIC y BIC. Esta comparación se hace para verificar vía simulación si el diseño  $\xi_{OC}$  hallado con la segunda estructura de correlación, además de tener mejor eficiencia estima mejor los parámetros del modelo, que el diseño  $\xi_{OI}$ .

Las comparaciones se realizaron de la siguiente manera:

1. Se construyen las matrices  $A_1$  y  $A_2$  con los diseños  $\xi_{OI}$  y  $\xi_{OC}$ , respectivamente, utilizando el caso 2 de la función de covarianza de Matérn. Cada matriz se descompone en la forma  $A = A^{1/2}A^{1/2}$ , para ello se utiliza la función *sqrtn* de la librería *expm* del R.
2. Se generan 1000 vectores de errores  $\varepsilon \sim N_N(0, \sigma^2 I)$  donde  $I$  es la matriz identidad de dimensión  $N \times N$ . Luego para garantizar que los errores tienen distribución  $N_N(0, \sigma^2 A_i)$  se calcula:  $A_i^{1/2} \varepsilon^T$ ,  $i = 1, 2$ . Este proceso se hace con el fin de garantizar que los errores se generaron de la misma distribución y para que tengan sentido las comparaciones que se hacen posteriormente.
3. Se reemplazan los parámetros estimados y los tiempos de cada diseño en la función del modelo y se le adiciona el vector de errores, este proceso se hace para los 1000 vectores de errores, obteniendo de esta manera 1000 vectores de respuesta.
4. Con cada diseño, se estiman los parámetros tomando como valores iniciales de  $\Theta_0^T = (\hat{\theta} \hat{\sigma}^2 \hat{\rho})^T = (\hat{\theta}_1 \hat{\theta}_2 \hat{\theta}_3 \hat{\sigma}^2 \hat{\rho})^T = (0.183871 \ 0.419717 \ 4.546761 \ 0.005467 \ 1.387299)^T$ , y los diseños  $\xi_{OI}$  y  $\xi_{OC}$ .
5. Por último, a cada vector de parámetros se le calcula la media, desviación estándar, error cuadrático medio y coeficiente de variación, para compararlos con el real y observar que tanta diferencia hay entre éstos.

En la tabla **4-2**, se muestra los parámetros estimados del estudio de simulación para el diseño con observaciones incorrelacionadas, se observa que los parámetros estimados tienen una variabilidad con respecto a  $\Theta_0$  del 0.49 %, 0.70 %, 1.13 %, 48.16 % y 16.18 % para  $\hat{\theta}_1$ ,  $\hat{\theta}_2$ ,  $\hat{\theta}_3$ ,  $\hat{\sigma}^2$  y  $\hat{\rho}$ , respectivamente. También se nota que la mayor variabilidad es la que está asociada al estimador de  $\sigma^2$  y la de menor es la de  $\hat{\theta}_1$ .

	$\theta_1$	$\theta_2$	$\theta_3$	$\sigma^2$	$\rho$
$\Theta_0$	0.183871	0.419717	4.546761	0.005467	1.387299
$\tilde{\Theta}$	0.183001	0.424397	4.535568	0.002871	1.160502
<i>de</i>	0.012904	0.034960	0.483936	0.002778	0.510251
<i>ECM</i>	0.000167	0.001242	0.234086	0.000014	0.311532
<i>ECMR</i>	0.004943	0.007055	0.011323	0.483667	0.161868

**Tabla 4-2.:** Parámetros estimados del estudio de simulación para el diseño con observaciones incorrelacionadas

En la tabla **4-3**, se muestra los parámetros estimados del estudio de simulación para el diseño con observaciones correlacionadas, se observa que los parámetros estimados tienen un error cuadrático medio relativo con respecto a  $\Theta_0$  del 0.19 %, 0.42 %, 0.16 %, 21.61 % y 0.24 % para  $\hat{\theta}_1$ ,  $\hat{\theta}_2$ ,  $\hat{\theta}_3$ ,  $\hat{\sigma}^2$  y  $\hat{\rho}$ , respectivamente. También se nota que la mayor variabilidad es la que está asociada al estimador de  $\sigma^2$  y la de menor es la de  $\hat{\theta}_3$ .

	$\theta_1$	$\theta_2$	$\theta_3$	$\sigma^2$	$\rho$
$\Theta_0$	0.183871	0.419717	4.546761	0.005467	1.387299
$\tilde{\Theta}$	0.181222	0.430021	4.363110	0.003658	1.367141
<i>de</i>	0.007781	0.025291	0.012442	0.001786	0.065664
<i>ECM</i>	0.000067	0.000745	0.033888	0.000006	0.004713
<i>ECMR</i>	0.001996	0.004230	0.001639	0.216190	0.002449

**Tabla 4-3.:** Parámetros estimados del estudio de simulación para el diseño con observaciones correlacionadas

Al comparar las estimaciones realizadas con el diseño  $\xi_{OI}$  y  $\xi_{OC}$ , se puede decir, que el diseño  $\xi_{OC}$  estima mejor los parámetros, pues presentan menos variabilidad y menos sesgo que el diseño  $\xi_{OI}$ . En conclusión, el diseño  $\xi_{OC}$  es mejor que el diseño  $\xi_{OI}$  para estimar los parámetros del modelo.

# 5. Conclusiones y recomendaciones

En este capítulo, se presentan de forma general las conclusiones de este trabajo, a fin de encontrar un diseño cuasi-óptimo eficiente para observaciones correlacionadas, utilizando la función de Matérn para modelar la estructura de covarianza del error.

## 5.1. Conclusiones

- La función de covarianza de Matérn, es una buena opción para modelar la estructura de covarianza del error, cuando se desconoce la forma de ésta, porque los diseños hallados mostraron buenas estimaciones.
- La estimación de un buen diseño cuasi-óptimo, depende de la estimación de los parámetros utilizados y de la estructura de la matriz de varianzas y covarianzas del error.
- Cuando no se tiene en cuenta la correlación entre las observaciones, el diseño con observaciones incorrelacionadas pierde eficiencia con respecto al de observaciones correlacionadas, en todos los casos propuestos en este trabajo.
- El diseño aplicado en el estudio de simulación, teniendo en cuenta el caso 2 de la función de covarianza de Matérn, arrojó buenas estimaciones de los parámetros, ya que presentaron variaciones muy pequeñas con respecto  $\Theta_0$  entre el 0.0007% y el 3.45%.
- El diseño para observaciones correlacionadas  $\xi_{OC}$  es más eficiente y estima mejor los parámetros del modelo con respecto al diseño, para observaciones incorrelacionadas  $\xi_{OI}$ .

## 5.2. Recomendaciones

A partir de este trabajo se pueden desarrollar algunos proyectos de investigación a futuro, entre los cuales se tiene:

- Buscar diseños cuasi-óptimos utilizando otros criterios de optimalidad, tales como el G-óptimo, el C-óptimo. Los diseños nuevos pueden ser comparados con los que se encontraron en este trabajo y ver que tan eficientes son.

- Utilizar métodos bayesianos, para calcular diseños cuasi-óptimos cuando los errores tienen una estructura de covarianza modelada por la función de Matérn.
- Proponer nuevas metodologías para modelar la estructura de covarianza del error.
- Se recomienda usar funciones del *R* tales como: *DEoptim*, *nls* y *nlminb* para hacer la estimación de los parámetros, cuando no se puedan despejar fácilmente de la función de verosimilitud.
- Buscar diseños cuasi-óptimos  $x$  espaciados utilizando la función de covarianza de Matérn, para modelar la estructura de covarianza del error.



# A. Anexo

```
#####  
## Script R  
# Librerias  
library(MASS)  
library(DEoptim)  
library(expm)  
library(proto)  
library(nls2)  
library(nlstools)  
library(nlme)  
# Observaciones incorrelacionadas  
# Funciones para encontrar los puntos óptimos del diseño  
# Estimación de theta para observaciones independientes  
esttheta<-function(theta,tiem,y){  
  t1<-theta[1]; t2<-theta[2];t3<-theta[3]  
  eta<-t3*(exp(-t1*tiem)-exp(-t2*tiem))  
  t(y-eta)%*(y-eta)  
}  
  
# función a minimizar para tiempos independientes  
function0<-function(tiem){  
  N<-length(tiem)  
  dt1<--tiem*thetaind[3]*exp(-thetaind[1]*tiem)  
  dt2<-tiem*thetaind[3]*exp(-thetaind[2]*tiem)  
  dt3<-exp(-thetaind[1]*tiem)-exp(-thetaind[2]*tiem)  
  F<-matrix(c(dt1,dt2,dt3),nrow=length(tiem))  
  Ft<-t(F)  
  M<-Ft%*F  
  logar<--log(det(M))  
  return(logar)  
}  
  
#####  
# Observaciones correlacionadas  
# Función a minimizar para calcular menos el logaritmo del
```

```

# determinante de la matriz de información de Fisher
# Caso 1
function21<-function(tiem,theta,rho){
N<-length(tiem)
# Matriz de varianzas y covarianzas para las observaciones correlacionadas
# utilizando uno de los casos de la función de Matern
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-exp(-abs(tiem[j]-tiem[i])/rho)}
# Derivada de A
dA<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
dA[j,i] <-exp(-abs(tiem[j]-tiem[i])/rho)*(abs(tiem[j]-tiem[i])/(rho^2))}
# valores propios
B<-solve(A)%*%dA
lambda<-eigen(B)$values
varlam<-((N-1)/N)*var(lambda)
#dtM22<-N*sum(diag(B%*%B))-sum(diag(B))^2
## Derivada de eta con respecto a cada parámetro
dt1<--tiem*theta[3]*exp(-theta[1]*tiem)
dt2<-tiem*theta[3]*exp(-theta[2]*tiem)
dt3<-exp(-theta[1]*tiem)-exp(-theta[2]*tiem)
F<-matrix(c(dt1,dt2,dt3),nrow=length(tiem))
Ft<-t(F)
## matriz de información de Fisher
M<-Ft%*%solve(A)%*%F
logar<--(log(det(M))+log((N^2)*varlam))
#logar<--(log(det(M))+log(dtM22))
return(logar)
}

# Caso 2
function22<-function(tiem,theta,rho){
N<-length(tiem)
# Matriz de varianzas y covarianzas para las observaciones correlacionadas
# utilizando uno de los casos de la función de Matern
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-exp(-(abs(tiem[j]-tiem[i])^2)/(2*rho^2))}

```

---

```

# Derivada de A
dA<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
dA[j,i] <-(exp(-(abs(tiem[j]-tiem[i])^2)/(2*rho^2)))*((abs(tiem[j]-tiem[i])^2)/(rho^3))}
# valores propios
B<-solve(A)%*%dA
lambda<-eigen(B)$values
varlam<-((N-1)/N)*var(lambda)
#dtM22<-N*sum(diag(B%*%B))-sum(diag(B))^2
## Derivada de eta con respecto a cada parámetro
dt1<--tiem*theta[3]*exp(-theta[1]*tiem)
dt2<-tiem*theta[3]*exp(-theta[2]*tiem)
dt3<-exp(-theta[1]*tiem)-exp(-theta[2]*tiem)
F<-matrix(c(dt1,dt2,dt3),nrow=length(tiem))
Ft<-t(F)
## matriz de información de Fisher
M<-Ft%*%solve(A)%*%F
logar<--(log(det(M))+log((N^2)*varlam))
#logar<--(log(det(M))+log(dtM22))
return(logar)
}

# Caso 3
function23<-function(tiem,theta,rho){
N<-length(tiem)
# Matriz de varianzas y covarianzas para las observaciones correlacionadas
# utilizando uno de los casos de la función de Matern
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-(1+(sqrt(3)*abs(tiem[j]-tiem[i])/rho))*exp(-sqrt(3)*abs(tiem[j]-tiem[i])/rho)}
# Derivada de A
dA<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
dA[j,i] <-exp(-sqrt(3)*abs(tiem[j]-tiem[i])/rho)*(3*((tiem[j]-tiem[i])^2)/(rho^3))}
# valores propios
B<-solve(A)%*%dA
lambda<-eigen(B)$values
varlam<-((N-1)/N)*var(lambda)
dt1<--tiem*theta[3]*exp(-theta[1]*tiem)

```

```

dt2<-tiem*theta[3]*exp(-theta[2]*tiem)
dt3<-exp(-theta[1]*tiem)-exp(-theta[2]*tiem)
F<-matrix(c(dt1,dt2,dt3),nrow=length(tiem))
Ft<-t(F)
## matriz de información de Fisher
M<-Ft%%solve(A)%%F
logar<--(log(det(M))+log((N^2)*varlam))
#logar<--(log(det(M))+log(dtM22))
return(logar)
}

# Valores de rho
# Caso 1
valoresrho1<-function(rho){
N<-length(tiem)
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-exp(-abs(tiem[j]-tiem[i])/rho)}
deterA<-det(A)
deterA
}

# Caso 2
valoresrho2<-function(rho){
N<-length(tiem)
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-exp(-(abs(tiem[j]-tiem[i])^2)/(2*rho^2))}
deterA<-det(A)
deterA
}

# Caso 3
valoresrho3<-function(rho){
N<-length(tiem)
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-(1+(sqrt(3)*abs(tiem[j]-tiem[i])/rho))*
exp(-sqrt(3)*abs(tiem[j]-tiem[i])/rho)}

```

```

deterA<-det(A)
deterA
}

# Función utilizada para calcular la verosimilitud
# Caso 1
estmpatros<-function(estima1,tiem,conc){
N<-length(tiem)
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-exp(-abs(tiem[j]-tiem[i])/estima1[5])}
eta<-estima1[3]*(exp(-estima1[1]*tiem)-exp(-estima1[2]*tiem))
LogL<-(-1/2)*log(det(estima1[4]*A))-(N/2)*log(2*pi)-(1/2)*
as.numeric(t(conc-eta)%*%solve(estima1[4]*A)%*%(conc-eta))
LogL1<--LogL
}

# Caso 2
estmpatros2<-function(estima1,tiem,conc){
N<-length(tiem)
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-exp(-(abs(tiem[j]-tiem[i])^2)/(2*estima1[5]^2))}
eta<-estima1[3]*(exp(-estima1[1]*tiem)-exp(-estima1[2]*tiem))
LogL<-(-1/2)*log(det(estima1[4]*A))-(N/2)*log(2*pi)-(1/2)*
as.numeric(t(conc-eta)%*%solve(estima1[4]*A)%*%(conc-eta))
LogL1<--LogL
}

# Caso 3
estmpatros3<-function(estima1,tiem,conc){
N<-length(tiem)
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-((1+(sqrt(3)*abs(tiem[j]-tiem[i])/estima1[5]))*
exp(-sqrt(3)*abs(tiem[j]-tiem[i])/estima1[5]))}
eta<-estima1[3]*(exp(-estima1[1]*tiem)-exp(-estima1[2]*tiem))
LogL<-(-1/2)*log(det(estima1[4]*A))-(N/2)*log(2*pi)-(1/2)*
as.numeric(t(conc-eta)%*%solve(estima1[4]*A)%*%(conc-eta))
}

```

```

LogL1<--LogL
}
#####
# Selección del mejor modelo
# Caso 1
rhoest<-1
estima1<-c(thetaind,sigestind,rhoest)
res<-nlminb(estima1,estmpatros,tiem=tiem,conc=conc)
parametrosest<-res$par
logvero1<--(res$objective)
AICML1<-2*(length(estima1))-2*logvero1
BICML1<-length(length(estima1))*log(N)-2*logvero1
# Caso 2
res2<-nlminb(estima1, estmpatros2,tiem=tiem,conc=conc)
parametrosest2<-res2$par
logvero2<--(res2$objective)
AICML2<-2*(length(estima1))-2*logvero2
BICML2<-length(length(estima1))*log(N)-2*logvero2
## Caso 3
res3<-nlminb(estima1, estmpatros3,tiem=tiem,conc=conc)
parametrosest3<-res3$par
logvero3<--(res3$objective)
AICML3<-2*(length(estima1))-2*logvero3
BICML3<-length(length(estima1))*log(N)-2*logvero3
#####
# Estudio de simulación
# Estimación de los parámetros
estmpatros<-function(estima1,conc,tiemposopt){
N<-length(tiemposopt)
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-exp(-(abs(tiemposopt[j]-tiemposopt[i])^2)/(2*estima1[5]^2))}
eta<-estima1[3]*(exp(-estima1[1]*tiemposopt)-exp(-estima1[2]*tiemposopt))
LogL<-(-1/2)*log(det(estima1[4]*A))-(N/2)*log(2*pi)-(1/2)*
as.numeric(t(conc-eta)%*%ginv(estima1[4]*A)%*%(conc-eta))
LogL1<--LogL
return(LogL1)
}
# la función estima los parámetros theta3,sigma y rho
# conociendo theta1 y theta2
estmpatros3<-function(estima1,conc,tiemposopt,thett){

```

---

```

N<-length(tiemposopt)
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-exp(-(abs(tiemposopt[j]-tiemposopt[i])^2)/(2*estima1[3]^2))}
eta<-estima1[1]*(exp(-thett[1]*tiemposopt)-exp(-thett[2]*tiemposopt))
detm<-det(estima1[2]*A)
if(detm>0){
LogL<-(-1/2)*log(detm)-(N/2)*log(2*pi)-(1/2)*
as.numeric(t(conc-eta)%*%ginv(estima1[2]*A)%*(conc-eta))}
else{LogL<-1}
LogL1<--LogL
return(LogL1)
}
# la función estima los parámetros theta1 y theta2
# conociendo theta3,sigma y rho
estmpatros4<-function(thett1,conc,tiemposopt,estima1){
N<-length(tiemposopt)
A<-matrix(0,nrow=N,ncol=N)
for (i in 1:N)
{for (j in 1:N)
A[j,i] <-exp(-(abs(tiemposopt[j]-tiemposopt[i])^2)/(2*estima1[3]^2))}
eta<-estima1[1]*(exp(-thett1[1]*tiemposopt)-exp(-thett1[2]*tiemposopt))
detm<-det(estima1[2]*A)
if(detm>0){
LogL<-(-1/2)*log(detm)-(N/2)*log(2*pi)-(1/2)*
as.numeric(t(conc-eta)%*%ginv(estima1[2]*A)%*(conc-eta))}
else{LogL<-1}
LogL1<--LogL
return(LogL1)
}
# Cálculo del error cuadrático medio
ECM<-function(z){
sum(z^2)/length(z)
}
# Funciones para generar las respuestas
yest<-functionerrores,tiemposexp){
theta[3]*(exp(-theta[1]*tiemposexp)-exp(-theta[2]*tiemposexp))+errores
}
# Matriz de correlación A
matrizA<-function(tiemposexp){
A<-matrix(0,nrow=N,ncol=N)

```

```

for (i in 1:N)
{for (j in 1:N)
A[j,i] <-exp(-(abs(tiemposexp[j]-tiemposexp[i])^2)/(2*rho^2))}
A
}

# Generación de observaciones
rho=parametrosest2[5]
theta=parametrosest2[1:3]
N<-length(tiemposopt)
mu<-rep(0,length(tiemposopt))
n<-1500
h<-1
while(h==1){
errores<-(mvrnorm(n,mu,matrizA(tiemposopt)*parametrosest2[4]))
conc11<-apply(errores,1,yest,tiemposexp=tiemposopt)
indec1<-numeric(0)
for(j in 1:ncol(conc11)){
longitud1<-length(as.numeric(which(conc11[,j]<0)))
if(longitud1>0){indec1<-cbind(indec1,j)}
}
conc1<-conc11[,-indec1]
nnn<-ncol(conc1)
if(nnn>=1000){h<-0
conc1<-conc1[,1:1000]}else{h<-1}
}
#####
# Estimación de los 1000 parámetros
estima2<-c(thetaind,sigestind,rho)
estima3<-c(thetaind[3],sigestind,rho)
estima4<-thetaind[1:2]
lower<-c(0,0,4,0.0035,1.35)
upper<- c(1,1,5,0.0062,1.4)
lower1<-c(0,0,0)
upper1<-c(10,0.0099,10)
n<-1000
estimacionespara<-matrix(0,nrow=n,ncol=length(estima2))
funobjetivo<-numeric(n)
diferencia<-numeric(n)
for(i in 1:n){
res<-nlminb(estima2,estmpatros,conc=conc1[,i],tiemposopt=tiemposopt,
lower=lower, upper=upper)
thett<-res$par[1:2]
}

```



```

dife1<-3
while(dife1>2){
res3<-nlminb(estima3,estmpatros3,conc=conc1[,i],tiemposopt=tiemposopt,
thett=thett,lower=lower1, upper=upper1)
res4<-nlminb(estima4,estmpatros4,conc=conc1[,i],tiemposopt=tiemposopt,
estima1=res3$par)
dife1<-abs(res4$objective-res3$objective)
thett<-res4$par
      }
estimacionespara[i,]<-c(res4$par,res3$par)
funobjetivo[i]<-res4$objective
diferencia[i]<-dife1
      }
#####
# Comparación de diseños
# Generación de observaciones
Z<-sqrtm(matrizA(tiemind))
Z1<-sqrtm(matrizA(tiemposopt))
n<-1500
h<-1
while(h==1){
set.seed(10)
errores0<-(mvrnorm(n,mu,parametrosest2[4]*diag(1,length(tiemind))))
errores<-(mvrnorm(n,mu,parametrosest2[4]*diag(1,length(tiemposopt))))
errores1<-Z%*%t(errores0)
errores2<-Z1%*%t(errores)
conc21<-apply(errores1,2,yest,tiemposexp=tiemind)
conc31<-apply(errores2,2,yest,tiemposexp=tiemposopt)
#ind1<-numeric(0)
ind2<-numeric(0)
for(j in 1:ncol(conc31)){
long3<-length(as.numeric(which(conc31[,j]<0)))
if(long3>0){ind2<-cbind(ind2,j)}
      }

conc2<-conc21[,1:1000]
conc3<-conc31[,1:1000]
nnn1<-ncol(conc3)
if(nnn1>=1000){h<-0
conc3<-conc3[,1:1000]}else{h<-1}
      }
## Estimación de los 1000 parámetros para los dos modelos
lower1<-c(0,0,0)

```

```

upper1<- c(10,0.0099,10)
n<-1000
estimacionespara2<-matrix(0,nrow=n,ncol=length(estima2))
funobjetivo2<-numeric(n)
diferencia2<-numeric(n)
estimacionespara3<-matrix(0,nrow=n,ncol=length(estima2))
funobjetivo3<-numeric(n)
diferencia3<-numeric(n)
for(i in 1:n){
res<-nlminb(estima2,estmpatros,conc=conc2[,i],tiemposopt=tiemind,
lower=lower, upper=upper)
thett<-res$par[1:2]
res1<-nlminb(estima2,estmpatros,conc=conc3[,i],tiemposopt=tiemposopt,
lower=lower, upper=upper)
thett1<-res1$par[1:2]
dife1<-3
dife2<-3
while(dife1>2 | dife2>2){
res3<-nlminb(estima3,estmpatros3,conc=conc2[,i],tiemposopt=tiemind,
thett=thett,lower=lower1, upper=upper1)
res4<-nlminb(estima4,estmpatros4,conc=conc2[,i],tiemposopt=tiemind,
estimal=res3$par)
dife1<-abs(res4$objective-res3$objective)
thett<-res4$par
res5<-nlminb(estima3,estmpatros3,conc=conc3[,i],tiemposopt=tiemposopt,
thett=thett1,lower=lower1, upper=upper1)
res6<-nlminb(estima4,estmpatros4,conc=conc3[,i],tiemposopt=tiemposopt,
estimal=res5$par)
dife2<-abs(res6$objective-res5$objective)
thett1<-res6$par
}
estimacionespara2[i,]<-c(res4$par,res3$par)
funobjetivo2[i]<-res4$objective
diferencia2[i]<-dife1
estimacionespara3[i,]<-c(res6$par,res5$par)
funobjetivo3[i]<-res6$objective
diferencia3[i]<-dife2
}

```

# Bibliografía

- Akaike, H. (1974), 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control*, **19**(6), 716–723.
- Allen, D. (1983), 'Parameter estimation for nonlinear models with emphasis on compartmental models', **39**(3), 629–637.
- Amo, M., López, J. & López-Ríos, V. (2012), 'Optimal designs for two nested pharmacokinetic models with correlated observations', *Communications in Statistics*, **41**(1), 944–963.
- Atkinson, A., Donev, A. & Tobias, R. (2007), *Optimum Experimental Designs with SAS*, New York.
- Bartlett, M. S. (1937), 'Properties of sufficiency and statistical tests', *Proceedings of the Royal Statistical Society Series A* **160**, 268–282.
- Bates, D. & Watts, D. (1988), *Nonlinear Regression Analysis and Its Applications*, John Wiley & Sons, New York.
- Bogacka, B. & Wright, F. (2005), 'Non-linear design problem in a chemical kinetic model with non-constant error variance', *Journal of Statistical Planning and Inference* **128**, 633–648.
- Buse, A. (1973), 'Goodness of fit in generalized least squares estimation', *The American Statistician* **27**(3), 106–108.
- Carroll, R. & Ruppert, D. (1988), *Transformation and Weighting in Regression*, Chapman and Hall, New York.
- Clark, B. & Smith, D. (1989), *An Introduction to Pharmacokinetics*, New York.
- Dette, H., Kunert, J. & Pepelyshev, A. (2008), 'Exact optimal designs for weights least squares analysis with correlated errors', *Statistica Sinica* **18**, 135–154.
- Dette, H., Pepelyshev, A. & Zhigljavsky, A. (2013), 'Optimal design for linear models with correlated observations', *The Annals of Statistics* **41**(1), 143–176.

- Fedorov, V. & Hackl, P. (1997), *Model - Oriented Design of Experiments. Lecture Notes in Statistics*, New York.
- Gabrielsson, J. & Weiner, D. (2000), *Pharmacokinetic and Pharmacodynamic Data Analysis*, Kristianstad Boktryckeri AB, Stockholm.
- Huet, S., Bouvier, A., Poursat, M.-A. & Jolivet, E. (2004), *Statistical Tools for Nonlinear Regression*, Springer, New York.
- Kaplan, S. A., Weinfeld, R. E., Abruzzo, C. W. & Lewis, M. (1972), 'Pharmacokinetic profile of sulfisoxazole following intravenous, intramuscular, and oral administration to man', *Journal of Pharmaceutical Sciences* **61**, 773–778.
- Kiefer, J. (1959), 'Optimum Experimental Designs', *Journal of the Royal Statistical Society* **21**(1), 272–319.
- Kiefer, J. & Wolfowitz, J. (1960), 'The Equivalence of two Extremum Problems', *Canadian Journal of Mathematics* **12**(1), 363–366.
- López-Ríos, V. (2008), Diseños Óptimos para Discriminación y Estimación en Modelos No Lineales, PhD thesis, CIMAT, Guanajuato, México.
- López-Ríos, V. & Ramos-Quiroga, R. (2007), 'Introducción a los Diseños Óptimos', *Revista Colombiana de Estadística* **30**(1), 37–51.
- Matern, B. (1960), *Spatial Variation*, Springer, New York.
- Minasny, T. B. & McBratney, A. (2005), 'The Matérn function as a general model for soil variograms', *Geoderma* **128**, 192–207.
- Montgomery, D. (2006), *Introducción al análisis de regresión lineal*, John Wiley & Sons, New York.
- Muller, W. & Pazman, A. (2003), 'Measures for design in experiments with correlated error', *Biometrika* **90**, 423–434.
- Patan, M. & Bogacka, B. (2007), 'Optimum experimental designs for dynamic systems in the presence of correlated errors', *Computational Statistics* **51**, 5644–5661.
- Pazman, A. (1986), *Foundations of Optimum Experimental Design*, D. Reidel Publishing Company, Dordrecht.
- Pazman, A. (2007), 'Criteria of optimal designs for small-samples experiments with correlated observations', *Kybernetika* **43**(4), 453–462.

- 
- R Development Core Team (2013), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.  
**URL:** <http://www.R-project.org>
- Ratkowsky, D. A. (1983), *Nonlinear Regression Modeling*, D. Dekker, New York.
- Ritz, C. & Streibig, J. (2008), *Nonlinear Regression with R*, Springer, New York.
- Schwarz, G. (1978), 'Estimating the dimension of a model', *Annals of Statistics* **6**(2), 461–464.
- Seber, G. & Wild, C. (2003), *Nonlinear Regression*, John Wiley & Sons, New Jersey.
- Shapiro, S. S. & Wilk, M. B. (1965), 'An analysis of variance test for normality', *Biometrika* **52**(3-4), 591–611.