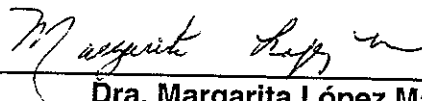


TESIS DEFENDIDA POR

Jonás de Dios De Basabe Delgado

Y APROBADA POR EL SIGUIENTE COMITÉ



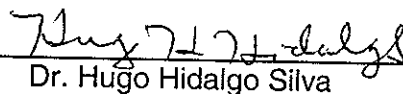
Dra. Margarita López Martínez

Director del Comité



Dr. Enrique Gómez Treviño

Miembro del Comité



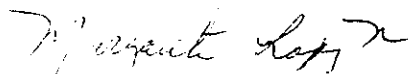
Dr. Hugo Hidalgo Silva

Miembro del Comité



Dr. Bodo Weber

Miembro del Comité



Dra. Margarita López Martínez

*Jefe del Departamento de
Geología*

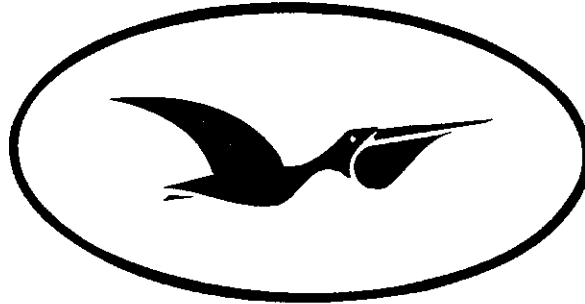


Dr. Luis Alberto Delgado Argote

Director de Estudios de Posgrado

21 de Febrero de 2002

CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE
EDUCACIÓN SUPERIOR DE ENSENADA



DIVISIÓN DE CIENCIAS DE LA TIERRA
DEPARTAMENTO DE GEOLOGÍA

**Regresión lineal con incertidumbres en
todas las variables: aplicaciones en
geocronología al cálculo de isocronas**

TESIS

para obtener el grado de Maestro en Ciencias que presenta:

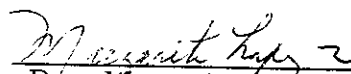
Jonás de Dios De Basabe Delgado

Febrero de 2002
Ensenada, Baja California, México

RESUMEN de la tesis de JONÁS DE DIOS DE BASABE DELGADO, presentada como requisito parcial para la obtención del grado de Maestro en Ciencias en Ciencias de la Tierra con orientación en Geología. Ensenada, Baja California, México. Febrero de 2002.

Regresión lineal con incertidumbres en todas las variables: aplicaciones en geocronología al cálculo de isocronas

Resumen aprobado por:


Dra. Margarita López Martínez
Directora de la tesis

En ciencias, existen numerosas aplicaciones en donde es necesario calcular una línea recta, minimizando las distancias de un conjunto de datos. Por citar un ejemplo, en geocronología se determinan experimentalmente un conjunto de variables (e.g. $^{39}\text{Ar}/^{36}\text{Ar}$ y $^{40}\text{Ar}/^{36}\text{Ar}$) a partir de las cuales es posible extraer información relevante: de la pendiente se determina la edad, y de la ordenada al origen, el $^{40}\text{Ar}/^{36}\text{Ar}$ atrapado al momento de su formación. Los datos recabados experimentalmente están sujetos a incertidumbres, de modo que es necesario tomarlas en consideración para incorporarlas en el resultado final.

Al trazar una línea recta a través de un conjunto de datos, por lo general se hacen las siguientes suposiciones: los datos pertenecen a una línea recta, sólo una de las variables está sujeta a errores, y éstos tienen una distribución gaussiana. En el caso de las isocronas, el segundo supuesto no es válido, ya que los datos están sujetos a errores en las dos variables, y estos están comúnmente correlacionados. Cuando se toma en cuenta esta característica de los datos al desarrollar las expresiones de los parámetros de la recta, el problema es no-lineal y se resuelve iterativamente.

Existen varias publicaciones que han abordado el tema de la regresión lineal con errores correlacionados utilizando el método de mínimos cuadrados, entre estas podemos resaltar: York (1969), Minster *et al.* (1979) y Mahon (1996). Además, Titterton y Halliday (1979) abordaron el problema utilizando el método de máxima verosimilitud. En dos de estas publicaciones aparecen expresiones diferentes para calcular la pendiente de la línea de regresión (York, 1969; Minster *et al.*, 1979). En las publicaciones arriba mencionadas se han presentado expresiones diferentes para el cálculo de los errores en la pendiente y la ordenada al origen. Uno de los objetivos de esta tesis es el de determinar las diferencias entre las expresiones para el cálculo de los errores y aplicar

estos conceptos al cálculo de isocronas, con el propósito de proporcionar al usuario la mejor estimación de los errores en los parámetros de la recta.

En esta tesis se desarrollan las expresiones para la pendiente y la ordenada al origen, para demostrar que todas las soluciones coinciden algebraicamente en la línea de regresión. Asimismo, se demuestra que todas las soluciones que utilizan el método de mínimos cuadrados concuerdan en las expresiones para las incertidumbres de los parámetros de la recta, si se hacen las correcciones necesarias.

También se demuestra que la razón de que existan diferencias entre la solución de mínimos cuadrados y la solución de máxima verosimilitud es debido a que la primera propaga el error utilizando los datos observados, y la segunda, utilizando los datos ajustados. Esta última demostración se presenta respaldada por pruebas algebraicas y numéricas.

Como parte de esta tesis, se desarrolló el programa *IsoCalc 3.0* para facilitar la tarea de calcular y graficar la línea de regresión y crear isocronas de diversos sistemas isotópicos.

Palabras clave: Geocronología, Isocrona, Regresión lineal, Mínimos cuadrados, Máxima verosimilitud, Correlación de errores, Propagación de errores.

ABSTRACT of the thesis of JONAS DE DIOS DE BASABE DELGADO, presented as partial requirement to obtain the Masters in Sciences degree in Geology. Ensenada, Baja California, Mexico, February 2002.

Straight-line regression with errors in both variables: applications in geochronology to Isochron calculation

ABSTRACT

There are a number of applications in science, where it is necessary to calculate a straight line that minimizes the distances of a data point set. An example is found in geochronology, where a set of variables (i.e., $^{39}\text{Ar}/^{36}\text{Ar}$ and $^{40}\text{Ar}/^{36}\text{Ar}$) is obtained experimentally. From these, important information can be extracted: the age can be determined from the slope, and the initial $^{40}\text{Ar}/^{36}\text{Ar}$, from the y-intercept. The data obtained experimentally are subject to errors, and it is important to take them into account in the final result.

When fitting a set of data points to a straight line, in most cases the following assumptions are made: the data define a straight line, only one of the variables is subject to errors, and these have a normal distribution. In the case of Isochron fitting, the second assumption is invalid, since the data are subject to errors in both variables, and these are usually correlated. When this characteristic of the data is taken into account when developing the expressions of the straight-line parameters, the problem is non linear, and is solved iteratively.

There are several publications on the subject of straight-line regression with correlated errors using the least squares estimation method (LSE), from these it can be noted: York (1969), Minster *et al.* (1979), and Mahon (1996). Titterington and Halliday, 1979 also discussed the maximum likelihood estimation method (MLE). Two of these publications presented different expressions for the slope of the regression line (York, 1969; Minster et al., 1979). Furthermore in the publications above mentioned, different expressions are given for the error in the slope and y-intercept. One of the objectives of this thesis is to determine the differences between the published expressions for calculating the errors and to apply this result in Isochron calculation with the aim to give the user the best estimation for the straight-line parameters.

In this thesis, the expressions for the slope and intercept are developed, demonstrating that all solutions come to a complete algebraic agreement for the slope and y-intercept. With respect to the uncertainties in the slope and y-intercept, it is also demonstrated that all the least squares method solutions published are the same, once some omissions, presumably typographic are taken into account.

Algebraic and numerical proofs are given to show that the differences between the LSE errors and MLE errors is entirely due to the fact that the first propagates the error from the observed data, while the second propagates the error from the adjusted data.

As part of this thesis, it was developed the program IsoCalc 3.0 to aid on the task of calculating and plotting the regression line and creating isochrones of different isotopic systems.

Keywords: Geochronology, Isochron, Straight-line regression, Least squares, Maximum Likelihood, Error correlation, Error propagation.

*“Dios hizo todo hermoso en su momento,
y puso en la mente humana el sentido del tiempo”
Salomón (ca. 971-931 a.C.)*

Agradecimientos

Estoy agradecido con muchas personas que me animaron y respaldaron en el desarrollo de este trabajo. Margarita López me apoyo, no sólo para realizar esta tesis, sino a lo largo de toda la maestría. A Enrique Gómez le agradezco haber participado y haberse entusiasmado con este proyecto, sus cursos e interesantes pláticas han influenciado profundamente este trabajo. Hugo Hidalgo y Bodo Weber me hicieron valiosas observaciones que se convirtieron en importantes mejoras para este documento.

Derek York y Norman Evensen compartieron conmigo sus hallazgos científicos y respondieron a mis inquietudes en la etapa más importante de mi investigación, les estoy agradecido por ello y por haber hecho tan agradable mi estancia en Toronto en Mayo de 2001.

Todo el personal y los estudiantes del CICESE fueron inmerecidamente amables y generosos conmigo durante mi estancia en esta institución, especialmente Guadalupe Martínez, quien siempre tuvo la mejor disposición para ayudarme y nunca me hizo sentir inoportuno (aunque yo se que aveces lo fui).

Agradezco el apoyo económico del CONACYT en los primeros dos años de la maestría, y del CICESE, en los últimos dos meses.

Luis Delgado es el responsable de que haya considerado estudiar la maestría que concluye con esta tesis, a él le agradezco su ejemplo, su entusiasmo y cada una de sus palabras, que tienen una resonancia especial en mi mente.

Mi familia tiene mucho que ver con esta tesis, el apoyo de mi mamá es el más generoso y el más discreto, siempre reparte las cosechas de su amor. Mis sobrinos y sobrinas, que no saben nada de regresiones lineales, pero bien que conocen esa vieja tradición de confabularse con su tío para todo tipo de travesuras. También agradezco el apoyo de mis hermanos, hermanas, cuñados y cuñadas, el de mis abuelitas, Beatriz y Celia, y el de mi suegra, Charito.

Le agradezco a Compañerismo Estudiantil por el apoyo espiritual y por la oportunidad de ser parte de este noble movimiento.

A Eunice, por creer en mí de esa manera tan dulcemente poderosa, todo lo que ella esté dispuesta a creer de mí tiene una virtud profética, en parte por eso estoy feliz de a compartir mi vida con ella.

Y a Jesucristo, el Héroe de la Historia, y de mi historia.

Índice General

I	Introducción	1
I.1	Antecedentes	2
I.1.1	Las isocronas en geocronología	2
I.1.2	Expresiones para calcular las isocronas	8
I.2	Organización de los capítulos	10
I.3	Objetivos	11
II	Regresión lineal por el método de mínimos cuadrados	12
II.1	Parámetros de la recta	13
II.2	Error en la pendiente	19
II.3	Error en la ordenada al origen	26
III	Regresión lineal por el método de máxima verosimilitud	32
III.1	Parámetros de la recta	33
III.2	Error en los parámetros de la recta	35
III.3	Relación entre los errores de MC y MV	36
III.4	Unificación de MC y MV para casos especiales	39
IV	Pruebas numéricas	42
IV.1	Convergencia con datos ajustados	42
IV.2	Convergencia con pesos especiales	44
IV.3	Datos que producen resultados divergentes	47
IV.3.1	Sesgo por el método jackknife	47
IV.3.2	Sesgo y correlación lineal	48
IV.4	Método de Monte Carlo para propagación de errores	50
V	Discusión	57

VI Conclusiones	59
Bibliografía	61
Apéndices	
A El programa <i>IsoCalc 3.0</i>	65
A.1 Controles de la ventana principal	67
A.1.1 Opciones del menú principal	67
A.1.2 Botones de la barra de herramientas	69
A.1.3 El área de trabajo	70
A.2 La ventana de la isocrona	72
A.3 Histogramas	74
A.4 Archivos de entrada	76
B Demostraciones complementarias	77
C Datos graficados	81

Índice de Figuras

1	Datos observados de Reed-i-2 y datos ajustados	45
2	Gráfica de sesgo contra Δ_{σ_b}	49
3	Correlación lineal de 129 grupos de datos.	49
4	Diagramas del algoritmo de Monte Carlo para propagación de errores. . .	51
5	Parámetros de la elipse de error.	52
6	Elipse de error y datos perturbados aleatoriamente.	52
7	Histogramas del algoritmo de Monte Carlo que utiliza los datos observados.	55
8	Histogramas del algoritmo de Monte Carlo que utiliza los datos ajustados.	56
A1	Ventana principal de <i>IsoCalc 3.0</i>	68
A2	Ventana de <i>IsoCalc 3.0</i> con la isocrona y los datos gráficos.	75
A3	Ventana del histograma de la pendiente	75
C1	Datos York-66, tomados de York (1966).	82
C2	Datos Mahon-1, tomados de Mahon (1996).	83
C3	Datos Mahon-2, tomados de Mahon (1996).	84
C4	Datos Reed-i-1, tomados de Reed (1989).	85
C5	Datos Reed-i-2, tomados de Reed (1989).	86

Índice de Tablas

I	Algunos ejemplos de sistemas de fechamiento y relaciones isotópicas que se grafican en abcisa y ordenada para las isocronas	4
II	6
III	Resumen de fórmulas en York (1969)	14
IV	Resumen de fórmulas en Minster <i>et al.</i> (1979)	15
V	Resumen de fórmulas en Mahon (1996)	15
VI	Resumen de fórmulas para los parámetros de la recta y sus errores	31
VII	Fórmulas de MV para los errores en los parámetros de la recta	33
VIII	Comparación de los errores calculados por las fórmulas de MC, las fórmulas de MV y los calculados con MC utilizando los datos ajustados.	43
IX	Referencias bibliográficas de los datos utilizados e información sobre las rectas calculadas.	43
X	Comparación de los errores calculados por las fórmulas de MC y las fórmulas de MV tomando en cuenta únicamente el error σ_{Y_i} (regresión de Y en X).	46
XI	Comparación de los errores calculados por las fórmulas de MC y las fórmulas de MV tomando en cuenta únicamente el error σ_{X_i} (regresión de X en Y).	46
XII	Comparación de los resultados de Monte Carlo con los errores calculados por las fórmulas de MC y las fórmulas de MV.	54
AI	Fórmulas de MC que utiliza el programa <i>IsoCalc 3.0</i>	66
AII	Fórmulas de MV que utiliza el programa <i>IsoCalc 3.0</i>	66
AIII	Ejemplo de un archivo TXT con datos isotópicos	76
CI	Resumen de resultados calculados para los datos que se utilizan como ejemplos	81

Capítulo I

Introducción

El problema de calcular una recta que minimice las distancias de un conjunto de datos se ha tratado de manera parcial en la literatura: la solución para calcular la recta cuando sólo una de las variables tiene errores es bien conocida (Taylor, 1982; Bevington, 1969; Daniel y Wood, 1980, por citar algunos ejemplos); pero en pocas ocasiones se toca el problema de la regresión lineal considerando el error en las dos variables, la independiente y la dependiente (Acton, 1959; York, 1966; McIntyre *et al.*, 1966; Reed, 1992; Press *et al.*, 1992; Macki, 1994), y cuando se desarrolla el tema, se realiza en la mayoría de los casos de una manera confusa o errónea (Press *et al.*, 1992).

Una de las razones por las que no es usualmente abordado el problema de la regresión lineal tomando en cuenta los errores en las dos variables es porque la solución a este problema no es lineal, no así cuando se supone que una de las variables está libre de error, lo cual es una ficción matemática en un sentido estricto (Acton, 1959). Hay muchas aplicaciones donde se pueden introducir errores sistemáticos en la solución, si se hace esta simplificación, por ejemplo, en geocronología, para calcular isocronas (York, 1966), en astronomía, para calibrar los colores de cúmulos estelares globulares en función de su tipo espectral (Reed, 1989), y en ingeniería eléctrica, al hacer identificación de sistemas, análisis y diseño de sistemas de control y procesamiento de señales (Macki, 1994).

Aun más limitadas son las ocasiones en las que se incluye la correlación de los errores

en el desarrollo de las expresiones para calcular la recta. Existen cuatro publicaciones que abordan este problema, todas ellas en relación con la geocronología: York (1969), Titterington y Halliday (1979), Minster *et al.* (1979) y Mahon (1996). Actualmente, hay una gran confusión sobre cuál de estas cuatro soluciones es la correcta, especialmente después de que Mahon (1996) propuso su solución al problema e hizo una comparación incompleta e inconsistente entre ésta y las soluciones de York (1969) y Titterington y Halliday (1979), destacando las diferencias entre ellas, pero argumentando pobremente en favor de la solución que se propone en ese artículo.

I.1 Antecedentes

I.1.1 Las isocronas en geocronología

En geocronología, los datos experimentales se presentan comúnmente en diagramas de correlación conocidos como isocronas. Por ejemplo, en el sistema $Rb-Sr$, se grafican las razones de $^{87}Rb/^{86}Sr$ en el eje de las abscisas y $^{87}Sr/^{86}Sr$ en el eje de las ordenadas. Si el conjunto de datos graficados corresponden a muestras de una misma edad geológica, con un mismo valor inicial de $^{87}Sr/^{86}Sr$ y han permanecido como sistemas cerrados (McIntyre *et al.*, 1966), entonces los datos graficados definen una línea recta de la forma

$$y = a + bx,$$

donde a es la ordenada al origen, que representa el valor de la relación $^{87}Sr/^{86}Sr$ al momento que el sistema se cierra e inicia el reloj isotópico, por ejemplo, en el momento de la cristalización, y b es la pendiente, la cual es función de la edad (t) y de la constante de decaimiento del ^{87}Rb (λ),

$$b = e^{\lambda t} - 1. \quad (1.1)$$

Dado un conjunto de datos ($^{87}Rb/^{86}Sr$, $^{87}Sr/^{86}Sr$), se pueden calcular los parámetros

de la recta, a y b , realizando una regresión lineal. Una vez calculada la pendiente, se puede obtener la edad de las muestras despejando t de la ecuación (1.1) de la siguiente manera

$$t = \frac{1}{\lambda} \ln(b + 1). \quad (1.2)$$

En la mayoría de los sistemas de fechamiento isotópico se pueden aplicar las isocronas de una forma similar, en la tabla I se muestran algunos ejemplos de sistemas de fechamiento y las relaciones isotópicas que se grafican en cada caso.

La interpretación geológica de una isocrona depende del sistema de fechamiento isotópico y de qué tanto cumplen los datos con las hipótesis, a continuación se discuten estos dos aspectos.

Tipos de isocronas y sistemas de fechamiento isotópico

En la tabla II se muestran las expresiones para calcular la edad y las constantes de decaimiento que se utilizan para algunos tipos de isocronas. Se puede observar en la tabla II que, aunque en la mayoría de los casos se utiliza la expresión (1.2) para calcular la edad, en algunos tipos de isocronas se utilizan expresiones diferentes, tal es el caso cuando los isótopos radiactivos tienen varios tipos de decaimiento isotópico. Por ejemplo, el ^{40}K y el ^{138}La son isótopos padres que producen cada uno dos isótopos hijo según el mecanismo de desintegración radiactiva, y requieren el uso de diferentes constantes de decaimiento para calcular la edad dependiendo del isótopo hijo en cuestión (Faure, 1986).

Se puede observar en la tabla II que existen dos tipos de isocronas de $^{40}\text{Ar} - ^{39}\text{Ar}$, conocidas como isocrona normal e isocrona inversa. En la isocrona normal la edad se calcula utilizando el valor de la pendiente y en la isocrona inversa, utilizando la abcisa al origen, la cual está dada por $c = -a/b$. Para calcular la edad de las isocronas de $^{40}\text{Ar} - ^{39}\text{Ar}$ es necesario conocer el valor del factor \mathcal{J} , el cual está dado por

$$\mathcal{J} = \frac{e^{\lambda t_m} - 1}{(^{40}\text{Ar}^*/^{39}\text{Ar}_K)_m}$$

Tabla I: Algunos ejemplos de sistemas de fechamiento y relaciones isotópicas que se grafican en abcisa y ordenada para las isocronas (ver: Faure, 1986; Dickin, 2000).

Sistema	Abcisa	Ordenada
<i>Rb - Sr</i>	$^{87}\text{Rb}/^{86}\text{Sr}$	$^{87}\text{Sr}/^{86}\text{Sr}$
<i>Sm - Nd</i>	$^{147}\text{Sm}/^{144}\text{Nd}$	$^{143}\text{Nd}/^{144}\text{Nd}$
<i>Lu - Hf</i>	$^{176}\text{Lu}/^{177}\text{Hf}$	$^{176}\text{Hf}/^{177}\text{Hf}$
<i>Re - Os</i>	$^{187}\text{Re}/^{186}\text{Os}$	$^{187}\text{Os}/^{186}\text{Os}$
<i>K - Ca</i>	$^{40}\text{K}/^{42}\text{Ca}$	$^{40}\text{Ca}/^{42}\text{Ca}$
<i>K - Ar</i>	$^{40}\text{K}/^{36}\text{Ar}$	$^{40}\text{Ar}/^{36}\text{Ar}$
$^{40}\text{Ar} - ^{39}\text{Ar}$ (isocr. normal)	$^{39}\text{Ar}/^{36}\text{Ar}$	$^{40}\text{Ar}/^{36}\text{Ar}$
$^{40}\text{Ar} - ^{39}\text{Ar}$ (isocr. inversa)	$^{39}\text{Ar}/^{40}\text{Ar}$	$^{36}\text{Ar}/^{40}\text{Ar}$
<i>Pb - Pb</i>	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$
<i>Pt - Os</i>	$^{190}\text{Pt}/^{188}\text{Os}$	$^{186}\text{Os}/^{188}\text{Os}$
<i>Th - Pb</i>	$^{232}\text{Th}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$
$^{238}\text{U} - ^{206}\text{Pb}$	$^{238}\text{U}/^{204}\text{Pb}$	$^{206}\text{Pb}/^{204}\text{Pb}$
$^{235}\text{U} - ^{207}\text{Pb}$	$^{235}\text{U}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$
<i>U - Th</i>	$^{234}\text{U}/^{232}\text{Th}$	$^{230}\text{Th}/^{232}\text{Th}$
<i>La - Ba</i>	$^{138}\text{La}/^{137}\text{Ba}$	$^{138}\text{Ba}/^{137}\text{Ba}$
<i>La - Ce</i>	$^{138}\text{La}/^{142}\text{Ce}$	$^{138}\text{Ce}/^{142}\text{Ce}$

y se calcula utilizando un monitor de irradiación de edad conocida (t_m) (McDougall y Harrison, 1988; Faure, 1986).

También se observa en la tabla II que la expresión para calcular la edad de las isocronas de $Pb - Pb$ es una función implícita, esta no se puede despejar para t sino que se debe de resolver numéricamente utilizando un algoritmo iterativo (Faure, 1986).

En todos los métodos geocronológicos mencionados es necesario realizar una regresión lineal para encontrar los parametros de la recta y sus incertidumbres, el tipo de regresión lineal que se debe de utilizar para calcular las isocronas tiene características particulares y es el tema principal de esta tesis.

A continuación se señalan algunos problemas generales que se encuentran al interpretar las isocronas en los métodos de fechamiento basados en radioisótopos. Para una discusión más amplia se sugiere al lector consultar los textos especializados en geología isotópica (*i.e.* Faure, 1986; McDougall y Harrison, 1988; Dickin, 2000).

Interpretación de las isocronas

Las hipótesis necesarias para que la edad calculada de una isocrona tenga un significado geológico son las siguientes:

1. Las muestras corresponden a una misma edad,
2. parten de una misma relación isotópica inicial (según el sistema de fechamiento isotópico), y
3. han permanecido como un sistema cerrado, por lo menos para los isótopos en cuestión.

Con relación a la primer hipótesis, se debe de tener la certeza (por ejemplo, por observaciones de campo) de que los datos son cogenéticos, si existe algún grado de incertidumbre se pueden calcular regresiones independientes para los grupos de datos que se conoce que son cogenéticos y comparar los resultados de las regresiones realizadas

Tabla II: Expresiones para calcular la edad y constantes de decaimiento de algunos tipos de isocronas (ver: Faure, 1986; Dickin, 2000).

Tipo de isocrona	Edad	Constante(s) de decaimiento
<i>Rb - Sr</i>	$t = \frac{1}{\lambda} \ln(b + 1)$	$\lambda = 1.42 \times 10^{-11}$
<i>Sm - Nd</i>	$t = \frac{1}{\lambda} \ln(b + 1)$	$\lambda = 6.54 \times 10^{-12}$
<i>Lu - Hf</i>	$t = \frac{1}{\lambda} \ln(b + 1)$	$\lambda = 1.94 \times 10^{-11}$
<i>Re - Os</i>	$t = \frac{1}{\lambda} \ln(b + 1)$	$\lambda = 1.64 \times 10^{-11}$
<i>K - Ca</i>	$t = \frac{1}{\lambda_T} \ln\left(\frac{\lambda_T}{\lambda_\beta} b + 1\right)$	$\lambda_T = 5.543 \times 10^{-10} \quad \lambda_\beta = 4.962 \times 10^{-10}$
<i>K - Ar</i>	$t = \frac{1}{\lambda_T} \ln\left(\frac{\lambda_T}{\lambda_{ec}} b + 1\right)$	$\lambda_T = 5.543 \times 10^{-10} \quad \lambda_{ec} = 0.581 \times 10^{-10}$
⁴⁰ Ar - ³⁹ Ar isocr. normal	$t = \frac{1}{\lambda} \ln(\mathcal{J}b + 1)$	$\lambda = 5.543 \times 10^{-10}$
⁴⁰ Ar - ³⁹ Ar isocr. inversa	$t = \frac{1}{\lambda} \ln(\mathcal{J}c + 1)$	$\lambda = 5.543 \times 10^{-10}$
<i>Pb - Pb</i>	$b = \frac{1}{137.88} \left(\frac{e^{\lambda_2 t} - 1}{e^{\lambda_1 t} - 1} \right)$	$\lambda_1 = 1.55125 \times 10^{-10} \quad \lambda_2 = 9.8485 \times 10^{-10}$
<i>Pt - Os</i>	$t = \frac{1}{\lambda} \ln(b + 1)$	$\lambda = 7.9 \times 10^{-13}$
<i>Th - Pb</i>	$t = \frac{1}{\lambda} \ln(b + 1)$	$\lambda = 0.49475 \times 10^{-10}$
²³⁸ U - ²⁰⁶ Pb	$t = \frac{1}{\lambda} \ln(b + 1)$	$\lambda = 1.55125 \times 10^{-10}$
²³⁵ U - ²⁰⁷ Pb	$t = \frac{1}{\lambda} \ln(b + 1)$	$\lambda = 9.8485 \times 10^{-10}$
<i>U - Th</i>	$t = \frac{1}{\lambda} \ln(b + 1)$	$\lambda = 9.8485 \times 10^{-10}$
<i>La - Ba</i>	$t = \frac{1}{\lambda_T} \ln\left(\frac{\lambda_T}{\lambda_{ec}} b + 1\right)$	$\lambda_T = 6.73 \times 10^{-12} \quad \lambda_{ec} = 4.44 \times 10^{-12}$
<i>La - Ce</i>	$t = \frac{1}{\lambda_T} \ln\left(\frac{\lambda_T}{\lambda_\beta} b + 1\right)$	$\lambda_T = 6.73 \times 10^{-12} \quad \lambda_\beta = 2.29 \times 10^{-12}$

para cada grupo de datos (Roddick y Compston, 1977), si los resultados coinciden es recomendable hacer la regresión con todos los datos para reducir el error de los resultados. En algunos métodos es más fácil de garantizar que los datos pertenecen a una misma edad, por ejemplo, en el fechamiento por $^{40}\text{Ar} - ^{39}\text{Ar}$, utilizando la técnica de calentamiento por pasos, se utilizan concentrados minerales, por lo que puede suponerse que los datos representan una misma edad (McDougall y Harrison, 1988).

La segunda hipótesis es difícil de respaldar en algunos casos, por ejemplo, en el sistema $\text{Rb} - \text{Sr}$, se han observado que la dispersión en los datos es consecuencia de las variaciones en la relación inicial de $^{87}\text{Sr}/^{86}\text{Sr}$ (McIntyre *et al.*, 1966). Se ha propuesto asignar pesos especiales a los datos para compensar en la isocrona para este tipo de variaciones (McIntyre *et al.*, 1966), pero los resultados son insatisfactorios, por lo que es más recomendable en estos casos agrupar los datos y calcular regresiones independientes siempre que sea posible (Brooks y Compston, 1965). Una de las razones por las que puede dejar de cumplirse la hipótesis de homogeneidad es por que el magma original tiene heterogeneidades heredadas del manto (Hofmann y Hart, 1978).

Típicamente se ha utilizado el factor MSWD, como una medida de la dispersión, para determinar si los datos definen una isocrona (Brooks *et al.*, 1972). El valor esperado de MSWD es de 1 para isocronas con un número grande de datos y cuando los errores analíticos de los datos son calculados de un número grande de réplicas, de no ser este el caso, el valor esperado de MSWD es mayor que 1 (ver tablas de valores esperados para MSWD según el número de réplicas en Brooks *et al.*, 1972).

Si el factor MSWD resulta demasiado alto para una isocrona, entonces se dice que la dispersión es mayor que la esperada y no se puede justificar con los errores experimentales únicamente. Las causas de esta dispersión pueden atribuirse a que los datos no corresponden a una misma edad, tienen diferentes relaciones isotópicas iniciales o a que la muestra ha sufrido algún tipo de contaminación o redistribución de isótopos por procesos químicos (McIntyre *et al.*, 1966).

Una limitación del factor MSWD es que, cuando los errores asignados a los datos son mayores que el error analítico, el factor MSWD se reduce artificialmente indicando un buen ajuste cuando no es necesariamente así, este comportamiento particular del factor MSWD incluye la posibilidad de ser manipulado fácilmente sobrestimando los errores analíticos con el propósito de reducir el valor de MSWD. Otro indicador estadístico que puede ser útil para identificar un buen ajuste de los datos es el factor de correlación lineal, el cual está relacionado con la probabilidad de que los datos pertenezcan a una línea recta y se calcula en función de los datos únicamente, sin tomar en cuenta sus incertidumbres.

Por lo general, el factor MSWD es un buen indicador para identificar si la isocrona representa una edad con significado geológico, sin embargo, hay casos en los que los datos pueden tener un buen ajuste a una recta que no representa una edad, tal es el caso cuando las muestras son el producto de una mezcla entre dos componentes con diferentes relaciones isotópicas de Rb/Sr y $^{87}Sr/^{86}Sr$, en este caso los datos no cumplen con la tercer hipótesis y la línea de regresión representa una línea de mezcla y no una isocrona (Faure, 1986).

I.1.2 Expresiones para calcular las isocronas

Las isocronas son una herramienta de utilidad creciente en geocronología, sin embargo, las expresiones para calcular una recta a través de un conjunto de relaciones isotópicas han sido objeto de debate en la literatura.

La razón de que exista esta falta de acuerdo se deriva de que los datos para las isocronas están sujetos a errores en las abscisas y ordenadas (con una distribución gaussiana), además de que por lo general los errores en las abscisas y las ordenadas están correlacionados. Si se incorporan estas consideraciones en el desarrollo de las expresiones mínimo cuadráticas, el problema para encontrar la pendiente es no-lineal, y se resuelve iterativamente (York, 1969). Debido a la forma complicada de las expresiones

para la pendiente, la propagación de los errores para la pendiente y ordenada al origen requiere de un riguroso y detallado desarrollo algebraico que ha conducido a soluciones aparentemente diferentes.

La primera solución del problema que tomó en cuenta la correlación de los errores fue publicada por D. York (1969). Este autor presentó dos tipos de fórmulas para calcular los errores en la pendiente y la ordenada al origen: las expresiones "exactas", desarrolladas por el método general de propagación de errores, y las expresiones que describe como "aproximaciones razonables". En las expresiones aproximadas hay un error tipográfico (Titterington y Halliday, 1979) y las expresiones exactas se muestran expresadas en términos de derivadas parciales, de las cuales no todas aparecen explícitamente en este célebre artículo. Debido a estas razones, a esta publicación le han seguido ocasionales reformulaciones argumentando tener soluciones diferentes (Titterington y Halliday, 1979; Minster *et al.*, 1979; Mahon, 1996).

A pesar de las diferencias algebraicas de las soluciones publicadas, se ha observado que, al calcular la pendiente y la ordenada al origen de diferentes grupos de datos, todas las soluciones coinciden numéricamente (Titterington y Halliday, 1979; Mahon, 1996). También se ha observado que, al calcular los errores de los parámetros de la recta, las diferencias son pequeñas en la mayoría de los casos (Titterington y Halliday, 1979), pero en algunos casos, las diferencias son considerables (Mahon, 1996).

En esta tesis se pretende esclarecer el tema, demostrando algebraicamente que las ecuaciones publicadas para los parámetros de la recta son equivalentes, y explicando cuáles son las razones de que existan diferencias entre las ecuaciones para calcular las incertidumbres en los parámetros de la recta. También como parte de esta tesis, se ha desarrollado un programa que permite graficar la línea de regresión y crear isocronas de diferentes sistemas isotópicos.

I.2 Organización de los capítulos

El trabajo realizado se encuentra organizado de la siguiente manera:

- En el *presente capítulo* se encuentra una introducción del tema a tratar.
- En el *segundo capítulo* se desarrollan las expresiones para calcular los parámetros de la recta y sus incertidumbres, utilizando el método de mínimos cuadrados (MC), con el fin de demostrar que todas las soluciones publicadas que utilizan este método son equivalentes.
- En el *tercer capítulo* se explica la relación entre los métodos de máxima verosimilitud (MV) y MC para la regresión lineal con errores correlacionados, señalando los casos especiales en los que coinciden ambos métodos.
- El *cuarto capítulo* contiene los resultados de las pruebas numéricas, realizadas para respaldar las demostraciones del tercer capítulo.
- En el *quinto capítulo* se discuten los resultados.
- En el *sexto capítulo* se presentan las conclusiones.
- El *apéndice A* es un manual para el programa *IsoCalc 3.0*, el cual es de utilidad para calcular la regresión lineal descrita en esta tesis y para crear isocronas.
- En el *apéndice B* se desarrollan algunas demostraciones algebraicas, relacionadas con el material que se presenta en el segundo capítulo.
- Finalmente, en el *apéndice C*, se muestran algunos datos graficados para ilustrar los puntos discutidos.

I.3 Objetivos

- Establecer si existe una relación entre las soluciones publicadas para la regresión lineal con errores correlacionados.
- Presentar un desarrollo algebraico completo para las expresiones de los parámetros de la recta y sus errores, utilizando el método de mínimos cuadrados y el método general de propagación de errores.
- Desarrollar un programa amigable que permita calcular la línea de regresión, graficar los datos y crear isocronas de $^{40}\text{Ar} - ^{39}\text{Ar}$, $\text{Lu} - \text{Hf}$, $\text{Rb} - \text{Sr}$, $\text{Re} - \text{Os}$ y $\text{Sm} - \text{Nd}$.

Capítulo II

Regresión lineal por el método de mínimos cuadrados

El problema de la regresión lineal con errores correlacionados ha sido abordado en diferentes ocasiones en la literatura, pero no siempre satisfactoriamente. Una solución completa del problema debe incluir expresiones para estimar parámetros, errores en los parámetros y una medida estadística para aceptar o rechazar el ajuste (Press *et al.*, 1992). En este sentido, la solución de York (1969) ha sido la más citada en la literatura, especialmente en geocronología, no solo por haber sido la primera sino por ser una solución completa (tabla III). Sin embargo, Minster *et al.* (1979) y Mahon (1996) han presentado soluciones mínimo cuadráticas aparentemente diferentes para el mismo problema.

Minster *et al.* (1979) minimizó la misma expresión que York (1969) encontrando expresiones más compactas para la pendiente y los errores de los parámetros introduciendo el factor β_i (tabla IV), coincidiendo únicamente en la expresión para la ordenada al origen. En esta publicación no se establece la relación con las expresiones de York (1969), pero es posible demostrar que las expresiones de Minster *et al.* (1979) se derivan de las expresiones de York (1969), a excepción de $\partial\varphi/\partial b$, denominado Q^{-1} en Minster *et al.* (1979), debido a que hay un término faltante en esta expresión (posiblemente omitido

por un error tipográfico).

Mahon (1996) no propone expresiones diferentes para los parámetros de la recta y sus errores, sino que cita para esto a York (1969), pero desarrolla las derivadas parciales que se utilizan para calcular los errores y llega a resultados diferentes que los de York (1969). En sus expresiones hay una serie de omisiones que probablemente son simplificaciones que consideró el autor que introducían un error sistemático muy pequeño, pero no presenta ninguna discusión para justificar estas simplificaciones (tabla V).

En este capítulo se muestra el desarrollo detallado de mínimos cuadrados (MC) para demostrar que las soluciones publicadas, una vez corregidas, convergen algebraicamente¹.

II.1 Parámetros de la recta

El método de MC surgió originalmente de la necesidad de ajustar un modelo matemático lineal a un conjunto de observaciones (Björck, 1996). Tradicionalmente se ha supuesto, al utilizar este método, que la variable independiente no tiene errores, por lo que el problema consiste en minimizar las distancias verticales del dato a la recta, pero se ha señalado que este enfoque es inconsistente en algunos casos (e.g. Acton, 1959; York, 1966; Macki, 1994).

En esta sección se utiliza el método de MC considerando que se conocen las incertidumbres en las dos variables y su covarianza. Utilizando la terminología de York (1969), la función a minimizar es

$$S = \sum \frac{(Y_i - a - bX_i)^2}{\text{Var}(Y_i - a - bX_i)}, \quad (2.1)$$

donde a y b son los parámetros de la recta, (X_i, Y_i) son los datos observados, y $\text{Var}(Y_i -$

¹Algunas partes del desarrollo presentado en este capítulo fueron tomadas de las notas personales de D. York, de la Universidad de Toronto, a quien le agradezco el haberme dado acceso a sus manuscritos originales.

Tabla III: Resumen de fórmulas en York (1969), corregido el error tipográfico en la expresión aproximada de σ_a^2 .

$$\hat{a} = \bar{Y} - b\bar{X}$$

$$\hat{b} = \frac{-\sum Z_i^2 \left(\frac{U_i^2}{\omega_{Y_i}} - \frac{V_i^2}{\omega_{X_i}} \right) + \sqrt{\left[\sum Z_i^2 \left(\frac{U_i^2}{\omega_{Y_i}} - \frac{V_i^2}{\omega_{X_i}} \right) \right]^2 + 4 \sum Z_i^2 \left(\frac{U_i V_i}{\omega_{X_i}} - \frac{r_i U_i^2}{\alpha_i} \right) \sum Z_i^2 \left(\frac{U_i V_i}{\omega_{Y_i}} - \frac{r_i V_i^2}{\alpha_i} \right)}}{2 \sum Z_i^2 \left(\frac{U_i V_i}{\omega_{X_i}} - \frac{r_i U_i^2}{\alpha_i} \right)}$$

$$\varphi = b^2 \sum Z_i^2 \left(\frac{U_i V_i}{\omega_{X_i}} - \frac{r_i U_i^2}{\alpha_i} \right) + b \sum Z_i^2 \left(\frac{U_i^2}{\omega_{Y_i}} - \frac{V_i^2}{\omega_{X_i}} \right) - \sum Z_i^2 \left(\frac{U_i V_i}{\omega_{Y_i}} - \frac{r_i V_i^2}{\alpha_i} \right) = 0$$

Errores en los parámetros:

$$\hat{\sigma}_a^2 = \sum \left[\frac{1}{\omega_{X_i}} \left(\frac{\partial a}{\partial X_i} \right)^2 + \frac{1}{\omega_{Y_i}} \left(\frac{\partial a}{\partial Y_i} \right)^2 + 2 \frac{r_i}{\alpha_i} \left(\frac{\partial a}{\partial X_i} \right) \left(\frac{\partial a}{\partial Y_i} \right) \right]$$

$$\hat{\sigma}_b^2 = \sum \left[\frac{1}{\omega_{X_i}} \left(\frac{\partial \varphi}{\partial X_i} \right)^2 + \frac{1}{\omega_{Y_i}} \left(\frac{\partial \varphi}{\partial Y_i} \right)^2 + 2 \frac{r_i}{\alpha_i} \left(\frac{\partial \varphi}{\partial X_i} \right) \left(\frac{\partial \varphi}{\partial Y_i} \right) \right] / \left(\frac{\partial \varphi}{\partial b} \right)^2$$

“Aproximaciones razonables” para los errores de los parámetros:

$$\hat{\sigma}_a^2 = \sigma_b^2 \frac{\sum Z_i X_i^2}{\sum Z_i}$$

$$\hat{\sigma}_b^2 = \frac{1}{\sum U_i^2 Z_i}$$

$$\frac{\partial \varphi}{\partial b} = \sum Z_i^2 \left[U_i^2 \left(\frac{1}{\omega_{Y_i}} - \frac{2br_i}{\alpha_i} \right) + \frac{V_i}{\omega_{X_i}} (2bU_i - V_i) \right] + b^2 \left(\frac{\partial A}{\partial b} \right) + b \left(\frac{\partial B}{\partial b} \right) - \left(\frac{\partial C}{\partial b} \right)$$

$$\frac{\partial \varphi}{\partial X_i} = \sum_j Z_j^2 \left(\delta_i^j - \frac{Z_i}{\sum_j Z_j} \right) \left(\frac{b^2 V_j}{\omega_{X_j}} - \frac{2b^2 r_j U_j}{\alpha_j} + \frac{2bU_j}{\omega_{Y_j}} - \frac{V_j}{\omega_{Y_j}} \right)$$

$$\frac{\partial \varphi}{\partial Y_i} = \sum_j Z_j^2 \left(\delta_i^j - \frac{Z_i}{\sum_j Z_j} \right) \left(\frac{b^2 U_j}{\omega_{X_j}} - \frac{2bV_j}{\omega_{X_j}} - \frac{U_j}{\omega_{Y_j}} + \frac{2r_j V_j}{\alpha_j} \right)$$

$$\frac{\partial a}{\partial X_i} = \frac{-bZ_i}{\sum_j Z_j} + \left[\frac{2}{\sum_j Z_j} \sum_j \frac{Z_j^2}{\alpha_j^2} (r_j \alpha_j - b\omega_{Y_j}) (V_j - bU_j) - \bar{X} \right] \left(-\frac{\partial \varphi}{\partial X_i} \right) / \left(\frac{\partial \varphi}{\partial b} \right)$$

$$\frac{\partial a}{\partial Y_i} = \frac{Z_i}{\sum_j Z_j} + \left[\frac{2}{\sum_j Z_j} \sum_j \frac{Z_j^2}{\alpha_j^2} (r_j \alpha_j - b\omega_{Y_j}) (V_j - bU_j) - \bar{X} \right] \left(-\frac{\partial \varphi}{\partial Y_i} \right) / \left(\frac{\partial \varphi}{\partial b} \right)$$

Tabla IV: Resumen de fórmulas en Minster *et al.* (1979), utilizando la terminología de York (1969).

$$\begin{aligned}\hat{a} &= \bar{Y} - b\bar{X} \\ \hat{b} &= \frac{\sum Z_i V_i \beta_i}{\sum Z_i U_i \beta_i} \\ \hat{\sigma}_a^2 &= \frac{1}{\sum Z_i} + 2(\bar{X} + 2\bar{\beta})\bar{\beta}Q + \sigma_b^2(\bar{X} + 2\bar{\beta})^2 \\ \hat{\sigma}_b^2 &= Q^2 \sum Z_i^2 \left[\frac{U_i^2}{\omega_{Y_i}} + \frac{V_i^2}{\omega_{X_i}} - 2\frac{r_i}{\alpha_i} U_i V_i \right] \\ Q^{-1} &= \sum Z_i \left[\frac{U_i V_i}{b} + 4(\beta_i - \bar{\beta})(\beta_i - U_i) \right] \\ \beta_i &= Z_i \left(\frac{U_i}{\omega_{Y_i}} + \frac{bV_i}{\omega_{X_i}} - (bU_i + V_i)\frac{r_i}{\alpha_i} \right)\end{aligned}$$

Tabla V: Resumen de fórmulas en Mahon (1996). El resto de las fórmulas de esta publicación son citadas de York (1969), tabla III.

$$\begin{aligned}\frac{\partial \varphi}{\partial b} &= \sum Z_i^2 \left[2b(U_i V_i \sigma_{X_i}^2 - U_i^2 \sigma_{XY_i}) + U_i^2 \sigma_{Y_i}^2 - V_i^2 \sigma_{X_i}^2 \right] + \\ & 4 \sum Z_i^3 (\sigma_{XY_i} - b\sigma_{X_i}^2) \left[b^2 (U_i V_i \sigma_{X_i}^2 - U_i^2 \sigma_{XY_i}) + \right. \\ & \left. b(U_i^2 \sigma_{Y_i}^2 - V_i^2 \sigma_{X_i}^2) - (U_i V_i \sigma_{Y_i}^2 - V_i^2 \sigma_{XY_i}) \right] \\ \frac{\partial a}{\partial X_i} &= \frac{-bZ_i}{\sum Z_j} - \bar{X} \left(\frac{\partial \varphi}{\partial X_i} \right) / \left(\frac{\partial \varphi}{\partial b} \right) \\ \frac{\partial a}{\partial Y_i} &= \frac{Z_i}{\sum Z_j} - \bar{X} \left(\frac{\partial \varphi}{\partial Y_i} \right) / \left(\frac{\partial \varphi}{\partial b} \right)\end{aligned}$$

$a - bX_i$) es la varianza total de los datos

$$\begin{aligned}\text{Var}(Y_i - a - bX_i) &= \text{Var}(Y_i) + b^2\text{Var}(X_i) - 2b\text{Var}(X_iY_i) \\ &= \sigma_{Y_i}^2 + b^2\sigma_{X_i}^2 - 2b\sigma_{XY_i} = \frac{1}{Z_i}.\end{aligned}$$

Donde Z_i es el peso total, el cual es posible reexpresar como

$$\begin{aligned}Z_i &= \frac{1}{\sigma_{Y_i}^2 + b^2\sigma_{X_i}^2 - 2b\sigma_{XY_i}} \\ &= \frac{\omega_{X_i}\omega_{Y_i}}{\omega_{X_i} + b^2\omega_{Y_i} - 2br_i\alpha_i},\end{aligned}\tag{2.2}$$

y donde r_i es el factor de correlación entre los errores,

$$r_i = \frac{\sigma_{XY_i}}{\sigma_{X_i}\sigma_{Y_i}}, \quad \omega_{X_i} = \frac{1}{\sigma_{X_i}^2}, \quad \omega_{Y_i} = \frac{1}{\sigma_{Y_i}^2}, \quad y \quad \alpha_i = \sqrt{\omega_{X_i}\omega_{Y_i}}.\tag{2.3}$$

Con base en lo anterior, se puede reexpresar (2.1) en función de Z_i de la siguiente manera

$$S = \sum Z_i(Y_i - a - bX_i)^2.\tag{2.4}$$

Para obtener los parámetros de la recta que minimizan a S se igualan a cero las derivadas parciales con respecto a los parámetros. Primeramente, derivando con respecto a a

$$\begin{aligned}\frac{\partial S}{\partial a} &= -2 \sum Z_i(Y_i - a - bX_i) = 0 \\ a \sum Z_i + b \sum Z_iX_i - \sum Z_iY_i &= 0.\end{aligned}$$

Despejando, se encuentra que el estimador de a que minimiza a S es

$$\begin{aligned}\hat{a} &= \frac{\sum Z_iY_i}{\sum Z_i} - b \frac{\sum Z_iX_i}{\sum Z_i}, \\ \hat{a} &= \bar{Y} - b\bar{X},\end{aligned}\tag{2.5}$$

donde \bar{X} y \bar{Y} son los promedios ponderados de X_i y Y_i respectivamente, demostrando así que la recta pasa por el centroide.

Derivando a S con respecto a b

$$\frac{\partial S}{\partial b} = -2 \sum Z_i X_i (Y_i - a - bX_i) + \sum (Y_i - a - bX_i)^2 \left(\frac{\partial Z_i}{\partial b} \right) = 0 \quad (2.6)$$

donde (ver apéndice B)

$$\frac{\partial Z_i}{\partial b} = -2Z_i^2 \left(\frac{b\omega_{Y_i} - r_i\alpha_i}{\omega_{X_i}\omega_{Y_i}} \right).$$

Se puede sustituir el estimador de a para hacer más compactas las expresiones

$$\begin{aligned} Y_i - a - bX_i &= Y_i - \bar{Y} + b\bar{X} - bX_i \\ &= V_i - bU_i, \end{aligned} \quad (2.7)$$

donde $U_i = X_i - \bar{X}$ y $V_i = Y_i - \bar{Y}$. Sustituyendo en (2.6) se obtiene que

$$\sum Z_i X_i (V_i - bU_i) + 2 \sum Z_i^2 \left(\frac{b\omega_{Y_i} - r_i\alpha_i}{\omega_{X_i}\omega_{Y_i}} \right) (V_i - bU_i)^2 = 0,$$

desarrollando esta expresión se obtiene la expresión cúbica de la pendiente (York, 1969)

$$\begin{aligned} &b^3 \sum \frac{Z_i^2 U_i^2}{\omega_{X_i}} - b^2 \sum Z_i^2 \left(\frac{2U_i V_i}{\omega_{X_i}} + \frac{r_i U_i^2}{\alpha_i} \right) \\ &- b \sum \left(Z_i U_i^2 - \frac{2Z_i^2 r_i U_i V_i}{\alpha_i} - \frac{Z_i^2 V_i^2}{\omega_{X_i}} \right) + \sum \left(Z_i U_i V_i - \frac{Z_i^2 r_i V_i^2}{\alpha_i} \right) = 0. \end{aligned}$$

La expresión cúbica se puede reducir a una cuadrática (York, 1969) haciendo las siguientes sustituciones

$$\begin{aligned} b \sum Z_i U_i^2 &= b \sum Z_i^2 U_i^2 \left(\frac{1}{\omega_{Y_i}} + \frac{b^2}{\omega_{X_i}} - \frac{2br_i}{\alpha_i} \right) \\ &= b^3 \sum \frac{Z_i^2 U_i^2}{\omega_{X_i}} - 2b^2 \sum \frac{Z_i^2 r_i U_i^2}{\alpha_i} + b \sum \frac{Z_i^2 U_i^2}{\omega_{Y_i}}, \end{aligned} \quad (2.8)$$

y

$$\begin{aligned} \sum Z_i U_i V_i &= \sum Z_i^2 U_i V_i \left(\frac{1}{\omega_{Y_i}} + \frac{b^2}{\omega_{X_i}} - \frac{2br_i}{\alpha_i} \right) \\ &= b^2 \sum \frac{Z_i^2 U_i V_i}{\omega_{X_i}} - 2b \sum \frac{Z_i^2 r_i U_i V_i}{\alpha_i} + \sum \frac{Z_i^2 U_i V_i}{\omega_{Y_i}}. \end{aligned} \quad (2.9)$$

Sustituyendo primero (2.8) en la expresión cúbica se tiene que

$$b^3 \sum \frac{Z_i^2 U_i^2}{\omega_{x_i}} - b^2 \sum Z_i^2 \left(\frac{2U_i V_i}{\omega_{x_i}} + \frac{r_i U_i^2}{\alpha_i} \right) + b \sum Z_i^2 \left(\frac{V_i^2}{\omega_{x_i}} + \frac{2r_i U_i V_i}{\alpha_i} \right) + \sum \left(Z_i U_i V_i - \frac{Z_i^2 r_i V_i^2}{\alpha_i} \right) - b^3 \sum \frac{Z_i^2 U_i^2}{\omega_{x_i}} + 2b^2 \sum \frac{Z_i^2 r_i U_i^2}{\alpha_i} - b \sum \frac{Z_i^2 U_i^2}{\omega_{y_i}} = 0,$$

y desarrollando

$$-b^2 \sum Z_i^2 \left(\frac{2U_i V_i}{\omega_{x_i}} - \frac{r_i U_i^2}{\alpha_i} \right) + b \sum Z_i^2 \left(\frac{V_i^2}{\omega_{x_i}} + \frac{2r_i U_i V_i}{\alpha_i} - \frac{U_i^2}{\omega_{y_i}} \right) + \sum \left(Z_i U_i V_i - \frac{Z_i^2 r_i V_i^2}{\alpha_i} \right) = 0.$$

Sustituyendo (2.9) en la expresión cuadrática anterior, se tiene que se debe cumplir

$$-b^2 \sum Z_i^2 \left(\frac{2U_i V_i}{\omega_{x_i}} - \frac{r_i U_i^2}{\alpha_i} \right) + b \sum Z_i^2 \left(\frac{V_i^2}{\omega_{x_i}} + \frac{2r_i U_i V_i}{\alpha_i} - \frac{U_i^2}{\omega_{y_i}} \right) - \sum \frac{Z_i^2 r_i V_i^2}{\alpha_i} + b^2 \sum \frac{Z_i^2 U_i V_i}{\omega_{x_i}} - 2b \sum \frac{Z_i^2 r_i U_i V_i}{\alpha_i} + \sum \frac{Z_i^2 U_i V_i}{\omega_{y_i}} = 0.$$

Definiendo φ como

$$\varphi = b^2 \sum Z_i^2 \left(\frac{U_i V_i}{\omega_{x_i}} - \frac{r_i U_i^2}{\alpha_i} \right) + b \sum Z_i^2 \left(\frac{U_i^2}{\omega_{y_i}} - \frac{V_i^2}{\omega_{x_i}} \right) - \sum Z_i^2 \left(\frac{U_i V_i}{\omega_{y_i}} - \frac{r_i V_i^2}{\alpha_i} \right),$$

entonces se debe cumplir que $\varphi = 0$.

Esta expresión se puede resolver para b utilizando la fórmula cuadrática (York, 1969) o se puede despejar la pendiente linealizando para obtener la expresión de Minster *et al.* (1979), sumandole a φ el término

$$b \sum Z_i^2 \frac{U_i V_i r_i}{\alpha_i} - b \sum Z_i^2 \frac{U_i V_i r_i}{\alpha_i} = 0$$

y reordenando los términos para obtener

$$\varphi = b \sum Z_i^2 U_i \left(\frac{U_i}{\omega_{y_i}} + \frac{bV_i}{\omega_{x_i}} - (bU_i + V_i) \frac{r_i}{\alpha_i} \right) - \sum Z_i^2 V_i \left(\frac{U_i}{\omega_{y_i}} + \frac{bV_i}{\omega_{x_i}} - (bU_i + V_i) \frac{r_i}{\alpha_i} \right) = 0.$$

Si se define

$$\beta_i = Z_i \left(\frac{U_i}{\omega_{y_i}} + \frac{bV_i}{\omega_{x_i}} - (bU_i + V_i) \frac{r_i}{\alpha_i} \right), \quad (2.10)$$

entonces, al sustituir β_i en φ , la expresión para b se reduce a la forma más simple

$$\varphi = b \sum Z_i \beta_i U_i - \sum Z_i \beta_i V_i = 0 \quad (2.11)$$

de donde se obtiene el estimador de b

$$\hat{b} = \frac{\sum Z_i \beta_i V_i}{\sum Z_i \beta_i U_i} \quad (2.12)$$

Esta expresión para \hat{b} , que corresponde perfectamente con la de Minster *et al.* (1979), es aparentemente lineal con respecto a los datos pero, aunque la expresión lo disimula de una forma elegante, el valor de b se utiliza para calcular los valores de Z_i y β_i , por lo que la pendiente se debe de calcular de forma recursiva, utilizando un valor inicial para la pendiente e iterando hasta que el valor de la pendiente haya convergido con la precisión necesaria.

Los resultados numéricos muestran que se obtiene el mismo valor para la pendiente iterando con la fórmula cúbica, cuadrática o lineal, respaldando el desarrollo algebraico anterior que demuestra que las expresiones para la pendiente publicadas en York (1969) y Minster *et al.* (1979) son equivalentes.

II.2 Error en la pendiente

Utilizando el método general de propagación de errores (Young, 1962; Bevington, 1969; Meyer, 1975) para calcular el error en la pendiente, se expande en serie de Taylor la función (2.12)

$$\begin{aligned} \hat{b} = \hat{b}_o &+ \sum \frac{\partial \hat{b}}{\partial X_i} (X_i - \delta X_i) + \sum \frac{\partial \hat{b}}{\partial Y_i} (Y_i - \delta Y_i) \\ &+ \sum \frac{\partial^2 \hat{b}}{\partial^2 X_i} \frac{(X_i - \delta X_i)^2}{2} + \sum \frac{\partial^2 \hat{b}}{\partial^2 Y_i} \frac{(Y_i - \delta Y_i)^2}{2} + \dots \end{aligned}$$

donde \hat{b}_o es la pendiente calculada con los datos observados. Suponiendo que la pendiente es aproximadamente lineal cuando los datos tienen variaciones pequeñas, se pueden

eliminar los términos de orden superior en la serie de Taylor, de tal manera que

$$\hat{b} = \hat{b}_o + \sum \frac{\partial \hat{b}}{\partial X_i} (X_i - \delta X_i) + \sum \frac{\partial \hat{b}}{\partial Y_i} (Y_i - \delta Y_i),$$

$$\hat{b} - \hat{b}_o = \sum \frac{\partial \hat{b}}{\partial X_i} (X_i - \delta X_i) + \sum \frac{\partial \hat{b}}{\partial Y_i} (Y_i - \delta Y_i).$$

Se observa que la varianza de \hat{b} es

$$\hat{\sigma}_b^2 = E [\hat{b} - \hat{b}_o]^2$$

$$= E \left[\sum \frac{\partial \hat{b}}{\partial X_i} (X_i - \delta X_i) + \sum \frac{\partial \hat{b}}{\partial Y_i} (Y_i - \delta Y_i) \right]^2,$$

donde E indica valor esperado. Sustituyendo $X_i - \delta X_i = \Delta X_i$ y $Y_i - \delta Y_i = \Delta Y_i$ y desarrollando

$$\hat{\sigma}_b^2 = E \left[\sum \frac{\partial \hat{b}}{\partial X_i} \Delta X_i \right]^2 + E \left[\sum \frac{\partial \hat{b}}{\partial Y_i} \Delta Y_i \right]^2 + 2E \left[\sum \frac{\partial \hat{b}}{\partial X_i} \Delta X_i \sum \frac{\partial \hat{b}}{\partial Y_i} \Delta Y_i \right] \quad (2.13)$$

Si los datos son independientes, entonces $\text{Cov}(X_i, X_j) = 0$, de esta manera

$$E \left[\frac{\partial \hat{b}}{\partial X_i} \frac{\partial \hat{b}}{\partial X_j} \Delta X_i \Delta X_j \right] = \frac{\partial \hat{b}}{\partial X_i} \frac{\partial \hat{b}}{\partial X_j} E [\Delta X_i \Delta X_j]$$

$$= \frac{\partial \hat{b}}{\partial X_i} \frac{\partial \hat{b}}{\partial X_j} \text{Cov}(X_i, X_j)$$

$$= 0. \quad (2.14)$$

Asimismo, $\text{Cov}(Y_i, Y_j) = 0$ y $\text{Cov}(X_i, Y_j) = 0$, por lo que

$$E \left[\frac{\partial \hat{b}}{\partial Y_i} \frac{\partial \hat{b}}{\partial Y_j} \Delta Y_i \Delta Y_j \right] = 0, \quad (2.15)$$

$$E \left[\frac{\partial \hat{b}}{\partial X_i} \frac{\partial \hat{b}}{\partial Y_j} \Delta X_i \Delta Y_j \right] = 0. \quad (2.16)$$

Utilizando (2.14), (2.15) y (2.16) al desarrollar (2.13), se tiene que

$$\hat{\sigma}_b^2 = \sum E \left[\frac{\partial \hat{b}}{\partial X_i} \Delta X_i \right]^2 + \sum E \left[\frac{\partial \hat{b}}{\partial Y_i} \Delta Y_i \right]^2 + 2 \sum E \left[\frac{\partial \hat{b}}{\partial X_i} \frac{\partial \hat{b}}{\partial Y_i} \Delta X_i \Delta Y_i \right]$$

$$= \sum E \left(\frac{\partial \hat{b}}{\partial X_i} \right)^2 E(\Delta X_i)^2 + \sum E \left(\frac{\partial \hat{b}}{\partial Y_i} \right)^2 E(\Delta Y_i)^2$$

$$\begin{aligned}
 & +2 \sum E \left(\frac{\partial \hat{b}}{\partial X_i} \right) E \left(\frac{\partial \hat{b}}{\partial Y_i} \right) E (\Delta X_i \Delta Y_i) \\
 & = \sum \left(\frac{\partial \hat{b}}{\partial X_i} \right)^2 \sigma_{x_i}^2 + \sum \left(\frac{\partial \hat{b}}{\partial Y_i} \right)^2 \sigma_{y_i}^2 + 2 \sum \left(\frac{\partial \hat{b}}{\partial X_i} \right) \left(\frac{\partial \hat{b}}{\partial Y_i} \right) \sigma_{x y_i}. \quad (2.17)
 \end{aligned}$$

Al evaluar las derivadas parciales en su valor esperado, se deben de sustituir los valores esperados de los datos. El resultado de sustituir el valor esperado por el dato ajustado se desarrolla en el siguiente capítulo, al considerar el método de máxima verosimilitud. En el caso presente, se utilizan los datos observados en las derivadas parciales, tal como se propone en Meyer (1975), que es la estrategia adoptada en York (1969), Minster *et al.* (1979) y Mahon (1996).

Utilizando las definiciones (2.3), la ecuación (2.17) queda expresada como

$$\hat{\sigma}_b^2 = \sum \left\{ \left(\frac{\partial \hat{b}}{\partial X_i} \right)^2 \frac{1}{\omega_{x_i}} + \left(\frac{\partial \hat{b}}{\partial Y_i} \right)^2 \frac{1}{\omega_{y_i}} + \left(\frac{\partial \hat{b}}{\partial X_i} \right) \left(\frac{\partial \hat{b}}{\partial Y_i} \right) \frac{2r_i}{\alpha_i} \right\}. \quad (2.18)$$

Como \hat{b} es una función implícita, sería conveniente reexpresar $\hat{\sigma}_b^2$ en función de derivadas parciales de φ . Para $\partial \hat{b} / \partial X_i$ se toma la derivada total de φ tomando como constantes todas las variables excepto X_i (*i.e.* $Y_j = C \quad \forall j = 1 \dots n, X_j = C \quad \forall j \neq i$)

$$\begin{aligned}
 \delta \varphi & = \left(\frac{\partial \varphi}{\partial X_i} \right) \delta X_i + \left(\frac{\partial \varphi}{\partial b} \right) \delta b = 0, \\
 \left(\frac{\partial \varphi}{\partial X_i} \right) \delta X_i & = - \left(\frac{\partial \varphi}{\partial b} \right) \delta b
 \end{aligned}$$

resolviendo para $\partial \hat{b} / \partial X_i$

$$\frac{\partial \hat{b}}{\partial X_i} = - \left(\frac{\partial \varphi}{\partial X_i} \right) / \left(\frac{\partial \varphi}{\partial b} \right).$$

Y de la misma manera, para $\partial b / \partial Y_i$, se toman como constantes todas las variables excepto Y_i (*i.e.* $X_j = C \quad \forall j = 1 \dots n, Y_j = C \quad \forall j \neq i$) y se tiene que

$$\delta \varphi = \left(\frac{\partial \varphi}{\partial Y_i} \right) \delta Y_i + \left(\frac{\partial \varphi}{\partial b} \right) \delta b = 0$$

y

$$\frac{\partial \hat{b}}{\partial Y_i} = - \left(\frac{\partial \varphi}{\partial Y_i} \right) / \left(\frac{\partial \varphi}{\partial b} \right).$$

Sustituyendo en (2.18) y factorizando

$$\hat{\sigma}_b^2 = \sum \left\{ \left(\frac{\partial \varphi}{\partial X_i} \right)^2 \frac{1}{\omega_{X_i}} + \left(\frac{\partial \varphi}{\partial Y_i} \right)^2 \frac{1}{\omega_{Y_i}} + \left(\frac{\partial \varphi}{\partial X_i} \right) \left(\frac{\partial \varphi}{\partial Y_i} \right) \frac{2r_i}{\alpha_i} \right\} / \left(\frac{\partial \varphi}{\partial b} \right)^2, \quad (2.19)$$

donde hace falta obtener las derivadas parciales. Estas se desarrollan, a partir de (2.11), de la siguiente manera:

$$\frac{\partial \varphi}{\partial X_i} = b \sum_j Z_j \left(\beta_j \frac{\partial U_j}{\partial X_i} + U_j \frac{\partial \beta_j}{\partial X_i} \right) - \sum_j Z_j \left(\beta_j \frac{\partial V_j}{\partial X_i} + V_j \frac{\partial \beta_j}{\partial X_i} \right) \quad (2.20)$$

donde

$$\frac{\partial U_j}{\partial X_i} = \delta_i^j - \frac{Z_i}{\sum_j Z_j}, \quad \frac{\partial V_j}{\partial X_i} = 0$$

y δ_i^j es la delta de Kroneker. Para desarrollar $\partial \beta_j / \partial X_i$, se puede reexpresar β_i en función de dos nuevas variables τ_j y ϵ_j que se definen de la siguiente manera

$$\beta_j = Z_j \left(\frac{U_j}{\tau_j} + \frac{V_j}{\epsilon_j} \right), \quad \tau_j = \frac{1}{\frac{1}{\omega_{Y_j}} - \frac{b r_j}{\alpha_j}} \quad \text{y} \quad \epsilon_j = \frac{1}{\frac{b}{\omega_{X_j}} - \frac{r_j}{\alpha_j}}.$$

De modo que

$$\frac{\partial \beta_j}{\partial X_i} = \frac{Z_j}{\tau_j} \left(\delta_i^j - \frac{Z_i}{\sum_j Z_j} \right).$$

Sustituyendo las derivadas parciales en (2.20)

$$\begin{aligned} \frac{\partial \varphi}{\partial X_i} &= b \sum_j Z_j \left\{ \beta_j \left(\delta_i^j - \frac{Z_i}{\sum_j Z_j} \right) + U_j \frac{Z_j}{\tau_j} \left(\delta_i^j - \frac{Z_i}{\sum_j Z_j} \right) \right\} - \sum_j \frac{Z_j^2 V_j}{\tau_j} \left(\delta_i^j - \frac{Z_i}{\sum_j Z_j} \right) \\ &= b Z_i \beta_i + \frac{Z_i^2}{\tau_i} (b U_i - V_i) - \frac{b Z_i}{\sum_j Z_j} \sum_j Z_j \beta_j - \frac{Z_i}{\sum_j Z_j} \sum_j \frac{Z_j^2}{\tau_j} (b U_j - V_j), \end{aligned} \quad (2.21)$$

la cual se puede simplificar utilizando la relación (ver apéndice B)

$$\frac{Z_j^2}{\tau_j} (b U_j - V_j) = Z_j (b \beta_j - V_j),$$

que, al sustituir en (2.21), permite que se eliminen algunos términos quedando

$$\frac{\partial \varphi}{\partial X_i} = Z_i (2b \bar{\beta}_i - V_i), \quad (2.22)$$

donde

$$\bar{\beta}_i = (\beta_i - \bar{\beta}) \quad \text{y} \quad \bar{\beta} = \frac{\sum Z_i \beta_i}{\sum Z_i}.$$

Regresando a (2.19), ahora se necesita $\partial\varphi/\partial Y_i$. Un desarrollo paralelo al que se realizó anteriormente da como resultado

$$\frac{\partial\varphi}{\partial Y_i} = Z_i (U_i - 2\bar{\beta}_i). \quad (2.23)$$

Sustituyendo (2.22) y (2.23) en (2.19), y realizando algunas simplificaciones, se encuentra que el numerador de (2.19) es

$$\begin{aligned} num &= \sum \left\{ \left(\frac{\partial\varphi}{\partial X_i} \right)^2 \frac{1}{\omega_{x_i}} + \left(\frac{\partial\varphi}{\partial Y_i} \right)^2 \frac{1}{\omega_{y_i}} + \left(\frac{\partial\varphi}{\partial X_i} \right) \left(\frac{\partial\varphi}{\partial Y_i} \right) \frac{2r_i}{\alpha_i} \right\} \\ &= 4 \sum Z_i \bar{\beta}_i (\bar{\beta}_i - \beta_i) + \sum Z_i^2 \left(\frac{U_i^2}{\omega_{y_i}} + \frac{V_i^2}{\omega_{x_i}} - 2 \frac{r_i}{\alpha_i} U_i V_i \right), \end{aligned}$$

pero $\sum Z_i \bar{\beta}_i (\bar{\beta}_i - \beta_i) = 0$ (ver apéndice B) por lo tanto

$$num = \sum_i Z_i^2 \left(\frac{U_i^2}{\omega_{y_i}} + \frac{V_i^2}{\omega_{x_i}} - 2 \frac{r_i}{\alpha_i} U_i V_i \right).$$

Al expresar el numerador así, se puede observar que corresponde con el numerador de $\hat{\sigma}_b^2$ en Minster *et al.* (1979) (tabla IV).

De las derivadas parciales que se deben de sustituir en (2.19) solo resta calcular $\partial\varphi/\partial b$. Para esto se utilizan las siguientes derivaciones (ver su desarrollo en el apéndice B)

$$\frac{\partial U_i}{\partial b} = \frac{2}{\sum_j Z_j} \sum_j \frac{Z_j^2 U_j}{\epsilon_j}, \quad (2.24)$$

$$\frac{\partial V_i}{\partial b} = \frac{2}{\sum_j Z_j} \sum_j \frac{Z_j^2 V_j}{\epsilon_j}, \quad (2.25)$$

$$\frac{\partial Z_i}{\partial b} = -2 \frac{Z_i^2}{\epsilon_i}, \quad (2.26)$$

y

$$\frac{\partial\beta_i}{\partial b} = \frac{-2Z_i\beta_i}{\epsilon_i} + Z_i \left(\frac{2}{r_i \sum_j Z_j} \sum_j \frac{Z_j^2 U_j}{\epsilon_j} + \frac{2}{\epsilon_i \sum_j Z_j} \sum_j \frac{Z_j^2 V_j}{\epsilon_j} + \frac{V_i}{\omega_{x_i}} - \frac{r_i U_i}{\alpha_i} \right). \quad (2.27)$$

A partir de (2.11) se observa que

$$\begin{aligned} \frac{\partial \varphi}{\partial b} &= \sum_i Z_i \beta_i U_i + \sum_i \beta_i \frac{\partial Z_i}{\partial b} (bU_i - V_i) + \sum_i Z_i \frac{\partial \beta_i}{\partial b} (bU_i - V_i) + \sum_i Z_i \beta_i \left(b \frac{\partial U_i}{\partial b} - \frac{\partial V_i}{\partial b} \right) \\ &= \sum_i Z_i \beta_i U_i + \sum_i \left(\beta_i \frac{\partial Z_i}{\partial b} + Z_i \frac{\partial \beta_i}{\partial b} \right) (bU_i - V_i) + \sum_i Z_i \beta_i \left(b \frac{\partial U_i}{\partial b} - \frac{\partial V_i}{\partial b} \right), \end{aligned} \quad (2.28)$$

donde, para simplificar, se definen

$$\begin{aligned} A &= \sum_i Z_i \beta_i U_i, \\ B &= \sum_i \left(\beta_i \frac{\partial Z_i}{\partial b} + Z_i \frac{\partial \beta_i}{\partial b} \right) (bU_i - V_i), \end{aligned}$$

y

$$C = \sum_i Z_i \beta_i \left(b \frac{\partial U_i}{\partial b} - \frac{\partial V_i}{\partial b} \right).$$

Sustituyendo (2.26) y (2.27) en B

$$B = \sum_i (bU_i - V_i) \left\{ -\frac{4Z_i^2 \beta_i}{\epsilon_i} + \frac{2}{\sum_j Z_j} \left(\frac{Z_i^2}{\tau_i} \sum_j \frac{Z_j^2 U_j}{\epsilon_j} + \frac{Z_i^2}{\epsilon_i} \sum_j \frac{Z_j^2 V_j}{\epsilon_j} \right) + \frac{Z_i^2 V_i}{\omega_{x_i}} - Z_i^2 \frac{r_i U_i}{\alpha_i} \right\}.$$

Desarrollando el segundo término de B

$$\sum_i (bU_i - V_i) \left\{ \frac{2}{\sum_j Z_j} \left(\frac{Z_i^2}{\tau_i} \sum_j \frac{Z_j^2 U_j}{\epsilon_j} + \frac{Z_i^2}{\epsilon_i} \sum_j \frac{Z_j^2 V_j}{\epsilon_j} \right) \right\} = 2\bar{\beta} \sum_i Z_i (U_i - \beta_i),$$

y sustituyendo de nuevo en B, queda

$$B = -4 \sum_i \frac{Z_i^2 \beta_i}{\epsilon_i} (bU_i - V_i) + 2\bar{\beta} \sum_i Z_i (U_i - \beta_i) + \sum_i Z_i^2 (bU_i - V_i) \left(\frac{V_i}{\omega_{x_i}} - \frac{r_i U_i}{\alpha_i} \right).$$

Ahora, sustituyendo (2.24) y (2.25) en C se obtiene

$$C = 2\bar{\beta} \sum_i \frac{Z_i^2}{\epsilon_i} (bU_i - V_i),$$

de esta manera, ahora es posible calcular B + C

$$\begin{aligned} B + C &= -4 \sum_i \frac{Z_i^2 \beta_i}{\epsilon_i} (bU_i - V_i) + 2\bar{\beta} \sum_i \frac{Z_i^2}{\epsilon_i} (bU_i - V_i) + 2\bar{\beta} \sum_i Z_i (U_i - \beta_i) \\ &\quad + \sum_i Z_i^2 (bU_i - V_i) \left(\frac{V_i}{\omega_{x_i}} - \frac{r_i U_i}{\alpha_i} \right). \end{aligned}$$

Utilizando la relación $\frac{Z_i^2}{\epsilon_i} (bU_i - V_i) = Z_i (U_i - \beta_i)$ (ver apéndice B) la expresión para $B + C$ se puede simplificar

$$\begin{aligned} B + C &= -4 \sum_i Z_i \beta_i (U_i - \beta_i) + 4\bar{\beta} \sum_i Z_i (U_i - \beta_i) + \sum_i Z_i^2 (bU_i - V_i) \left(\frac{V_i}{\omega_{x_i}} - \frac{r_i U_i}{\alpha_i} \right) \\ &= 4 \sum_i Z_i (\beta_i - U_i) (\beta_i - \bar{\beta}) + \sum_i Z_i^2 (bU_i - V_i) \left(\frac{V_i}{\omega_{x_i}} - \frac{r_i U_i}{\alpha_i} \right). \end{aligned}$$

Ahora, sustituyendo A y $B + C$ en (2.28) se tiene que

$$\frac{\partial \varphi}{\partial b} = \sum_i Z_i \beta_i U_i + 4 \sum_i Z_i (\beta_i - U_i) (\beta_i - \bar{\beta}) + \sum_i Z_i^2 (bU_i - V_i) \left(\frac{V_i}{\omega_{x_i}} - \frac{r_i U_i}{\alpha_i} \right).$$

En esta expresión, se pueden reordenar el primer término junto con el último término de la siguiente manera

$$\begin{aligned} &\sum_i Z_i \beta_i U_i + \sum_i Z_i^2 (bU_i - V_i) \left(\frac{V_i}{\omega_{x_i}} - \frac{r_i U_i}{\alpha_i} \right) \\ &= \frac{1}{b} \sum_i Z_i U_i V_i - \frac{1}{b} \sum_i Z_i^2 \frac{r_i}{\alpha_i} (bU_i - V_i)^2 \end{aligned}$$

para expresar $\partial \varphi / \partial b$ como

$$\begin{aligned} \frac{\partial \varphi}{\partial b} &= \frac{1}{b} \sum_i Z_i U_i V_i - \frac{1}{b} \sum_i Z_i^2 \frac{r_i}{\alpha_i} (bU_i - V_i)^2 + 4 \sum_i Z_i (\beta_i - U_i) (\beta_i - \bar{\beta}) \\ &= \sum_i Z_i \left[\frac{1}{b} U_i V_i - \frac{Z_i r_i}{b \alpha_i} (bU_i - V_i)^2 + 4 (\beta_i - U_i) (\beta_i - \bar{\beta}) \right]. \end{aligned} \quad (2.29)$$

Esta expresión corresponde con Q^{-1} en Minster *et al.* (1979) excepto el segundo término, el cual pudo ser omitido debido a un error tipográfico. Regresando a la expresión (2.19), se tiene finalmente que el estimador para el error en la pendiente esta dado por

$$\hat{\sigma}_b^2 = \frac{\sum_i Z_i^2 \left[\frac{U_i^2}{\omega_{y_i}} + \frac{V_i^2}{\omega_{x_i}} - 2 \frac{r_i U_i V_i}{\alpha_i} \right]}{\left\{ \sum_i Z_i \left[\frac{1}{b} U_i V_i - \frac{Z_i r_i}{b \alpha_i} (bU_i - V_i)^2 + 4 (\beta_i - U_i) (\beta_i - \bar{\beta}) \right] \right\}^2}. \quad (2.30)$$

Una vez corregido el estimador de σ_b^2 de Minster *et al.* (1979) por la omisión aquí señalada, éste arroja resultados numéricamente idénticos a los de York (1969).

La expresión para σ_b^2 de Mahon (1996) arroja resultados semejantes, sin embargo en su desarrollo algebraico aplica simplificaciones injustificadas, lo cual resulta contradictorio pues Mahon (1996) afirma que sus expresiones son más precisas que las publicadas en York (1969).

II.3 Error en la ordenada al origen

Utilizando de nuevo el método general de propagación de errores

$$\hat{\sigma}_a^2 = \sum_i \left\{ \left(\frac{\partial \hat{a}}{\partial X_i} \right)^2 \frac{1}{\omega_{X_i}} + \left(\frac{\partial \hat{a}}{\partial Y_i} \right)^2 \frac{1}{\omega_{Y_i}} + \frac{2r_i}{\alpha_i} \left(\frac{\partial \hat{a}}{\partial X_i} \right) \left(\frac{\partial \hat{a}}{\partial Y_i} \right) \right\} \quad (2.31)$$

donde

$$\frac{\partial \hat{a}}{\partial X_i} = \frac{\partial \bar{Y}}{\partial X_i} - \bar{X} \frac{\partial \hat{b}}{\partial X_i} - b \frac{\partial \bar{X}}{\partial X_i}, \quad (2.32)$$

$$\begin{aligned} \frac{\partial \bar{Y}}{\partial X_i} &= \frac{1}{(\sum Z_j)^2} \left[\left(\sum_j Z_j \right) \left(\sum_j Y_j \frac{\partial Z_j}{\partial X_i} \right) - \left(\sum_j Z_j Y_j \right) \left(\sum_j \frac{\partial Z_j}{\partial X_i} \right) \right] \\ &= \frac{1}{\sum Z_j} \sum_j V_j \frac{\partial Z_j}{\partial X_i}, \end{aligned}$$

$$\frac{\partial \bar{X}}{\partial X_i} = \frac{1}{\sum Z_j} \left[Z_i + \sum_j U_j \frac{\partial Z_j}{\partial X_i} \right],$$

$$\begin{aligned} \frac{\partial Z_j}{\partial X_i} &= \frac{-\omega_{x_j} \omega_{y_j} (2b\omega_{y_j} - 2r_j \alpha_j)}{(\omega_{x_j} + b^2 \omega_{y_j} - 2br_j \alpha_j)^2} \frac{\partial \hat{b}}{\partial X_i} \\ &= \frac{-2Z_j^2}{\alpha_j^2} (b\omega_{y_j} - r_j \alpha_j) \frac{\partial \hat{b}}{\partial X_i} \\ &= \frac{2Z_j^2}{\alpha_j^2} (c_j - b\omega_{y_j}) \frac{\partial \hat{b}}{\partial X_i}, \end{aligned}$$

y $c_j = \alpha_j r_j$. Sustituyendo las expresiones anteriores en (2.32)

$$\begin{aligned} \frac{\partial \hat{a}}{\partial X_i} &= \frac{-bZ_i}{\sum Z_j} + \frac{1}{\sum Z_j} \left[\sum \frac{\partial Z_j}{\partial X_i} (V_j - bU_j) \right] - \bar{X} \frac{\partial \hat{b}}{\partial X_i} \\ &= \frac{-bZ_i}{\sum Z_j} + \left[\frac{2}{\sum Z_j} \sum_j \frac{Z_j^2}{\alpha_j^2} (c_j - b\omega_{y_j}) (V_j - bU_j) - \bar{X} \right] \frac{\partial \hat{b}}{\partial X_i}, \end{aligned}$$

$$\begin{aligned}
 &= \frac{-bZ_i}{\sum Z_j} + \left[\frac{2}{\sum Z_j} \sum_j \frac{Z_j^2}{\epsilon_j} (bU_j - V_j) - \bar{X} \right] \frac{\partial \hat{b}}{\partial X_i} \\
 &= \frac{-bZ_i}{\sum Z_j} + \left[\frac{2}{\sum Z_j} \sum_j Z_j (U_j - \beta_j) - \bar{X} \right] \frac{\partial \hat{b}}{\partial X_i} \\
 &= \frac{-bZ_i}{\sum Z_j} + \left[\frac{2}{\sum Z_j} \left(\sum_j Z_j U_j - \sum_j Z_j \beta_j \right) - \bar{X} \right] \frac{\partial \hat{b}}{\partial X_i},
 \end{aligned}$$

pero $\sum Z_j U_j = 0$ (ver apéndice B), por lo tanto

$$\begin{aligned}
 \frac{\partial \hat{a}}{\partial X_i} &= \frac{-bZ_i}{\sum Z_j} + \left(\frac{-2 \sum_j Z_j \beta_j}{\sum Z_j} - \bar{X} \right) \frac{\partial b}{\partial X_i} \\
 &= \frac{-bZ_i}{\sum Z_j} - (\bar{X} + 2\bar{\beta}) \frac{\partial \hat{b}}{\partial X_i} \\
 &= \frac{-bZ_j}{\sum Z_j} + (\bar{X} + 2\bar{\beta}) \left(\frac{\partial \varphi}{\partial X_i} \right) / \left(\frac{\partial \varphi}{\partial \hat{b}} \right). \tag{2.33}
 \end{aligned}$$

En Mahon (1996), al desarrollar $\partial a / \partial X_i$ se consideró que $\partial Z_j / \partial X_i = 0$, sin embargo, como Z_i es función de b , y a su vez b es función de X_i , por lo tanto la derivación de Mahon es errónea y por esta razón su expresión para $\partial a / \partial X_i$ (ver tabla V) es diferente a la expresión (2.33). Regresando a (2.31), es necesario desarrollar

$$\frac{\partial \hat{a}}{\partial Y_i} = \frac{\partial \bar{Y}}{\partial Y_i} - \bar{X} \frac{\partial b}{\partial Y_i} - b \frac{\partial \bar{X}}{\partial Y_i} \tag{2.34}$$

donde

$$\begin{aligned}
 \frac{\partial \bar{Y}}{\partial Y_i} &= \frac{1}{\sum Z_j} \left(Z_i + \sum_j V_j \frac{\partial Z_j}{\partial Y_i} \right), \\
 \frac{\partial \bar{X}}{\partial Y_i} &= \frac{1}{\sum Z_j} \sum_j U_j \frac{\partial Z_j}{\partial Y_i},
 \end{aligned}$$

y

$$\frac{\partial Z_j}{\partial Y_i} = \frac{2Z_j^2}{\alpha_j^2} (c_j - b\omega_{Y_j}) \frac{\partial \hat{b}}{\partial Y_i},$$

donde de nuevo, Mahon (1969) utilizó erróneamente $\partial Z_j / \partial Y_i = 0$, pero sustituyendo la expresión completa en (2.34)

$$\frac{\partial \hat{a}}{\partial Y_i} = \frac{Z_i}{\sum Z_j} + \frac{1}{\sum Z_j} \left[\sum_j \frac{\partial Z_j}{\partial Y_i} (V_j - bU_j) \right] - \bar{X} \frac{\partial \hat{b}}{\partial Y_i}$$

$$\begin{aligned}
 &= \frac{Z_i}{\sum Z_j} + \left[\frac{2}{\sum Z_j} \sum_j \frac{Z_j^2}{\alpha_j^2} (c_j - b\omega_{yj}) (V_j - bU_j) - \bar{X} \right] \frac{\partial \hat{b}}{\partial Y_i} \\
 &= \frac{Z_i}{\sum Z_j} + \left[\frac{2}{\sum Z_j} \sum_j \frac{Z_j^2}{\epsilon_j} (bU_j - V_j) - \bar{X} \right] \frac{\partial \hat{b}}{\partial Y_i} \\
 &= \frac{Z_i}{\sum Z_j} + \left[\frac{2}{\sum Z_j} \sum_j Z_j (U_j - \beta_j) - \bar{X} \right] \frac{\partial \hat{b}}{\partial Y_i} \\
 &= \frac{Z_i}{\sum Z_j} - (\bar{X} + 2\bar{\beta}) \frac{\partial \hat{b}}{\partial Y_i} \\
 &= \frac{Z_i}{\sum Z_j} + (\bar{X} + 2\bar{\beta}) \left(\frac{\partial \varphi}{\partial Y_i} \right) / \left(\frac{\partial \varphi}{\partial b} \right). \tag{2.35}
 \end{aligned}$$

Antes de sustituir las expresiones (2.33) y (2.35) en (2.31), se pueden definir

$$\begin{aligned}
 A &= \frac{Z_i}{\sum Z_j}, \\
 B &= (\bar{X} + 2\bar{\beta}), \\
 C &= \frac{\partial \varphi}{\partial X_i} = Z_i (2b\bar{\beta}_i - V_i), \\
 D &= \left(\frac{\partial \varphi}{\partial b} \right)^{-1},
 \end{aligned}$$

y

$$E = \frac{\partial \varphi}{\partial Y_i} = Z_i (U_i - 2\bar{\beta}_i).$$

C y E corresponden con las expresiones (2.22) y (2.23). De esta forma se pueden reescribir (2.33) y (2.35) de la siguiente manera

$$\frac{\partial \hat{a}}{\partial X_i} = -bA + BCD \quad \text{y} \quad \frac{\partial \hat{a}}{\partial Y_i} = A + BED.$$

Los componentes de la ecuación (2.31) se pueden representar entonces como

$$\left(\frac{\partial \hat{a}}{\partial X_i} \right)^2 \frac{1}{\omega_{x_i}} = [b^2 A^2 - 2bABCD + B^2 C^2 D^2] \frac{1}{\omega_{x_i}}, \tag{2.36}$$

$$\left(\frac{\partial \hat{a}}{\partial Y_i} \right)^2 \frac{1}{\omega_{y_i}} = [A^2 + 2ABED + B^2 E^2 D^2] \frac{1}{\omega_{y_i}}, \tag{2.37}$$

y

$$\frac{2r_i}{\alpha_i} \left(\frac{\partial \hat{a}}{\partial X_i} \right) \left(\frac{\partial \hat{a}}{\partial Y_i} \right) = [-bA^2 + ABCD - bABED + B^2CED^2] \frac{2r_i}{\alpha_i}. \quad (2.38)$$

De modo que la expresión (2.31) se puede representar

$$\begin{aligned} \hat{\sigma}_a^2 &= \sum_i \left\{ \frac{b^2 A^2}{\omega_{x_i}} - \frac{2bABCD}{\omega_{x_i}} + \frac{B^2 C^2 D^2}{\omega_{x_i}} + \frac{A^2}{\omega_{y_i}} + \frac{2ABED}{\omega_{y_i}} + \frac{B^2 E^2 D^2}{\omega_{y_i}} \right. \\ &\quad \left. - \frac{2r_i b A^2}{\alpha_i} + \frac{2r_i ABCD}{\alpha_i} - \frac{2r_i b ABED}{\alpha_i} + \frac{2r_i B^2 CED^2}{\alpha_i} \right\} \\ &= \sum_i \left\{ A^2 \left(\frac{b^2}{\omega_{x_i}} + \frac{1}{\omega_{y_i}} - \frac{2r_i b}{\alpha_i} \right) - 2ABD \left(\frac{bC}{\omega_{x_i}} - \frac{E}{\omega_{y_i}} + (bE - C) \frac{r_i}{\alpha_i} \right) + \right. \\ &\quad \left. B^2 D^2 \left(\frac{C^2}{\omega_{x_i}} + \frac{E^2}{\omega_{y_i}} + \frac{2r_i CE}{\alpha_i} \right) \right\}. \end{aligned} \quad (2.39)$$

Definiendo los términos

$$F = \frac{bC}{\omega_{x_i}} - \frac{E}{\omega_{y_i}} + (bE - C) \frac{r_i}{\alpha_i} \quad \text{y} \quad H = \frac{C^2}{\omega_{x_i}} + \frac{E^2}{\omega_{y_i}} + \frac{2r_i CE}{\alpha_i}$$

y sustituyendo C y E en F y H se tiene que

$$\begin{aligned} F &= \frac{2b^2 Z_i \bar{\beta}_i}{\omega_{x_i}} - \frac{bZ_i V_i}{\omega_{x_i}} - \frac{Z_i U_i}{\omega_{y_i}} + \frac{2Z_i \bar{\beta}_i}{\omega_{y_i}} + \frac{br_i Z_i U_i}{\alpha_i} - \frac{4br_i Z_i \bar{\beta}_i}{\alpha_i} + \frac{r_i Z_i V_i}{\alpha_i} \\ &= 2Z_i \bar{\beta}_i \left(\frac{b^2}{\omega_{x_i}} + \frac{1}{\omega_{y_i}} - \frac{2br_i}{\alpha_i} \right) - Z_i \left(\frac{U_i}{\omega_{y_i}} + \frac{bV_i}{\omega_{x_i}} - (bU_i + V_i) \frac{r_i}{\alpha_i} \right) \\ &= 2\bar{\beta}_i - \beta_i, \\ H &= \frac{4b^2 Z_i^2 \bar{\beta}_i^2}{\omega_{x_i}} - \frac{4bZ_i^2 \bar{\beta}_i V_i}{\omega_{x_i}} + \frac{Z_i^2 V_i^2}{\omega_{x_i}} + \frac{Z_i^2 U_i^2}{\omega_{y_i}} - \frac{4Z_i^2 \bar{\beta}_i U_i}{\omega_{y_i}} + \frac{4Z_i^2 \bar{\beta}_i^2}{\omega_{y_i}} \\ &\quad + \frac{4br_i Z_i^2 U_i \bar{\beta}_i}{\alpha_i} - \frac{2r_i Z_i^2 U_i V_i}{\alpha_i} - \frac{8br_i Z_i^2 \bar{\beta}_i^2}{\alpha_i} + \frac{4r_i Z_i^2 \bar{\beta}_i V_i}{\alpha_i} \\ &= 4Z_i^2 \bar{\beta}_i^2 \left(\frac{b^2}{\omega_{x_i}} + \frac{1}{\omega_{y_i}} - \frac{2br_i}{\alpha_i} \right) + Z_i^2 \left(\frac{V_i^2}{\omega_{x_i}} + \frac{U_i^2}{\omega_{y_i}} - \frac{2r_i U_i V_i}{\alpha_i} \right) \\ &\quad - 4\bar{\beta}_i Z_i^2 \left(\frac{U_i}{\omega_{y_i}} + \frac{bV_i}{\omega_{x_i}} - (bU_i + V_i) \frac{r_i}{\alpha_i} \right) \\ &= 4Z_i \bar{\beta}_i^2 - 4Z_i \bar{\beta}_i \beta_i + Z_i^2 \left(\frac{V_i^2}{\omega_{x_i}} + \frac{U_i^2}{\omega_{y_i}} - \frac{2r_i U_i V_i}{\alpha_i} \right) \\ &= 4Z_i \bar{\beta}_i (\bar{\beta}_i - \beta_i) + Z_i^2 \left(\frac{V_i^2}{\omega_{x_i}} + \frac{U_i^2}{\omega_{y_i}} - \frac{2r_i U_i V_i}{\alpha_i} \right). \end{aligned}$$

Al sustituir en (2.39) se tiene que

$$\begin{aligned}
 \hat{\sigma}_a^2 &= \sum_i \left\{ \frac{A^2}{Z_i} - 2ABD (2\bar{\beta}_i - \beta_i) + 4B^2 D^2 Z_i \bar{\beta}_i (\bar{\beta}_i - \beta_i) \right. \\
 &\quad \left. + B^2 D^2 Z_i^2 \left(\frac{V_i^2}{\omega_{x_i}} + \frac{U_i^2}{\omega_{y_i}} - \frac{2r_i U_i V_i}{\alpha_i} \right) \right\} \\
 &= \sum_i \left\{ \frac{Z_i}{(\sum Z_j)^2} + \frac{2Z_i (\bar{X} + 2\bar{\beta}) (\beta_i - 2\bar{\beta}_i)}{(\sum Z_j) \left(\frac{\partial \varphi}{\partial b} \right)} + \right. \\
 &\quad \left. \frac{4(\bar{X} + 2\bar{\beta})^2 Z_i \bar{\beta}_i (\bar{\beta}_i - \beta_i)}{\left(\frac{\partial \varphi}{\partial b} \right)^2} + (\bar{X} + 2\bar{\beta})^2 \frac{Z_i^2}{\left(\frac{\partial \varphi}{\partial b} \right)^2} \left(\frac{V_i^2}{\omega_{x_i}} + \frac{U_i^2}{\omega_{y_i}} - \frac{2r_i U_i V_i}{\alpha_i} \right) \right\} \\
 &= \frac{\sum Z_i}{(\sum Z_j)^2} + 2(\bar{X} + 2\bar{\beta}) \left(\frac{\partial \varphi}{\partial b} \right)^{-1} \frac{\sum_i Z_i (\beta_i - 2\bar{\beta}_i)}{\sum Z_j} + \\
 &\quad \frac{4(\bar{X} + 2\bar{\beta})^2}{\left(\frac{\partial \varphi}{\partial b} \right)^2} \sum_i Z_i \bar{\beta}_i (\bar{\beta}_i - \beta_i) + (\bar{X} + 2\bar{\beta})^2 \sigma_b^2,
 \end{aligned}$$

pero $\sum Z_i \bar{\beta}_i (\bar{\beta}_i - \beta_i) = 0$ y $\sum Z_i (\beta_i - 2\bar{\beta}_i) / (\sum Z_j) = \bar{\beta}$ (ver apéndice B), entonces se obtiene finalmente

$$\hat{\sigma}_a^2 = \frac{1}{\sum Z_i} + (\bar{X} + 2\bar{\beta})^2 \sigma_b^2 + 2\bar{\beta} (\bar{X} + 2\bar{\beta}) \left(\frac{\partial \varphi}{\partial b} \right)^{-1}, \quad (2.40)$$

donde σ_b^2 y $\partial \varphi / \partial b$ se calculan usando las expresiones (2.30) y (2.29) respectivamente.

Esta expresión concuerda con la publicada por Minster *et al.* (1979), y es equivalente a la publicada en el apéndice de York (1969), las diferencias numéricas en los resultados se deben completamente al término faltante en la expresión para $\partial \varphi / \partial b$ (Q^{-1} en Minster *et al.*, 1979). La expresión publicada en Mahon (1996) para σ_a^2 difiere algebraica y numéricamente debido a los errores al derivar $\partial a / \partial X_i$ y $\partial a / \partial Y_i$ señalados en este capítulo.

Del desarrollo presentado se concluye que, una vez corregidos los errores tipográficos y las omisiones, las soluciones mínimo cuadráticas citadas convergen algebraicamente con la solución original de York (1969). En la tabla VI se presentan en resumen las fórmulas para los parámetros de la recta y sus errores. Estas expresiones, como aquí se

demuestra, son algebraicamente equivalentes a las de York (1969), solo difieren en su presentación más simple y fácil de programar.

Tabla VI: Resumen de fórmulas para los parámetros de la recta y sus errores.

$$\begin{aligned} \hat{a} &= \bar{Y} - b\bar{X} \\ \hat{b} &= \frac{\sum Z_i \beta_i V_i}{\sum Z_i \beta_i U_i} \\ \hat{\sigma}_a^2 &= \frac{1}{\sum Z_i} + (\bar{X} + 2\bar{\beta})^2 \hat{\sigma}_b^2 + 2\bar{\beta} (\bar{X} + 2\bar{\beta}) / \left(\frac{\partial \varphi}{\partial b} \right) \\ \hat{\sigma}_b^2 &= \left\{ \sum_i Z_i^2 \left[\frac{U_i^2}{\omega_{Y_i}} + \frac{V_i^2}{\omega_{X_i}} - 2 \frac{r_i}{\alpha_i} U_i V_i \right] \right\} / \left(\frac{\partial \varphi}{\partial b} \right)^2 \\ \frac{\partial \varphi}{\partial b} &= \sum_i Z_i \left[\frac{1}{b} U_i V_i - \frac{Z_i r_i}{b \alpha_i} (b U_i - V_i)^2 + 4 (\beta_i - U_i) (\beta_i - \bar{\beta}) \right] \end{aligned}$$

Capítulo III

Regresión lineal por el método de máxima verosimilitud

Se ha considerado generalmente el problema de ajustar una línea recta a un grupo de datos como un problema de mínimos cuadrados (MC), especialmente en la literatura geológica, pero desde el punto de vista estadístico existe otro método desde el cual también se puede abordar este problema, el método de máxima verosimilitud (MV).

Para el caso de la regresión lineal en la que las dos variables están sujetas a errores y estos están correlacionados, el método de MV arroja expresiones idénticas a las de MC para la pendiente y ordenada al origen, suponiendo que los datos tienen errores con distribuciones normales y que cada dato (X_i, Y_i) es independiente (Titterington y Halliday, 1979); no sucede lo mismo con las expresiones para los errores, las cuales no solo son diferentes algebraicamente (comparar la tabla VII con las expresiones correspondientes de la tabla VI) sino que también numéricamente los resultados pueden diferir.

Las diferencias numéricas en los errores calculados con ambos métodos por lo general son pequeñas, pero se pueden incrementar en algunos casos dependiendo de qué tan bien definen los datos a la recta: si los datos son dispersos los resultados divergen, mientras que con datos perfectamente colineales los resultados son idénticos.

El objetivo de este capítulo es explicar la relación que existe entre estos dos métodos

en el caso particular de la regresión lineal, con especial interés en las expresiones para calcular los errores de los parámetros de la recta.

Tabla VII: Fórmulas de MV para los errores en los parámetros de la recta (Titterington y Halliday, 1979). La pendiente y ordenada al origen se calculan con las fórmulas de la tabla VI.

$$\tilde{\sigma}_b^2 = \frac{\sum Z_i}{\sum Z_i \sum Z_i x_i^2 - (\sum Z_i x_i)^2}$$

$$\tilde{\sigma}_a^2 = \frac{\sum Z_i x_i^2}{\sum Z_i \sum Z_i x_i^2 - (\sum Z_i x_i)^2}$$

III.1 Parámetros de la recta

El método de MV consiste en maximizar la función de probabilidad conjunta de las observaciones L , la cual es función de los parámetros a obtener. Considerando que las observaciones son extraídas de procesos independientes y que los errores de las observaciones tienen una distribución Gaussiana, L se obtiene multiplicando las probabilidades de cada dato:

$$L = \prod_{i=1}^n \frac{1}{C_i \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \theta_i^T C_i^{-1} \theta_i \right\}, \quad (3.1)$$

donde C_i es la matriz de covarianza y θ_i representa la diferencia entre el dato observado (X_i, Y_i) y el "real" (x_i, y_i) :

$$C_i = \begin{pmatrix} \sigma_{X_i}^2 & \sigma_{XY_i} \\ \sigma_{XY_i} & \sigma_{Y_i}^2 \end{pmatrix}, \quad \theta_i = \begin{pmatrix} x_i - X_i \\ y_i - Y_i \end{pmatrix}.$$

Como no se conocen los datos reales, entonces se utilizan estimadores, por ejemplo, los obtenidos de la siguiente manera (York, 1969; Titterington y Halliday, 1979):

$$x_i = X_i + \frac{Z_i(a + bX_i - Y_i)(r_i \alpha_i - b \omega_{Y_i})}{\omega_{X_i} \omega_{Y_i}}, \quad y_i = a + bx_i.$$

Al invertir la matriz C_i se tiene que

$$\begin{aligned} C_i^{-1} &= \frac{1}{\sigma_{x_i}^2 \sigma_{y_i}^2 - \sigma_{xy_i}^2} \begin{pmatrix} \sigma_{y_i}^2 & -\sigma_{xy_i} \\ -\sigma_{xy_i} & \sigma_{x_i}^2 \end{pmatrix} \\ &= \frac{1}{1 - r_i^2} \begin{pmatrix} \omega_{x_i} & -r_i \alpha_i \\ -r_i \alpha_i & \omega_{y_i} \end{pmatrix} \end{aligned}$$

donde ω_{x_i} , ω_{y_i} y α_i están definidos en (2.3). Dado que los parámetros de la recta que maximizan a L también maximizan a su logaritmo, se puede trabajar convenientemente con $\ln L$

$$\begin{aligned} \ln L &= -\ln \left(\prod_{i=1}^n C_i \sqrt{2\pi} \right) - \frac{1}{2} \sum_{i=1}^n (\theta_i^T C_i^{-1} \theta_i) \\ &= -\ln \left(\prod_{i=1}^n C_i \sqrt{2\pi} \right) - \frac{1}{2} S. \end{aligned}$$

El primer término es independiente de a y b , por lo que se toma como constante. Desarrollando la sumatoria del segundo término

$$\begin{aligned} S &= \sum \frac{1}{1 - r_i^2} \begin{pmatrix} x_i - X_i & y_i - Y_i \end{pmatrix} \begin{pmatrix} \omega_{x_i} & -r_i \alpha_i \\ -r_i \alpha_i & \omega_{y_i} \end{pmatrix} \begin{pmatrix} x_i - X_i \\ y_i - Y_i \end{pmatrix} \quad (3.2) \\ &= \sum \frac{\omega_{x_i} (x_i - X_i)^2 + \omega_{y_i} (y_i - Y_i)^2 - 2r_i \alpha_i (x_i - X_i)(y_i - Y_i)}{1 - r_i^2}. \end{aligned}$$

Equivalentemente, se puede utilizar S de la expresión (2.4) (York, 1969; Titterington y Halliday, 1979). Los parámetros de la recta se obtienen minimizando S , que a su vez maximizan a L , por esta razón coinciden los métodos de MC y MV en los parámetros de la recta.

En general, la relación entre los estimadores de MC y MV ha sido señalada, e.g. Press *et al.* (1992) y Bevington (1969). También en el caso particular de la regresión lineal con errores correlacionados, se ha demostrado que con ambos métodos se obtienen las mismas ecuaciones para la pendiente y la ordenada al origen (Titterington y Halliday, 1979).

III.2 Error en los parámetros de la recta

La estimación de los errores en los parámetros con el método de MV se calcula utilizando la desigualdad de Cramér-Rao (Kendall y Stuart, 1961), la cual está relacionada con el límite inferior de la varianza de un estimador de MV y está dada por

$$\sigma_{\theta}^2 \geq \frac{1}{E\left[\left(\frac{\partial \log L}{\partial \theta}\right)^2\right]} = \frac{-1}{E\left[\frac{\partial^2 \log L}{\partial \theta^2}\right]}, \quad (3.3)$$

donde E indica el valor esperado. La expresión (3.3) se puede escribir en forma matricial de la siguiente manera:

$$\sigma_{\theta}^2 \geq F^{-1}, \quad (3.4)$$

donde F es la matriz de información de Fisher y sus elementos están dados por

$$F_{ij} = E\left[\frac{\partial^2 \log L}{\partial \theta_i \partial \theta_j}\right].$$

Siguiendo el desarrollo presentado en el apéndice de Titterington y Halliday (1979) para el caso de la regresión lineal con errores correlacionados, se tiene que la matriz F se representa como:

$$F = \begin{bmatrix} \sum Z_i & \sum Z_i x_i \\ \sum Z_i x_i & \sum Z_i x_i^2 \end{bmatrix}. \quad (3.5)$$

Aplicando la desigualdad (3.4), para los errores en los estimadores \hat{a} y \hat{b} se obtienen

$$\tilde{\sigma}_b^2 \geq \frac{\sum Z_i}{\sum Z_i \sum Z_i x_i^2 - (\sum Z_i x_i)^2} \quad (3.6)$$

y

$$\tilde{\sigma}_a^2 \geq \frac{\sum Z_i x_i^2}{\sum Z_i \sum Z_i x_i^2 - (\sum Z_i x_i)^2}. \quad (3.7)$$

La condición necesaria para que se cumplan estas desigualdades es que los estimadores de a y b no estén sesgados (Freund y Walpole, 1980; Varadhan, 1974), y la condición necesaria para que la igualdad se cumpla es que los estimadores sean conjuntamente suficientes (Eadie *et al.*, 1971).

Estas expresiones son algebraicamente diferentes de las obtenidas por MC (tabla VI), pero los resultados numéricos difieren muy poco cuando se cumple la hipótesis de linealidad.

III.3 Relación entre los errores de MC y MV

Acerca de las diferencias señaladas anteriormente entre las expresiones para los errores de MC y MV, se comenta en Titterington y Halliday (1979) que ambos métodos son aproximaciones para calcular los errores y que por lo general las diferencias numéricas no son significativas. También se comenta la ventaja que tienen las expresiones de MV de ser más sencillas (pero esto sin tomar en cuenta que es necesario primero calcular los datos ajustados (x_i, y_i) , es decir, los datos que cumplen con la recta para cada dato). En esta sección se presenta la demostración de que las expresiones de los errores de MV se pueden obtener exactamente de las expresiones de MC.

Normalmente se evalúan las expresiones de MC en los datos observados (X_i, Y_i) , pero evaluándolas en los datos ajustados (x_i, y_i) se tiene que el numerador de la ecuación (2.30) se convierte en

$$\sum Z_i^2 \left(\frac{u_i^2}{\omega_{Y_i}} + \frac{v_i^2}{\omega_{X_i}} - \frac{2r_i u_i v_i}{\alpha_i} \right),$$

donde u_i y v_i se calculan sobre los datos ajustados

$$u_i = x_i - \bar{x}, \quad \text{y} \quad v_i = y_i - \bar{y}.$$

Cuando se calculan de esta manera se cumple que $v_i = bu_i$ (ver apéndice B). Entonces, sustituyendo para el numerador

$$\begin{aligned} \sum Z_i^2 \left(\frac{u_i^2}{\omega_{Y_i}} + \frac{b^2 u_i^2}{\omega_{X_i}} - \frac{2br_i u_i^2}{\alpha_i} \right) &= \sum Z_i^2 u_i^2 \left(\frac{1}{\omega_{Y_i}} + \frac{b^2}{\omega_{X_i}} - \frac{2br_i}{\alpha_i} \right) \\ &= \sum Z_i^2 u_i^2 \left(\frac{1}{Z_i} \right) \\ &= \sum Z_i u_i^2. \end{aligned} \tag{3.8}$$

Ahora, evaluando $\partial\varphi/\partial b$ para los datos ajustados

$$\frac{\partial\varphi}{\partial b} = \sum Z_i \left[\frac{1}{b} u_i v_i - \frac{Z_i r_i}{b\alpha_i} (bu_i - v_i)^2 + 4(\beta_i^{ad} - u_i)(\beta_i^{ad} - \bar{\beta}^{ad}) \right], \tag{3.9}$$

donde β_i^{ad} y $\bar{\beta}^{ad}$ son evaluados en los datos ajustados:

$$\begin{aligned}
 \beta_i^{ad} &= Z_i \left(\frac{u_i}{\omega_{y_i}} + \frac{bv_i}{\omega_{x_i}} - (bu_i + v_i) \frac{r_i}{\alpha_i} \right) \\
 &= Z_i \left(\frac{u_i}{\omega_{y_i}} + \frac{b^2 u_i}{\omega_{x_i}} - (bu_i + bu_i) \frac{r_i}{\alpha_i} \right) \\
 &= Z_i u_i \left(\frac{1}{\omega_{y_i}} + \frac{b^2}{\omega_{x_i}} - \frac{2br_i}{\alpha_i} \right) \\
 &= Z_i u_i \left(\frac{1}{Z_i} \right) \\
 &= u_i,
 \end{aligned} \tag{3.10}$$

y

$$\begin{aligned}
 \bar{\beta}^{ad} &= \frac{\sum Z_i \beta_i^{ad}}{\sum Z_i} \\
 &= \frac{\sum Z_i u_i}{\sum Z_i},
 \end{aligned}$$

pero $\sum Z_i u_i = 0$ (ver apéndice B), por lo tanto $\bar{\beta}^{ad} = 0$. Sustituyendo en (3.9)

$$\begin{aligned}
 \frac{\partial \varphi}{\partial b} &= \sum Z_i \left[\frac{1}{b} u_i b u_i - \frac{Z_i r_i}{b \alpha_i} (b u_i - b u_i)^2 + 4 (u_i - u_i) u_i \right] \\
 &= \sum Z_i u_i^2.
 \end{aligned} \tag{3.11}$$

Por lo que $\hat{\sigma}_b^2$ ahora es

$$\begin{aligned}
 \hat{\sigma}_b^2 &= \frac{\sum Z_i u_i^2}{(\sum Z_i u_i^2)^2} \\
 &= \frac{1}{\sum Z_i u_i^2}.
 \end{aligned} \tag{3.12}$$

Usando las definiciones desarrolladas arriba, para $\hat{\sigma}_a^2$ se tiene

$$\begin{aligned}
 \hat{\sigma}_a^2 &= \frac{1}{\sum Z_i} + (\bar{x} + 2(0))^2 \hat{\sigma}_b^2 + 2(0) (\bar{x} + 2(0)) \left(\frac{1}{\sum Z_i u_i^2} \right) \\
 &= \frac{1}{\sum Z_i} + \bar{x}^2 \hat{\sigma}_b^2 \\
 &= \hat{\sigma}_b^2 \left(\frac{\sum Z_i u_i^2}{\sum Z_i} + \bar{x}^2 \right)
 \end{aligned}$$

$$\begin{aligned}
 &= \hat{\sigma}_b^2 \left(\frac{\sum Z_i u_i^2 + \bar{x}^2 \sum Z_i}{\sum Z_i} \right) \\
 &= \hat{\sigma}_b^2 \frac{\sum Z_i (x_i^2 - 2\bar{x}x_i + 2\bar{x}^2)}{\sum Z_i} \\
 &= \hat{\sigma}_b^2 \frac{\sum Z_i x_i^2 - 2\bar{x} (\sum Z_i x_i - \bar{x} \sum Z_i)}{\sum Z_i} \\
 &= \hat{\sigma}_b^2 \frac{\sum Z_i x_i^2 - 2\bar{x} \sum Z_i x_i}{\sum Z_i} \\
 &= \hat{\sigma}_b^2 \frac{\sum Z_i x_i^2}{\sum Z_i}. \tag{3.13}
 \end{aligned}$$

Para demostrar que las expresiones (3.12) y (3.13) son equivalentes a las de Titterington y Halliday (1979), se puede expresar $\tilde{\sigma}_b^2$ (ver tabla VII) de una forma más conveniente en función de u_i

$$\begin{aligned}
 \tilde{\sigma}_b^2 &= \frac{\sum Z_i}{\sum Z_i \sum Z_i x_i^2 - (\sum Z_i x_i)^2} \\
 &= \frac{\sum Z_i}{\sum Z_i (\sum Z_i x_i^2 - \bar{x} \sum Z_i x_i)} \\
 &= \frac{1}{\sum Z_i x_i^2 - 2\bar{x} \sum Z_i x_i + \bar{x} \sum Z_i x_i} \\
 &= \frac{1}{\sum Z_i x_i^2 - 2\bar{x} \sum Z_i x_i + \bar{x}^2 \sum Z_i} \\
 &= \frac{1}{\sum Z_i (x_i^2 - 2\bar{x}x_i + \bar{x}^2)} \\
 &= \frac{1}{\sum Z_i u_i^2}. \tag{3.14}
 \end{aligned}$$

Expresando $\tilde{\sigma}_a^2$ en función de $\tilde{\sigma}_b^2$, se tiene que

$$\begin{aligned}
 \tilde{\sigma}_a^2 &= \frac{\sum Z_i x_i^2}{\sum Z_i \sum Z_i x_i^2 - (\sum Z_i x_i)^2} \\
 &= \frac{\sum Z_i}{\sum Z_i \sum Z_i x_i^2 - (\sum x_i Z_i)^2} \frac{\sum Z_i x_i^2}{\sum Z_i} \\
 &= \tilde{\sigma}_b^2 \frac{\sum Z_i x_i^2}{\sum Z_i}. \tag{3.15}
 \end{aligned}$$

Expresadas de esta manera se puede ver claramente como al sustituir los datos ajustados en las expresiones de MC, estas coinciden con las de MV, demostrando que

las expresiones de MV son una solución especial de MC. Además, es importante notar que al sustituir los datos observados en las expresiones de MV se llega a las expresiones que York (1969) presentó como soluciones aproximadas para $\hat{\sigma}_a^2$ y $\hat{\sigma}_b^2$ (Titterington y Halliday, 1979; York *et al.*, en preparación).

Con esta demostración es posible concluir que si las expresiones de los errores de MC se evalúan en los datos ajustados, en lugar de ser evaluados en los datos observados, los resultados numéricos llegan a ser idénticos a los de MV. Este resultado explica las diferencias numéricas entre los errores de estos métodos, las cuales se deben completamente a las pequeñas diferencias que existen entre los datos observados y los datos ajustados, explicando asimismo la razón de que las diferencias numéricas se incrementen cuando los datos están alejados de la recta (York *et al.*, en preparación).

III.4 Unificación de MC y MV para casos especiales

En esta sección se demuestra que para los casos de las regresiones de Y en X y X en Y , los casos más comúnmente discutidos en la literatura, las soluciones de MC y MV convergen algebraicamente.

Para el primer caso, suponiendo que no hay errores en X y consecuentemente $r_i = 0$ y $\omega_{x_i} = \infty$, las ecuaciones de Z_i , β_i y $\partial\varphi/\partial b$ se reducen a

$$Z_i = \omega_{y_i}, \quad \beta_i = U_i, \quad y \quad \frac{\partial\varphi}{\partial b} = \frac{1}{b} \sum \omega_{y_i} U_i V_i,$$

sustituyendo ahora en $\hat{\sigma}_b^2$

$$\begin{aligned} \hat{\sigma}_b^2 &= \frac{\sum \omega_{y_i}^2 \left(\frac{U_i^2}{\omega_{y_i}} \right)}{\left(\frac{1}{b} \sum \omega_{y_i} U_i V_i \right)^2} \\ &= \frac{b^2 \sum \omega_{y_i} U_i^2}{\left(\sum \omega_{y_i} U_i V_i \right)^2}. \end{aligned} \tag{3.16}$$

Asimismo la ecuación de la pendiente (2.12) se convierte en

$$\hat{b} = \frac{\sum \omega_{y_i} U_i V_i}{\sum \omega_{y_i} U_i^2},$$

$$\hat{b} \sum \omega_{Y_i} U_i^2 = \sum \omega_{Y_i} U_i V_i,$$

sustituyendo en el denominador de (3.16)

$$\begin{aligned} \hat{\sigma}_b^2 &= \frac{b^2 \sum \omega_{Y_i} U_i^2}{(b \sum \omega_{Y_i} U_i^2)^2} \\ &= \frac{1}{\sum \omega_{Y_i} U_i^2}. \end{aligned} \quad (3.17)$$

Para $\hat{\sigma}_a^2$ en el caso que $\sigma_{X_i} = 0$, se tiene de la ecuación (2.40)

$$\hat{\sigma}_a^2 = \frac{1}{\sum Z_i} + (\bar{X} + 2\bar{\beta})^2 \hat{\sigma}_b^2 + 2\bar{\beta} (\bar{X} + 2\bar{\beta}) \left(\frac{\partial \varphi}{\partial b} \right)^{-1}, \quad (3.18)$$

pero se ha demostrado que $\bar{\beta} = 0$ cuando $\beta_i = U_i$, por lo tanto

$$\hat{\sigma}_a^2 = \frac{1}{\sum \omega_{Y_i}} + \bar{X}^2 \hat{\sigma}_b^2.$$

Siguiendo el desarrollo que se utilizó para obtener (3.13) se puede reexpresar esta ecuación como

$$\begin{aligned} \hat{\sigma}_a^2 &= \hat{\sigma}_b^2 \frac{\sum Z_i X_i^2}{\sum Z_i} \\ &= \hat{\sigma}_b^2 \frac{\sum \omega_{Y_i} X_i^2}{\sum \omega_{Y_i}}. \end{aligned} \quad (3.19)$$

Sustituyendo $Z_i = \omega_{Y_i}$ en las expresiones (3.14) y (3.15) de MV, se tiene

$$\hat{\sigma}_b^2 = \frac{1}{\sum \omega_{Y_i} u_i^2}, \quad (3.20)$$

y

$$\hat{\sigma}_a^2 = \hat{\sigma}_b^2 \frac{\sum \omega_{Y_i} x_i^2}{\sum \omega_{Y_i}}. \quad (3.21)$$

Se observa que $X_i = x_i$ y $U_i = u_i$ cuando no existen errores en X , por lo tanto las expresiones (3.16) y (3.19) son idénticas a (3.20) y (3.21), demostrando así la convergencia de los métodos de MC y MV en el caso particular de la regresión Y en X . Esto mismo se cumple para el caso de la regresión de X en Y .

Dado que los métodos de MC y MV coinciden perfectamente en los casos más comunmente discutidos en la literatura, que son las regresiones de Y en X y X en Y , no es necesario seleccionar uno u otro método. Sin embargo, en el caso general, cuando se tienen errores en las dos variables y estos pueden estar correlacionados, es importante señalar que los errores calculados con el método de MC se calculan en base a los datos observados, mientras que el método de MV los calcula a partir de los datos ajustados, por lo que su valor numérico puede ser diferente.

Capítulo IV

Pruebas numéricas

En los capítulos anteriores se ha demostrado que hay diferencias entre las fórmulas para calcular los errores en los parámetros de la recta. El número de formas diferentes para calcular los errores, sin embargo, es menor del que aparentemente existe en la literatura, ya que, como se demostró en el capítulo 2, todas las soluciones que usan el método de mínimos cuadrados (MC) son equivalentes, por lo que para el caso general de la regresión lineal, cuando ambas variables están sujetas a errores y estos pueden estar correlacionados, solo hay dos alternativas para calcular los errores en los parámetros de la recta: las fórmulas de MC y las de máxima verosimilitud (MV).

Como no es posible descartar ninguna de estas dos soluciones, ya que ambas están respaldadas por una base teórica, se realizaron algunas pruebas numéricas para ver cuál de los dos métodos es más consistente y cuál corresponde mejor con la intuición. A continuación se describen las pruebas numéricas realizadas y se presentan los resultados obtenidos.

IV.1 Convergencia con datos ajustados

Se ha demostrado anteriormente (ver sección III.3) que, al introducir los datos ajustados en las expresiones de MC, éstas son idénticas a las de MV. Para respaldar esta

Tabla VIII: Comparación de los errores calculados por las fórmulas de MC, las fórmulas de MV y los calculados con MC utilizando los datos ajustados.

Datos	MC		MV		MC con datos ajustados	
	$\hat{\sigma}_a$	$\hat{\sigma}_b$	$\tilde{\sigma}_a$	$\tilde{\sigma}_b$	$\hat{\sigma}_a$	$\hat{\sigma}_b$
York-66	0.291934	0.0576167	0.294971	0.0579850	0.294971	0.0579850
Mahon-1	0.000349769	0.00952059	0.000320288	0.00871306	0.000320288	0.00871306
Mahon-2	0.000375911	0.0102713	0.000364644	0.00996480	0.000364644	0.00996480
Reed-i-1	0.0324699	0.0354464	0.0280406	0.0305674	0.0280406	0.0305674
Reed-i-2	55.1930	10.9604	26.9173	5.33265	26.9173	5.33265

Tabla IX: Referencias bibliográficas de los datos utilizados e información sobre las rectas calculadas.

Datos	Referencia	n	S	$MSWD$	Corr	r	a	b
York-66	York (1966)	10	11.866	1.483	-0.9765	0	5.47991	-0.480533
Mahon-1	Mahon (1996)	8	10.382	1.730	-0.9740	0	0.00364111	-0.0763849
Mahon-2	Mahon (1996)	8	6.266	1.044	-0.9740	0.7071	0.00368416	-0.0775157
Reed-i-1	Reed (1989)	27	578.047	23.122	0.8202	0	-0.365155	1.16684
Reed-i-2	Reed (1989)	10	13.956	1.744	-0.1708	0	-17.4835	4.54366

demostración se muestran en la tabla VIII los resultados numéricos de 6 grupos de datos donde se hicieron las sustituciones de datos ajustados en las expresiones de MC.

Los datos utilizados son casos de prueba típicos tomados de la literatura, en el apéndice C se muestran las gráficas. En la tabla IX se presentan las referencias bibliográficas e información relacionada a las rectas que ajustan a los datos: número de datos, sumatoria de residuales, correlación lineal, correlación de los errores (para todos estos casos se aplicó la misma correlación en los errores, en cada conjunto de datos), ordenada al origen y pendiente.

En la tabla VIII se puede observar cómo los resultados de las últimas dos columnas, calculados con las expresiones de MC pero utilizando los datos ajustados, coinciden perfectamente con los resultados de MV. También se puede observar que los errores calculados con las expresiones de MV son inferiores en la mayoría de los casos mostrados a los de MC. Esto es debido a que los errores calculados por MV son por definición el límite inferior de la incertidumbre.

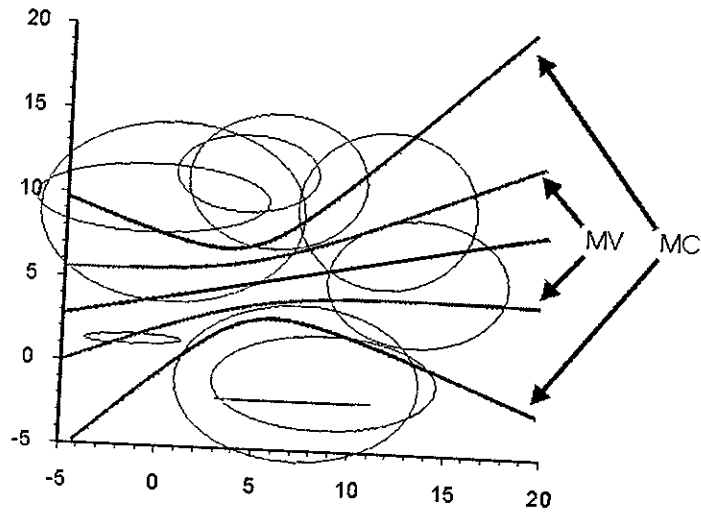
De los ejemplos mostrados, la mayor diferencia numérica entre los resultados de MC y MV se produce con los datos Reed-i-2, los cuales tienen la menor correlación lineal (ver fig. 1). En la sección IV.3 se presenta un intento de tipificar los datos que cumplen con la característica de producir resultados divergentes al calcular los errores con las fórmulas de MC y MV.

Los resultados numéricos de esta prueba confirman que la diferencia entre las expresiones de MC y MV para las incertidumbres en los parámetros de la recta se deben a que el método de MC propaga los errores de los datos observados y MV propaga los errores de los datos ajustados, como se demostró algebraicamente en el capítulo anterior.

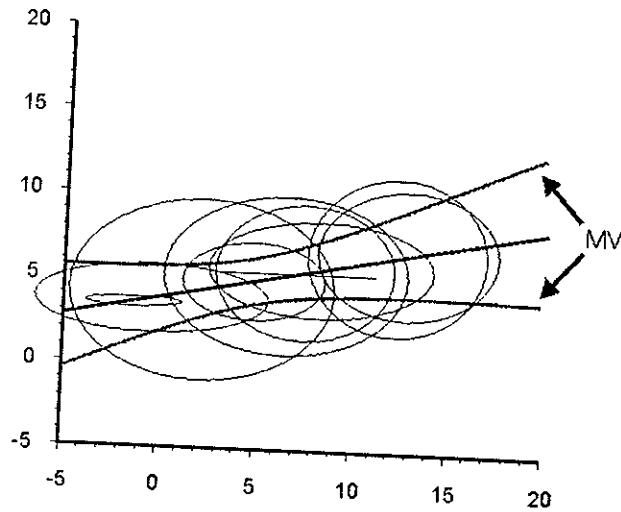
IV.2 Convergencia con pesos especiales

En la sección III.4 se demostró que los métodos de MC y MV coinciden algebraicamente para los casos de la regresión de Y en X y X en Y , es decir, cuando el error de una de

Reed-i-2



Datos observados



Datos ajustados

Figura 1: Datos observados de Reed-i-2 y datos ajustados. En la gráfica se han invertido los ejes debido a una limitación del programa utilizado para graficar los datos. Las hiperbolas representan los intervalos de confianza definidos por las incertidumbres calculadas por los métodos de MC y MV.

Tabla X: Comparación de los errores calculados por las fórmulas de MC y las fórmulas de MV tomando en cuenta únicamente el error σ_{y_i} (regresión de Y en X).

Datos	Recta		MC		MV	
	\hat{a}	\hat{b}	$\hat{\sigma}_a$	$\hat{\sigma}_b$	$\tilde{\sigma}_a$	$\tilde{\sigma}_b$
York-66	6.10011	-0.610813	0.204663	0.0300874	0.204663	0.0300874
Mahon-1	0.00341742	-0.0705731	0.000244564	0.00631974	0.000244564	0.00631974
Reed-i-1	-0.150072	0.931057	0.0171638	0.0186968	0.0171638	0.0186968
Reed-i-2	3.75850	-1.61327	0.140611	0.0675261	0.140611	0.0675261

Tabla XI: Comparación de los errores calculados por las fórmulas de MC y las fórmulas de MV tomando en cuenta únicamente el error σ_{x_i} (regresión de X en Y).

Datos	Recta		MC		MV	
	\hat{a}	\hat{b}	$\hat{\sigma}_a$	$\hat{\sigma}_b$	$\tilde{\sigma}_a$	$\tilde{\sigma}_b$
York-66	5.94505	-0.630429	0.0160165	0.00833718	0.0160165	0.00833718
Mahon-1	0.00374386	-0.0791517	5.13102e-05	0.00194319	5.13102e-05	0.00194319
Reed-i-1	-0.563135	1.38387	0.0288983	0.0315442	0.0288983	0.0315442
Reed-i-2	-34.7951	7.98903	79.7734	16.0023	79.7734	16.0023

las variables es igual a 0. Para respaldar esto numéricamente, se utilizaron los grupos de datos citados en la tabla IX, pero tomando en cuenta únicamente la incertidumbre en una de las variables e ignorando la incertidumbre en la otra variable y su correlación.

Ya que el programa calcula el peso de cada dato con el inverso de la varianza (i.e. $\omega_{x_i} = 1/\sigma_{x_i}$ y $\omega_{y_i} = 1/\sigma_{y_i}$), ninguna de las varianzas puede ser igual a 0 en el programa, por esta razón se asignó a la incertidumbre un valor de 1×10^{-10} .

Para la regresión de Y en X se tomó en cuenta únicamente σ_{y_i} y se ignoró la incertidumbre σ_{x_i} . Se puede observar en la tabla X que, hasta la precisión que se muestra, corresponden perfectamente los resultados de MC con los de MV. Lo mismo sucede cuando sólo se toma en cuenta σ_{x_i} , lo cual equivale a la regresión de X en Y , estos resultados se muestran en la tabla XI.

IV.3 Datos que producen resultados divergentes

Se ha señalado en diversas ocasiones que, en la mayoría de los casos, las expresiones de MC y MV para las incertidumbres arrojan resultados muy cercanos, pero se puede observar en la tabla VIII que en algunos casos las diferencias entre ambos métodos son importantes. A continuación se muestra que los datos que producen diferencias significativas entre ambos métodos tienen, por lo general, también otras características en común.

Los cálculos se hicieron únicamente para el error en la pendiente, pero las conclusiones a las que se llega pueden hacerse extensivas al error en la ordenada en el origen, ya que este es función del error en la pendiente (ver ecuaciones (2.40) y (3.15)).

IV.3.1 Sesgo por el método jackknife

El método jackknife es un método de remuestreo que permite conocer la variabilidad de un estimador estadístico calculando su sesgo (*Bias*), que es la diferencia que existe entre su valor esperado y su valor calculado. La idea principal del método jackknife

es tomar submuestras de los datos descartando una observación a la vez y recalculando el estimador para cada una de estas submuestras, la diferencia entre el valor estimado para la muestra de datos y el valor promedio de los calculados de las submuestras es una aproximación del sesgo del estimador (Efron, 1982). Por ejemplo, para el caso del estimador $b = b(d_1, \dots, d_n)$ (la pendiente de la recta mínimo cuadrática que pasa por un conjunto de n datos), el sesgo de b está dado por

$$B = \frac{n-1}{n} \sum (b_i - b), \quad (4.1)$$

donde b_i es la pendiente calculada sin tomar en cuenta el i -ésimo dato,

$$b_i = b(d_1, \dots, d_{i-1}, d_{i+1}, \dots, d_n),$$

y d_i es el i -ésimo dato, conformado por abcisa, ordenada, errores y su correlación,

$$d_i = (x_i, y_i, \sigma_{x_i}, \sigma_{y_i}, r_i).$$

IV.3.2 Sesgo y correlación lineal

Con el propósito de identificar los casos en los cuales las incertidumbres calculadas por los métodos de MC y MV producen resultados diferentes, se calculó la diferencia Δ_{σ_b} ,

$$\Delta_{\sigma_b} = \left| \frac{\hat{\sigma}_b - \tilde{\sigma}_b}{b} \right|, \quad (4.2)$$

y se graficó contra B (ecuación 4.1). La fig. 2 muestra la gráfica con los resultados de esta prueba realizada con 129 grupos de datos ¹. Dado que los datos tienen diferencias de varios órdenes de magnitud entre ellos, se graficaron $\log |B|$ y $\log \Delta_{\sigma_b}$ y se encontró una correlación de 0.4224, esta correlación indica que los datos sesgados tienden a producir diferencias entre los métodos de MC y MV.

También se encontró una relación entre la correlación lineal de los datos y Δ_{σ_b} . En la fig. 3 se muestran las correlaciones lineales de los mismos 129 grupos de datos y Δ_{σ_b} ,

¹Los datos fueron tomados principalmente del laboratorio de Geocronología, de la División de Ciencias de la Tierra del CICESE.

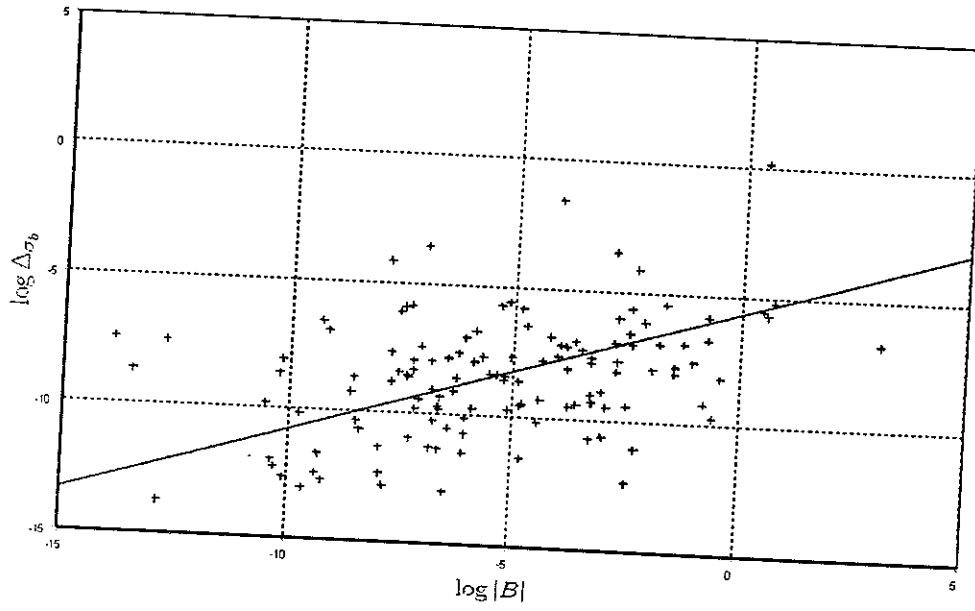


Figura 2: Gráfica de sesgo contra $\Delta\sigma_b$.

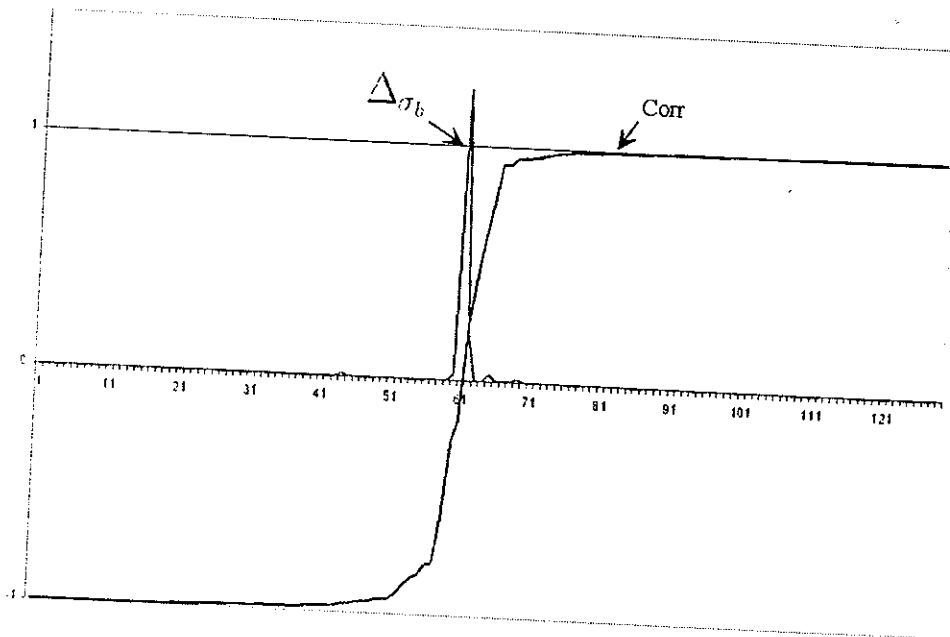


Figura 3: Correlación lineal de 129 grupos de datos.

ordenados por su correlación lineal para una mejor apreciación visual. Se puede ver claramente en la fig. 3 que cuando la correlación lineal es cercana a 0, la diferencia Δ_{σ_b} tiene valores máximos.

IV.4 Método de Monte Carlo para propagación de errores

Un algoritmo de Monte Carlo es cualquier algoritmo numérico que utiliza secuencias de números aleatorios. En esta sección se describe la adaptación del método de Monte Carlo para propagación de errores (Anderson, 1976) para la regresión lineal con errores correlacionados².

Si se conoce el error de los datos de entrada, se puede estimar el error de los parámetros calculándolos repetidas veces, variando aleatoriamente los datos de entrada dentro de su incertidumbre en cada repetición del cálculo (Anderson, 1976).

La desviación estándar de los parámetros calculados con los datos perturbados aleatoriamente es una medida del error de los parámetros. En la fig. 4(a) se muestra un diagrama que explica el funcionamiento del algoritmo de Monte Carlo para propagación de errores.

La propagación de errores por este método es computacionalmente ineficiente pero conceptualmente muy simple por lo que es un método atractivo para propagar los errores cuando es necesario verificar los resultados de fórmulas complicadas, o cuando es difícil desarrollar algebraicamente las ecuaciones de las incertidumbres (Anderson, 1976).

En el caso particular de la regresión lineal con errores correlacionados, se puede utilizar este algoritmo de Monte Carlo para verificar los resultados de las expresiones de MC. También se puede modificar el algoritmo para propagar el error de los datos ajustados, de tal manera que, con esta modificación, se pueda utilizar este método para

²Partes de esta adaptación fueron implementadas originalmente por Norman Evensen, de la Universidad de Toronto, a quien le agradezco haberme facilitado el código fuente de sus programas.

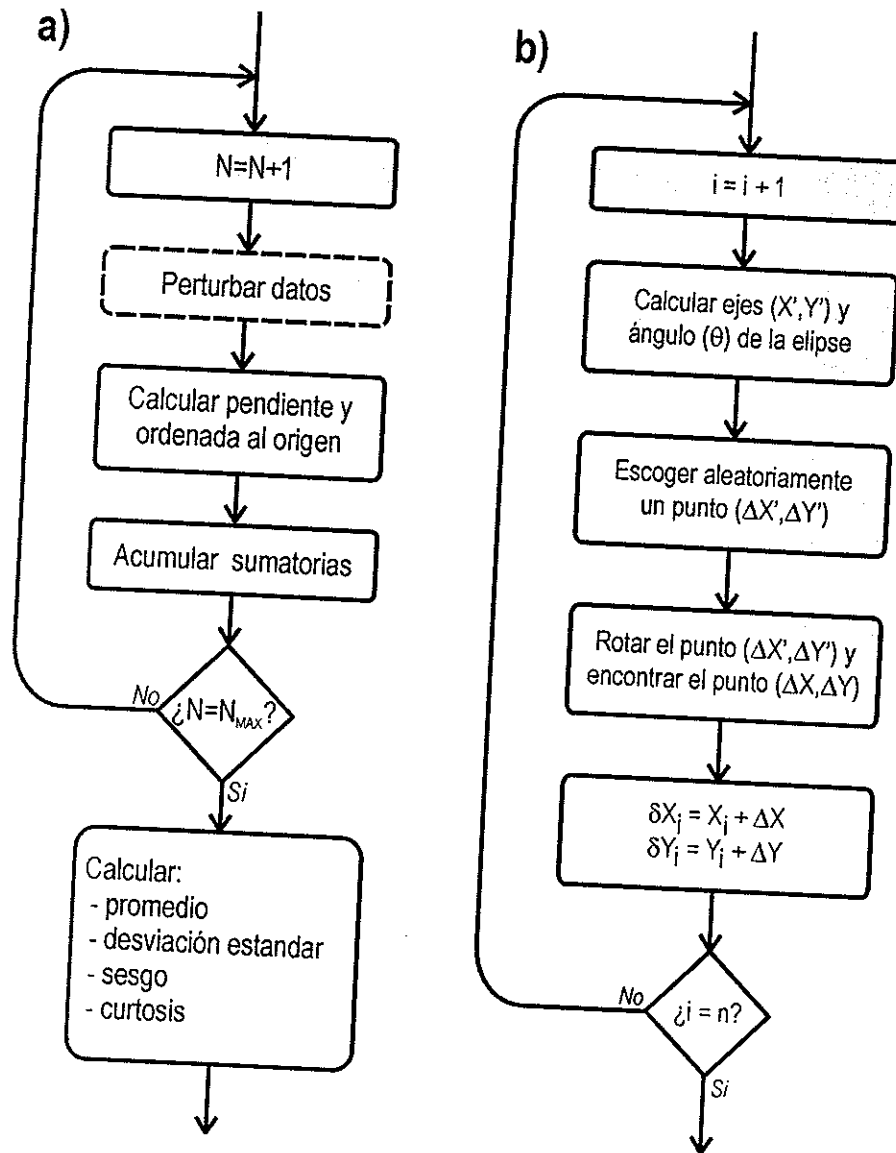


Figura 4: (a) Algoritmo de Monte Carlo para propagación de errores, N_{MAX} es el número de iteraciones. (b) Perturbación de datos del algoritmo de Monte Carlo, n es el número de datos.

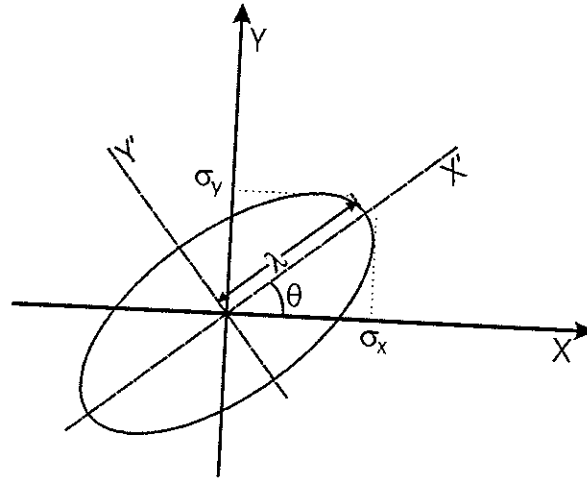


Figura 5: Parámetros de la elipse de error.

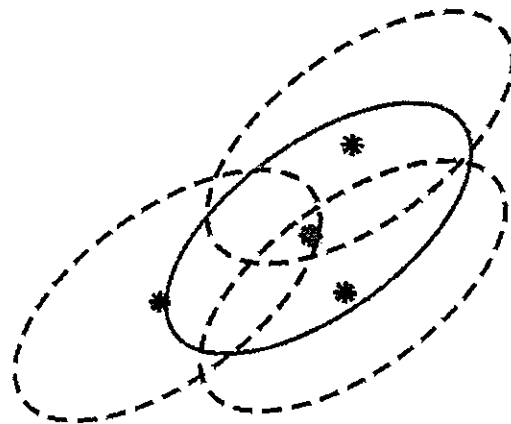


Figura 6: Elipse de incertidumbre inicial (área sombreada) y 3 valores perturbados aleatoriamente.

verificar los resultados de las expresiones de MV.

Para generar datos perturbados dentro de su incertidumbre es importante tomar en cuenta la correlación de los errores. La fig. 5 muestra la elipse de error de un dato, donde λ y θ están dados por

$$\lambda^2 = \frac{\sigma_{x_i}^2 + \sigma_{y_i}^2 + \sqrt{(\sigma_{x_i}^2 + \sigma_{y_i}^2)^2 + 4r_i\sigma_{x_i}^2\sigma_{y_i}^2}}{2} \quad \text{y} \quad \theta = \tan^{-1} \left(\frac{\lambda^2 - \sigma_{x_i}^2}{r_i\sigma_{x_i}\sigma_{y_i}} \right).$$

En el diagrama de la fig. 4(b) se muestra el algoritmo que perturba los datos, a continuación se explican los pasos que sigue el algoritmo:

1. Encontrar los ejes X' y Y' de la elipse de error y el ángulo θ que forma el eje mayor con el eje de las abcisas (ver fig. 5),
2. Utilizando los ejes de la elipse, tomar dos números $\Delta X'$ y $\Delta Y'$, donde $\Delta X'$ se toma aleatoriamente de una muestra con distribución normal $\mathcal{N}(0, \sigma_{x_i})$, y $\Delta Y'$ proviene de una muestra con distribución normal $\mathcal{N}(0, \sigma_{y_i})$,
3. Calcular la perturbación $(\Delta X, \Delta Y)$ rotando el punto $(\Delta X', \Delta Y')$ con el ángulo que se calculó en el paso 1,
4. Calcular el dato perturbado $(\delta X_i, \delta Y_i)$ con las siguientes fórmulas:

$$\delta X_i = X_i + \Delta X,$$

$$\delta Y_i = Y_i + \Delta Y.$$

Estos pasos se siguen con cada dato para generar un conjunto de datos perturbados aleatoriamente, de esta manera los datos perturbados tienen los mismos errores y correlación que los datos de entrada después de un número significativo de repeticiones.

En la fig. 6 se muestra la elipse de incertidumbre de un dato y algunos posibles valores perturbados aleatoriamente. Utilizando un número grande de variaciones aleatorias se reproduce la elipse de incertidumbre inicial.

En la tabla XII se muestran los resultados de los algoritmos de Monte Carlo que utilizan datos observados y datos ajustados, con 1,000,000 iteraciones realizadas, utilizando los datos Reed-i-1 (ver fig. C4 en el apéndice C). También se muestran en la tabla XII las incertidumbres de los parámetros calculadas con las expresiones de MC y MV. Se observa que el algoritmo que utiliza los datos observados arroja resultados cercanos a los de MC, y el algoritmo que utiliza datos ajustados, a los de MV.

Tabla XII: Comparación de los resultados de Monte Carlo con los errores calculados por las fórmulas de MC y las fórmulas de MV.

	Monte Carlo		MC	MV
	obser.	ajust.		
σ_a	0.032619	0.028154	0.032470	0.028041
σ_b	0.035608	0.030690	0.035446	0.030567

En la fig. 7 se muestran los histogramas de la pendiente y ordenada al origen utilizando los datos observados, y en la fig. 8 se muestran los histogramas del algoritmo que utiliza los datos ajustados.

Los resultados del método de Monte Carlo para propagación de errores son consistentes con las conclusiones de las secciones anteriores: las fórmulas para las incertidumbres desarrolladas por el método de MC propagan el error de los datos observados y las de MV de los datos ajustados.

Histogramas de Monte Carlo utilizando datos observados

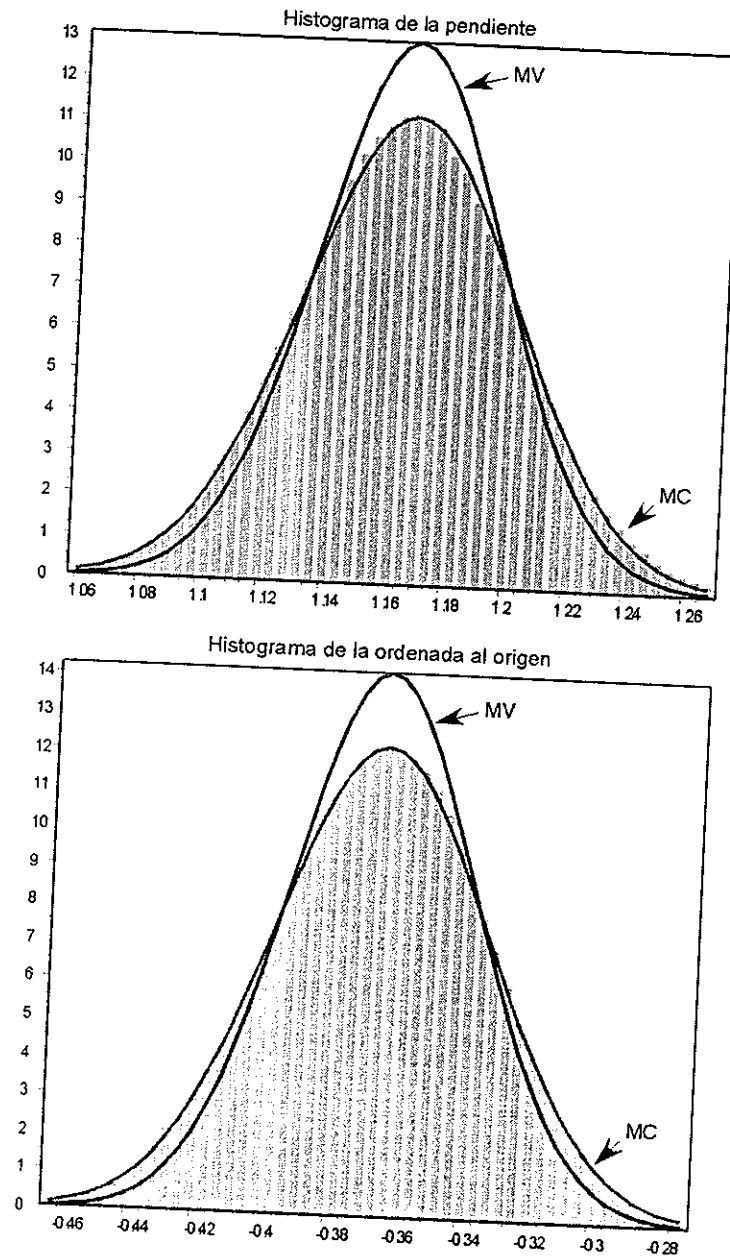


Figura 7: Histogramas del algoritmo de Monte Carlo que utiliza los datos observados. Las curvas representan la distribución normal con la desviación estándar calculada por las fórmulas de propagación de errores. MC es mínimos cuadrados y MV es máxima verosimilitud.

Histogramas de Monte Carlo utilizando datos ajustados

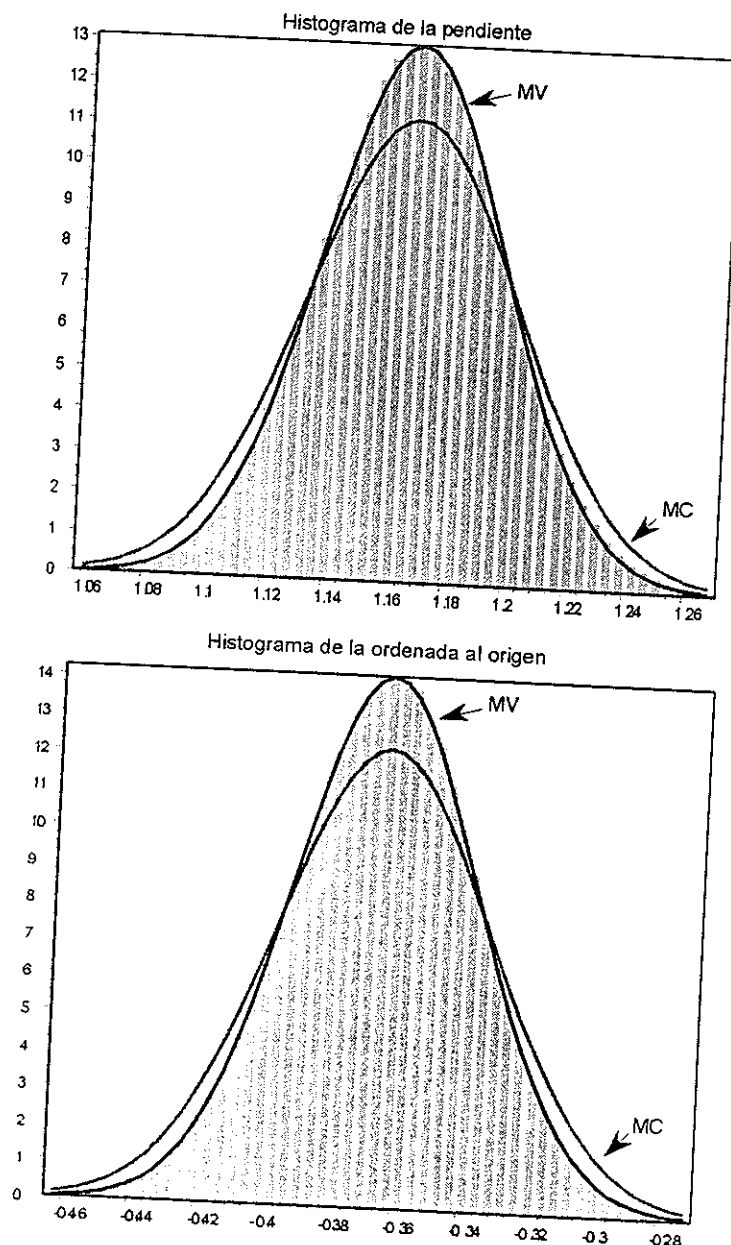


Figura 8: Histogramas del algoritmo de Monte Carlo que utiliza los datos ajustados. Las curvas representan la distribución normal con la desviación estándar calculada por las fórmulas de propagación de errores. MC es mínimos cuadrados y MV es máxima verosimilitud.

Capítulo V

Discusión

En esta tesis se han presentado pruebas algebraicas que demuestran que todas las alternativas para calcular la pendiente y la ordenada al origen de la recta mínimo cuadrática que ajusta a un grupo de datos con errores correlacionados convergen en una misma solución, la cual está respaldada por los métodos de MC y MV.

También se ha mostrado en esta tesis que existen dos maneras de calcular los errores de los parámetros de la recta: propagando el error con los datos observados o con los datos ajustados; la primera corresponde con la solución de MC y la segunda con MV.

Cuál de las dos formas de calcular los errores se debe de utilizar, es una pregunta difícil de responder, y de gran importancia, ya que, de acuerdo con el material presentado en el capítulo 4, los resultados pueden ser distintos en algunos casos, y entonces se debe de escoger un método u otro.

Para responder a esta cuestión, se deben de considerar dos casos: uno, en el que los datos tienen una baja correlación lineal, y otro, con datos altamente correlacionados.

En el primer caso se puede argumentar que es absurdo intentar calcular el error de datos que están pobremente correlacionados, pero esta no es la pregunta que se intenta responder sino cuál es el método correcto para propagar el error, y está claro que no tiene sentido propagar el error con los datos *ajustados* a la recta cuando no se tiene la certeza de que los datos *pertenezcan* a una recta, por lo que se deduce que al utilizar

los datos observados se obtiene una mejor estimación del error.

En el segundo caso, la hipótesis de linealidad está sustentada, dado que los datos están fuertemente correlacionados. En este caso es razonable pensar que se deberían de utilizar los datos ajustados, ya que el valor esperado del dato es el que cumple con la ecuación de la recta, por lo que el método de MV es el que se recomienda aquí. No obstante, se ha observado que, cuando la correlación lineal es alta, las diferencias numéricas entre MC y MV son mínimas (ver sección IV.3), de tal manera que los errores calculados con las fórmulas de MC también son una buena aproximación.

Intuitivamente, se podría decir que el método de MC tiene una aplicación más general, como se sugiere de la demostración que se presenta en la sección III.3, pero derivar de esta apreciación una afirmación que descarte cualquiera de los dos métodos está fuera del alcance de esta tesis.

Capítulo VI

Conclusiones

Con base en los resultados obtenidos en el desarrollo de esta tesis, se pueden concluir los puntos que a continuación se enumeran:

1. Las publicaciones que utilizan el método de MC para ajustar una recta a un conjunto de datos con errores correlacionados coinciden en las fórmulas para los parámetros de la recta y sus incertidumbres, si se hacen las correcciones necesarias en Minster *et al.* (1979) y Mahon (1996).
2. El método de MV coincide con el de MC en las expresiones para los parámetros de la recta, pero las expresiones para las incertidumbres de los parámetros son diferentes algebraicamente. Las diferencias numéricas son pequeñas en muchos casos, pero también existen casos en los que las diferencias son significativas.
3. Las diferencias entre las expresiones de las incertidumbres de MC y MV se debe *completamente* a que el método de MC propaga los error de los datos observados y el de MV, el de los datos ajustados.
4. Las expresiones de las incertidumbres de MV son una solución particular de MC, ya que estas se pueden obtener exactamente al sustituir los datos ajustados en las expresiones de los errores calculadas por MC.

5. Las expresiones de las incertidumbres de MC y MV coinciden perfectamente en los casos especiales de las regresiones de Y en X y X en Y .
6. Las incertidumbres de los parámetros calculadas por MV representan, en general, el error mínimo de los parámetros, y pueden ser una buena aproximación de la varianza cuando los datos tienen una alta correlación lineal y no están sesgados. Sin embargo, es importante enfatizar que, en este caso, las incertidumbres de los parámetros de la recta calculadas por ambos métodos son indistinguibles.

Bibliografía

- Acton, F.S. (1959) Analysis of straight-line data. Dover Publications, New York, 267 pp.
- Anderson, G.M. (1976) Error propagation by the Monte Carlo method in geochemical calculations. *Geochim. Cosmochim. Acta* 40: 1533-1538 p.
- Bevington, P.R. (1969) Data Reduction and error analysis for the physical sciences. McGraw-Hill, New York, 336 pp.
- Björck, A. (1996) Numerical methods for least squares problems. SIAM, Philadelphia, 408 pp.
- Brooks, C., y Compston (1965) The Age and initial $^{87}\text{Sr}/^{86}\text{Sr}$ of the Heemiskirk granite, western Tasmania. *J. Geophys. Res.* 70: 6249-6262 p.
- Brooks, C., Hart, S.R., y Wendt, I. (1972) Realistic use of two-error regression treatments as applied to rubidium-strontium data. *Rev. of Geophys. Space Phys.* 10: 551-577 p.
- Daniel, C. y Wood, F.S. (1980) Fitting equations to data. John Wiley & Sons, New York, 458 pp.
- De Basabe Delgado, J. y López Martínez, M. (1999) Cálculo de isocronas de $^{40}\text{Ar}/^{39}\text{Ar}$, $\text{Rb} - \text{Sr}$ y $\text{Sm} - \text{Nd}$ utilizando el programa Isochron. *GEOS*, 19 (4): 247 p.
-

- Dickin, A.P. (2000) Radiogenic isotope geology. Cambridge University Press, Cambridge, 490 pp.
- Eadie, W.T., Dryard, D., James, F.E., Roos, M. y Saboulet, B. (1971) Statistical methods in experimental physics. North Holland publishing company, Amsterdam, 296 pp.
- Efron, B. (1982) The Jackknife, the bootstrap and other resampling plans. Society for industrial and applied mathematics, Philadelphia, 92 pp.
- Faure, G. (1986) Principles of isotope geology, segunda edición. John Wiley & Sons, New York, 589 pp.
- Freund, J.E. y Walpole, R.E. (1980) Mathematical statistics. Prentice-Hall, tercera edición, New Jersey, 548 pp.
- Hofmann, A.W., y Hart, S.R. (1978) An assessment of local and regional isotopic equilibrium in the mantle. *Earth Planet. Sci. Lett.* 38: 44-62 p.
- Kendall, M.G. y Stuart, A. (1961) The advanced theory of statistics, Vol. 2. Hafner Publishing Company, New York, 676 pp.
- López Martínez, M., De Basabe Delgado, J. y York, D. (2000) Everything old is new again. *EOS* 81 (19): 423 p.
- Ludwig K.R. (1980) Calculation of uncertainties of U-Pb isotope data. *Earth Planet. Sci. Lett.* 46: 212-220 p.
- Macki, J.W. (1994) Total least squares: state-of-the-art regression in numerical analysis. *SIAM Review* 36 (2): 258-264 p.
- Mahon, K.I. (1996) The new "York" regression: Application of an improved statistical method to geochemistry. *Int. Geol. Rev.* 38: 293-303 p.

- McDougall, I. y Harrison, T.M. (1988) Geochronology and Thermochronology by the $^{40}\text{Ar} - ^{39}\text{Ar}$ method. Oxford University Press, New York, 212 pp.
- McIntyre G.A., Brooks, C., Compston, W. y Turek, A. (1966) The statistical assessment of $\text{Rb} - \text{Sr}$ isochrons. *J. Geophys. Res.* 71: 5459-5468 p.
- Meyer, S.L. (1975) Data analysis for scientists and engineers. John Wiley & Sons, New York, 513 pp.
- Minster J.F., Ricard, L.P. y Allgre, C.J. (1979) $^{87}\text{Rb} - ^{87}\text{Sr}$ chronology of enstatite meteorites. *Earth Planet. Sci. Lett.* 44: 420-440 p.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T. y Flannery, B.P. (1992) Numerical recipes in C, the art of scientific computing. Cambridge University Press, New York, 994 pp.
- Reed, B.C. (1989) Linear least-squares fits with errors in both coordinates. *American J. Phys.* 57 (7): 642-646 p.
- Reed, B.C. (1992) Linear least-squares fits with errors in both coordinates. II: coments on parameter variances. *American J. Phys.* 60 (1): 59-62 p.
- Roddick, J.C., y Compston W. (1977) Strontium isotopic equilibration: a solution to a paradox. *Earth Planet. Sci. Lett.* 34: 2338-246 p.
- Taylor, J. R. (1982) An introduction to error analysis : the study of uncertainties in physical measurements. University Science Books, Mill Valley, Calif. 253 pp.
- Titterington, D. y Halliday, A.N. (1979) On the fitting of parallel isochrons and the method of maximum likelihood. *Chem. Geol.* 26: 183-195 p.
- Varadhan, S. (1974) Mathematical statistics. Courant Institute of Mathematical Sciences, New York, 287 pp.

- York, D. (1966) Least-squares fitting of a straight line. *Can. J. Phys.* 44: 1079-1086 p.
- York, D. (1969) Least-squares fitting of a straight line with correlated errors. *Earth Planet. Sci. Lett.* 5: 320-324 p.
- York, D., Evensen, N.E., López Martínez, M. y De Basabe Delgado, J. (en preparación)
The unification of the Least Squares Estimation (LSE) and Maximum Likelihood Estimation (MLE) of the errors in the slope and intercept of the best straight line.
- Young, H.D. (1962) *Statistical treatment of experimental data.* McGraw-Hill, New York, 172 pp.

Apéndice A

El programa *IsoCalc 3.0*

Con el propósito de facilitar la tarea de calcular y gráficar isocronas, se desarrolló el programa *IsoCalc 3.0*. Este utiliza las fórmulas de mínimos cuadrados (MC) de la tabla AI para calcular la pendiente y ordenada al origen. Los errores de los parámetros de la recta también se calculan utilizando las expresiones de la tabla AI, el error de la intercepción con el eje de las abcisas se calcula intercambiando los ejes y utilizando la expresión para σ_a . El programa también cuenta con la opción de calcular los errores con las fórmulas de máxima verosimilitud (MV, tabla AII) o por el método de Monte Carlo.

IsoCalc 3.0 está diseñado para trabajar en el sistema operativo *Windows 95* o superior y fue desarrollado con el lenguaje C++, utilizando las herramientas de *Borland C++ Builder 4.0*. Una de las prioridades en el diseño del programa fue la facilidad de uso, por lo que las principales operaciones se pueden acceder fácilmente a través de botones en la ventana principal. El programa cuenta con un sistema de ayudas interactivas en la parte inferior de la ventana que proporciona información adicional cuando se selecciona un control de la ventana.

Este programa es una versión totalmente rediseñada de las versiones anteriores (De Basabe Delgado y López Martínez, 1999) ya que las versiones anteriores estaban diseñadas para trabajar en la plataforma *Windows 3.1*. En esta nueva versión, además

Tabla AI: Fórmulas de mínimos cuadrados (MC) para los parámetros de la recta y sus errores. Una explicación detallada se encuentra en el capítulo 2.

$$\begin{aligned}
 a &= \bar{Y} - b\bar{X} \\
 b &= \frac{\sum Z_i \beta_i V_i}{\sum Z_i \beta_i U_i} \\
 \sigma_a^2 &= \frac{1}{\sum Z_i} + (\bar{X} + 2\bar{\beta})^2 \sigma_b^2 + 2\bar{\beta} (\bar{X} + 2\bar{\beta}) \left(\frac{\partial \varphi}{\partial b} \right)^{-1} \\
 \sigma_b^2 &= \left\{ \sum_i Z_i^2 \left[\frac{U_i^2}{\omega_{y_i}} + \frac{V_i^2}{\omega_{x_i}} - 2 \frac{r_i}{\alpha_i} U_i V_i \right] \right\} \left(\frac{\partial \varphi}{\partial b} \right)^{-1} \\
 \frac{\partial \varphi}{\partial b} &= \sum_i Z_i \left[\frac{1}{b} U_i V_i - \frac{1}{b} Z_i \frac{r_i}{\alpha_i} (b U_i - V_i)^2 + 4 (\beta_i - U_i) (\beta_i - \bar{\beta}) \right]
 \end{aligned}$$

Tabla AII: Fórmulas de máxima verosimilitud (MV) para los errores de los parámetros de la recta (Titterington y Halliday, 1979).

$$\begin{aligned}
 \sigma_b^2 &= \frac{\sum Z_i}{\sum Z_i \sum x_i^2 Z_i - (\sum x_i Z_i)^2} = \frac{1}{\sum u_i^2 Z_i} \\
 \sigma_a^2 &= \frac{\sum x_i^2 Z_i}{\sum Z_i \sum x_i^2 Z_i - (\sum x_i Z_i)^2} = \sigma_b^2 \frac{\sum x_i^2 Z_i}{\sum Z_i}
 \end{aligned}$$

de la interface nueva, se han incluido más opciones para la isocrona, incluyendo más tipos de isocronas, opciones para los errores calculados y para los errores de los datos de entrada, se agregó el formato TXT para archivos de entrada y salida e histogramas de la pendiente y ordenada al origen.

A.1 Controles de la ventana principal

La ventana principal de *IsoCalc 3.0* tiene tres partes principales (fig. A1):

1. menú principal,
2. barra de herramientas y
3. área de trabajo.

A.1.1 Opciones del menú principal

file ⇒

Open: abrir un archivo TXT o HTM con datos X, Y , sus errores y la correlación de los errores.

Save As: guardar los datos en un archivo TXT o HTM. En el archivo que se crea solo se guardan los datos de la tabla, no se guardan los resultados ni las opciones de configuración.

Exit: salir del programa.

Tools ⇒

Isochron: leer los datos de la tabla, calcular la isocrona y abrir una ventana con la gráfica (fig. A2).

Clear: reiniciar el programa, borrando los datos y los resultados.

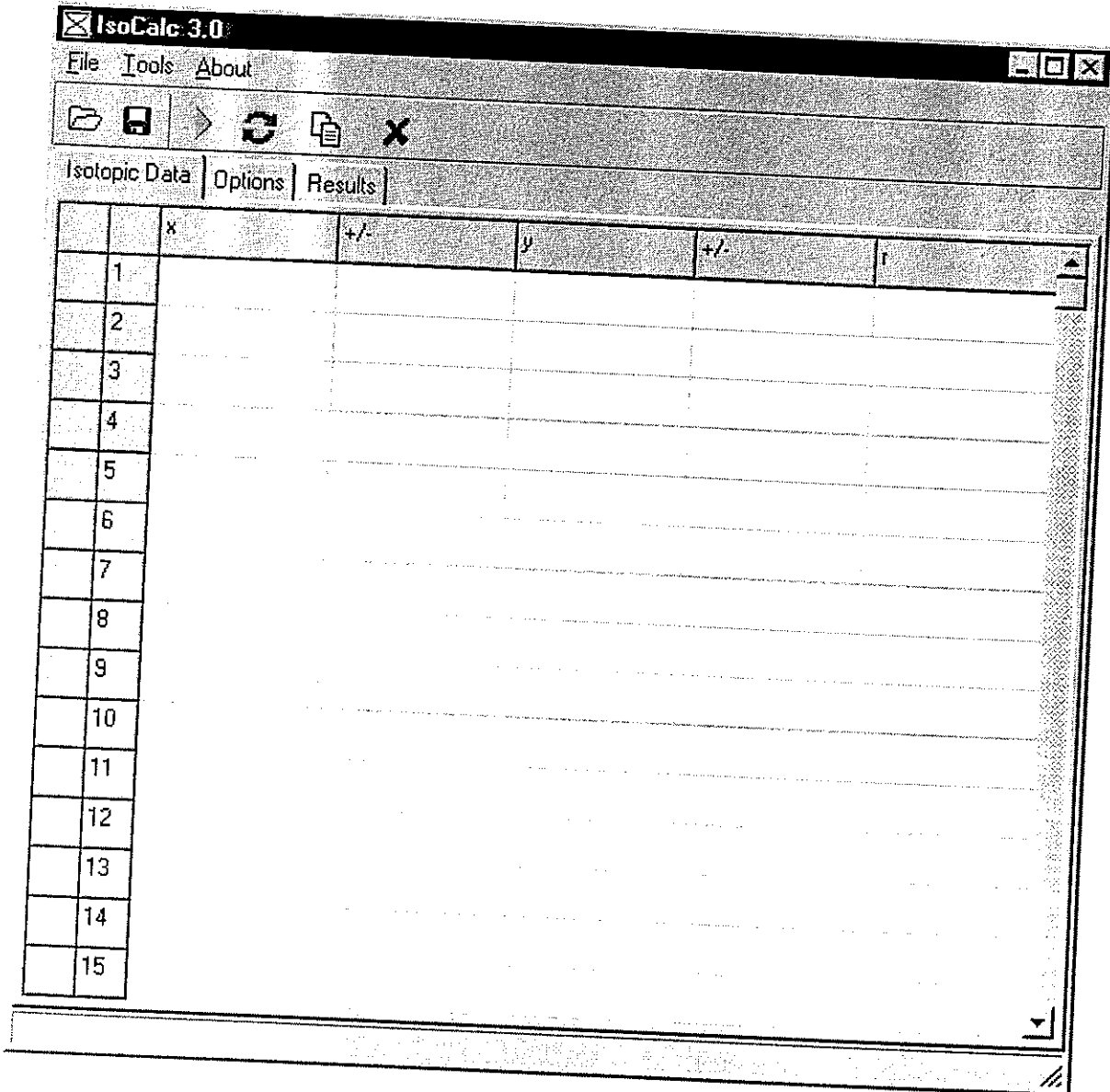


Figura A1: Ventana principal de *IsoCalc 3.0*.

Copy: copiar los resultados al portapapeles de Windows, para pegarlos en otro programa.

Options: activar la ventana de opciones en el área de trabajo.

Slope Histogram: abrir una nueva ventana con el histograma de la pendiente. Esta opción se habilita cuando se han calculado los errores por el método de Monte Carlo.

Y-int Histogram: abrir una nueva ventana con el histograma de la ordenada al origen. Esta opción se habilita cuando se han calculado los errores por el método de Monte Carlo.

About: abrir la ventana con información del programa, como la fecha de la última versión e información sobre los colaboradores del programa.

A.1.2 Botones de la barra de herramientas

Open: abrir un archivo TXT o HTM con datos X, Y y sus errores.

Save: guardar los datos en un archivo TXT o HTM, en el archivo que se crea solo se guardan los datos de la tabla. Al presionar este botón, el archivo no se sobrescribe sino que se abre una ventana para seleccionar un nombre de archivo.

Isochron: leer los datos de la tabla, calcular la isocrona y abrir una ventana con la gráfica (fig. A2).

Clear: reiniciar el programa, borrando los datos y los resultados.

Copy: copiar los resultados al portapapeles de Windows, para pegarlos en otro programa.

Exit: salir del programa.

A.1.3 El área de trabajo

En la parte superior del área de trabajo se pueden ver tres pestañas, al iniciar el programa está seleccionada la de datos isotópicos (*Isotopic Data*) y se puede ver una tabla con 5 columnas para los datos, en esta tabla se pueden escribir los datos directamente o se pueden tomar de un archivo. El orden de los datos es X , σ_X , Y , σ_Y , y r (r es la correlación de los errores). Si existe alguna duda sobre las relaciones isotópicas que se deben de tomar como X y Y en un sistema isotópico se puede seleccionar un tipo de isocrona de las opciones, de esta manera aparece la razón isotópica indicada en el renglón superior, arriba de los datos.

La primera columna de la tabla es para seleccionar o eliminar datos de la isocrona, si aparece el signo '+' el dato está seleccionado, si aparece el signo '-' o aparece en blanco el dato no será tomado en cuenta para los cálculos ni aparecerá en la gráfica. Para seleccionar un dato es necesario que los errores sean diferentes de 0 y que la correlación de los errores esté entre -1 y 1, de otra manera aparece un mensaje de error al seleccionar el dato.

Al seleccionar la pestaña de opciones (*Options*) se mostraran en el área de trabajo las configuraciones posibles del programa, agrupadas de la siguiente manera:

Type: en *Type* se puede seleccionar un sistema isotópico de la lista, en la parte derecha aparece la constante de decaimiento correspondiente al sistema isotópico seleccionado.

Los tipos de isocrona disponibles actualmente son: isocrona normal de $^{40}\text{Ar} - ^{39}\text{Ar}$, isocrona inversa de $^{40}\text{Ar} - ^{39}\text{Ar}$, $\text{Rb} - \text{Sr}$, $\text{Sm} - \text{Nd}$, $\text{Re} - \text{Os}$ y $\text{Lu} - \text{Hf}$. Por omisión está seleccionada la opción $X - Y$, en la cual no se calcula la edad de la isocrona sino solo los parámetros de la recta.

Ar/Ar Data: Si se seleccionó uno de los dos tipos de isocrona de $^{40}\text{Ar} - ^{39}\text{Ar}$ es necesario escribir un valor para el factor de irradiación y su error para que el

programa pueda calcular la edad correspondiente. Si no se dan estos valores, la edad calculada sera 0.

Input Errors: El programa acepta diferentes tipos de errores de entrada, pero se deben de especificar el tipo de errores que se están utilizando para que se puedan realizar correctamente los cálculos. Por omisión se toman los errores de entrada como absolutos y 1-sigma, pero también se pueden seleccionar errores relativos (porcentuales) y errores de 2-sigma.

Output Errors: Se pueden calcular cuatro tipos de errores para la edad de la isocrona y los parámetros de la recta. Por omisión se calculan los errores mediante las fórmulas de MC de la tabla (AI) pero si se selecciona la opción de propagar los errores de los datos ajustados (*Calculate errors from adjusted data*) se calculan los errores con las fórmulas de MV (tabla AII).

Si se selecciona la opción de Monte Carlo (*Monte Carlo method for error propagation*) aparece una ventana de diálogo para escribir la semilla y el número de iteraciones para el algoritmo de Monte Carlo. Por omisión la semilla es 1 y en la mayoría de los casos no es necesario modificar este valor, ya que los resultados no cambian al modificar la semilla si el programa realiza un número suficiente de iteraciones.

Las iteraciones son el número de veces que el programa va a calcular los parámetros de la recta (con los datos perturbados aleatoriamente), se recomienda dar un valor entre 10,000 y 100,000, si se da aquí un número menor que 10,000 es probable que el algoritmo no llegue a convergir y si se da un número mayor que 100,000 el programa se puede tardar varios minutos para realizar los cálculos (en algunos casos puede tomar más de 5 min. para realizar 100,000 iteraciones, dependiendo de la computadora y de los datos).

Si no está seleccionada la opción de propagar los errores de los datos ajustados se

utiliza el algoritmo de Monte Carlo que toma los datos observados para calcular el error (Anderson, 1976). Si está seleccionada la opción de propagar el error de los datos ajustados se utiliza un algoritmo diferente de Monte Carlo, el cual utiliza los datos ajustados para propagar el error.

Al seleccionar la pestaña de Resultados (*Results*) se leen los datos de la tabla y las opciones para calcular la isocrona y la edad. Es importante notar que aparecen dos tipos de errores para la edad y los parámetros de la recta, el primero es el error estandar y el segundo está multiplicado por el factor $\sqrt{S/(n-2)}$ (York, 1969). En la parte inferior derecha se puede ver un botón para guardar estos resultados en un archivo, este botón abre una ventana de diálogo en la que se puede seleccionar el nombre del archivo que se desea crear.

A.2 La ventana de la isocrona

La ventana que contiene la gráfica de la isocrona (fig. A2) tiene una barra de herramientas con 5 botones, que realizan las siguientes funciones:

Options: abrir una ventana de diálogo donde se pueden modificar las opciones de la isocrona, en esta ventana se pueden escoger los valores mínimos y máximos de los ejes, se pueden eliminar los números de los datos y se pueden graficar los intervalos de confianza ¹.

Save: abrir una ventana de diálogo para seleccionar el nombre de un archivo para guardar la isocrona y el formato del archivo. Los tipos de archivos disponibles son: Metafile de Windows (WMF y EMF) y bitmap (BMP).

¹Los intervalos de confianza se calculan según Ludwig (1980). Se ha observado en algunos casos que las hipérbolas se pueden distorsionar considerablemente debido a la no-linealidad de los parámetros de la recta.

Print: abrir una ventana de diálogo en la que se muestra la forma en la que la gráfica se va a imprimir, en esta ventana se pueden modificar las opciones de impresión y enviar la gráfica a la impresora.

Copy: copiar la gráfica en formato Metafile de Windows al portapapeles de Windows, la gráfica se puede pegar después en cualquier programa que soporte este formato².

Close: cerrar la ventana de la isocrona.

Se pueden modificar otras opciones de visualización de la isocrona, con el ratón de la siguiente manera:

Borrar datos: si se presiona el botón izquierdo del ratón sobre uno de los datos, se recalcula y grafica de nuevo la isocrona sin ese dato. Se pueden identificar fácilmente los datos que han sido eliminados porque aparecen en la gráfica de color rojo. Para incluir de nuevo un dato solo se presiona de nuevo el botón izquierdo del ratón sobre el dato.

Acercamiento: se puede hacer más grande un área de la gráfica, al presionar el botón izquierdo sobre la esquina superior izquierda y arrastrarlo hasta la esquina inferior derecha del área que se quiere ampliar. Para regresar a la forma inicial solo se presiona el botón izquierdo del ratón en cualquier parte de la gráfica y se arrastra hacia cualquier punto que quede hacia arriba y a la izquierda.

Mover: para acomodar la gráfica se puede presionar el botón derecho y arrastrarlo en cualquier dirección hasta que la gráfica se muestre de la forma deseada.

²Se ha observado una incompatibilidad, por razones desconocidas, al tratar de pegar la gráfica en el programa Corel Draw.

A.3 Histogramas

Si se calculan los errores con el algoritmo de Monte Carlo se habilitan las opciones de histogramas del menú *Tools*, estas opciones abren las ventanas con la gráfica del histograma correspondiente, el de la pendiente o el de la ordenada al origen (fig. A3).

En esta gráfica se pueden comparar los errores de MC y de MV con los de Monte Carlo. Las barras azules son la distribución de los parametros de la recta. En algunos casos la curva de Monte Carlo (en color rojo) no se ajusta al histograma, esto es debido a que la distribución del histograma puede no ser exactamente normal, sino que puede estar sesgada o tener una curtosis anormalmente alta o baja, los valores de sesgo y curtosis se pueden ver en el archivo `montecarlo.log` que se crea en el directorio donde está el archivo de entrada. En este archivo se registran 10 renglones con resultados parciales que muestran la forma en la que los resultados de Monte Carlo convergen. Este es un archivo temporal, que se sobrescribe al utilizar la rutina de Monte Carlo, si se desea consultar posteriormente esta información es necesario renombrar el archivo.

La barra de herramientas de los histogramas tiene los siguientes botones:

Save: abrir una ventana de diálogo para seleccionar el nombre de un archivo para guardar el histograma y el formato del archivo. Los tipos de archivos disponibles son: Metafile de Windows (WMF y EMF) y bitmap (BMP).

Print: abrir una ventana de diálogo en la que se muestra la forma en la que el histograma se va a imprimir, en esta ventana se pueden modificar las opciones de impresión y enviar la gráfica a la impresora.

Copy: copiar el histograma en formato Metafile de Windows al portapapeles de Windows, se puede pegar después el histograma en cualquier programa que soporte este formato ³.

³Se ha observado un error, por razones desconocidas, al tratar de pegar el histograma en el programa Corel Draw.

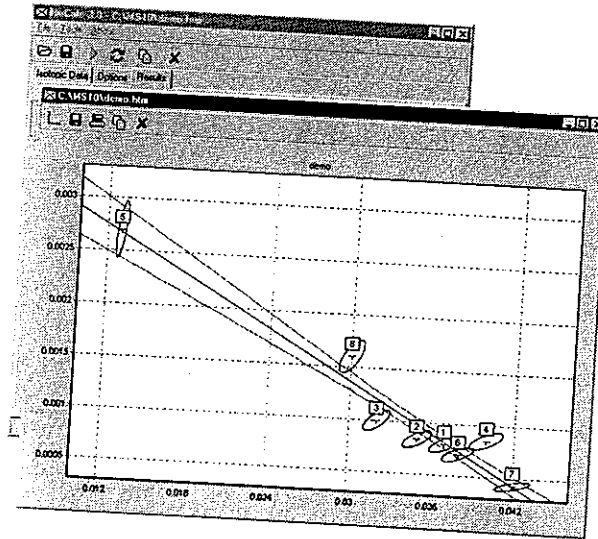


Figura A2: Ventana de *IsoCalc 3.0* con la isocrona y los datos gráficos.

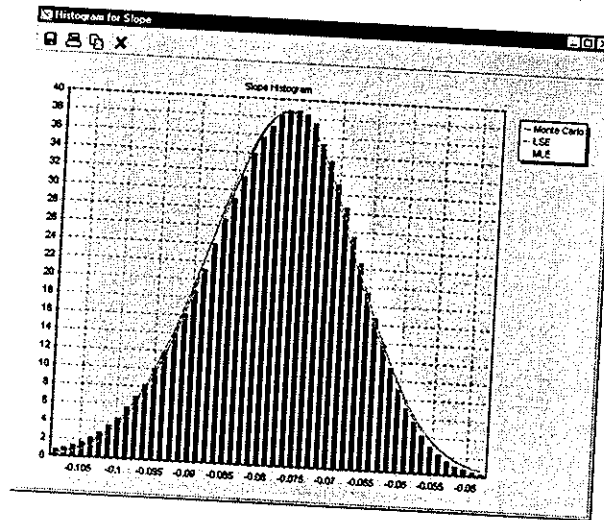


Figura A3: Ventana del histograma de la pendiente. LSE corresponde al método de mínimos cuadrados y MLE al de máxima verosimilitud.

Close: cerrar la ventana del histograma.

A.4 Archivos de entrada

El programa acepta dos tipos de archivos para las relaciones isotópicas: archivos TXT y archivos HTM. El archivo TXT es un archivo de texto con extensión TXT, en el primer renglón, que es ignorado por el programa, puede tener información de la muestra o información pertinente sobre los datos, en los renglones siguientes deben estar los datos ordenados de forma tabular, con n renglones y 5 columnas (donde n es el número de datos). El orden de las columnas es: X , σ_x , Y , σ_y y r (tabla AIII).

El archivo HTM debe contener una tabla formateada debidamente en lenguaje HTML de n renglones por 4 o 5 columnas (si la tabla tiene 4 columnas se asume una correlación igual a 0). El primer renglón es ignorado y puede tener información relacionada con los datos. El orden de los datos es el mismo que para el archivo TXT.

Tabla AIII: Ejemplo de un archivo TXT con datos isotópicos

Isocrona Inversa de Ar/Ar. Datos tomados de Mahon (1996)				
3.70e-02	1.11e-03	8.00e-04	8.00e-05	7.071e-01
3.50e-02	1.05e-03	8.40e-04	8.40e-05	7.071e-01
3.20e-02	9.60e-04	1.00e-03	1.00e-04	7.071e-01
4.00e-02	1.20e-03	8.50e-04	8.50e-05	7.071e-01
1.30e-02	3.90e-04	2.70e-03	2.70e-04	7.071e-01
3.80e-02	1.14e-03	7.10e-04	7.10e-05	7.071e-01
4.20e-02	1.26e-03	4.30e-04	4.30e-05	7.071e-01
3.00e-02	9.00e-04	1.60e-03	1.60e-04	7.071e-01

Apéndice B

Demostraciones complementarias

Demostración 1 $\partial Z_i / \partial b$

$$\begin{aligned}\frac{\partial Z_i}{\partial b} &= \frac{\partial}{\partial b} \left(\frac{\omega_{x_i} \omega_{y_i}}{\omega_{x_i} + b^2 \omega_{y_i} - 2br_i \alpha_i} \right) \\ &= \frac{-\omega_{x_i} \omega_{y_i} (2b\omega_{y_i} - 2r_i \alpha_i)}{(\omega_{x_i} + b^2 \omega_{y_i} - 2br_i \alpha_i)^2} \\ &= -2Z_i^2 \left(\frac{b\omega_{y_i} - r_i \alpha_i}{\omega_{x_i} \omega_{y_i}} \right) \\ &= -2 \frac{Z_i^2}{\epsilon_i}\end{aligned}$$

Demostración 2 $\partial U_i / \partial b$

$$\begin{aligned}\frac{\partial U_i}{\partial b} &= \frac{\partial}{\partial b} (X_i - \bar{X}) \\ &= -\frac{\partial \bar{X}}{\partial b} \\ &= -\frac{\partial \sum Z_j X_j}{\partial b \sum Z_j} \\ &= -\frac{\sum Z_j \sum \frac{\partial Z_j}{\partial b} X_j - \sum Z_j X_j \sum \frac{\partial Z_j}{\partial b}}{(\sum Z_j)^2} \\ &= -\frac{\sum \frac{\partial Z_j}{\partial b} X_j - \bar{X} \sum \frac{\partial Z_j}{\partial b}}{\sum Z_j} \\ &= -\frac{\sum \frac{\partial Z_j}{\partial b} U_j}{\sum Z_j}\end{aligned}$$

$$\begin{aligned}
 &= -\frac{1}{\sum Z_j} \sum \frac{-2Z_j^2}{\epsilon_j} U_j \\
 &= \frac{2}{\sum Z_j} \sum \frac{Z_j^2 U_j}{\epsilon_j}
 \end{aligned}$$

Demostración 3 $\partial V_i / \partial b$

Siguiendo el mismo desarrollo que en la demostración 2

$$\begin{aligned}
 \frac{\partial V_i}{\partial b} &= \frac{\partial}{\partial b} (Y_i - \bar{Y}) \\
 &= -\frac{\sum \frac{\partial Z_j}{\partial b} Y_j - \bar{Y} \sum \frac{\partial Z_j}{\partial b}}{\sum Z_j} \\
 &= -\frac{\sum \frac{\partial Z_j}{\partial b} V_j}{\sum Z_j} \\
 &= \frac{2}{\sum Z_j} \sum \frac{Z_j^2 V_j}{\epsilon_j}
 \end{aligned}$$

Demostración 4 $\partial \beta_i / \partial b$

$$\begin{aligned}
 \frac{\partial \beta_i}{\partial b} &= \frac{\partial}{\partial b} \left[Z_i \left(\frac{U_i}{\omega_{Y_i}} + \frac{bV_i}{\omega_{X_i}} - (bU_i + V_i) \frac{r_i}{\alpha_i} \right) \right] \\
 &= \frac{-2Z_i^2}{\epsilon_i} \left(\frac{U_i}{\omega_{Y_i}} + \frac{bV_i}{\omega_{X_i}} - (bU_i + V_i) \frac{r_i}{\alpha_i} \right) + \\
 &\quad Z_i \left(\frac{\partial U_i}{\partial b} \left(\frac{1}{\omega_{Y_i}} - \frac{br_i}{\alpha_i} \right) + \frac{\partial V_i}{\partial b} \left(\frac{b}{\omega_{X_i}} - \frac{r_i}{\alpha_i} \right) + \frac{V_i}{\omega_{X_i}} - \frac{r_i U_i}{\alpha_i} \right) \\
 &= -2 \frac{Z_i \beta_i}{\epsilon_i} + Z_i \left(\frac{\partial U_i}{\partial b} \left(\frac{1}{\tau_i} \right) + \frac{\partial V_i}{\partial b} \left(\frac{1}{\epsilon_i} \right) + \frac{V_i}{\omega_{X_i}} - \frac{r_i U_i}{\alpha_i} \right) \\
 &= -2 \frac{Z_i \beta_i}{\epsilon_i} + Z_i \left(\frac{2}{\tau_i \sum_j Z_j} \sum_j \frac{Z_j^2 U_j}{\epsilon_j} + \frac{2}{\epsilon_i \sum_j Z_j} \sum_j \frac{Z_j^2 V_j}{\epsilon_j} + \frac{V_i}{\omega_{X_i}} - \frac{r_i U_i}{\alpha_i} \right)
 \end{aligned}$$

Demostración 5 $v_i = bu_i$

Dado que los datos ajustados (x_i, y_i) cumplen con la ecuación de la recta $y_i = a + bx_i$

$$\begin{aligned}
 \sum Z_i y_i &= a \sum Z_i + b \sum Z_i x_i \\
 \bar{y} &= a + b\bar{x}
 \end{aligned}$$

restando esta ecuación a la ecuación de la recta

$$y_i = \bar{y} - b\bar{x} + bx_i$$

$$y_i - \bar{y} = b(x_i - \bar{x})$$

$$v_i = bu_i$$

Demostración 6 Sea $\bar{\delta}_i$ una diferencia de la forma $\bar{\delta}_i = \delta_i - \bar{\delta}$, donde $\bar{\delta} = (\sum Z_i \delta_i) / \sum Z_i$, entonces

$$\begin{aligned} \sum Z_i \bar{\delta}_i &= \sum Z_i (\delta_i - \bar{\delta}) \\ &= \sum Z_i \left(\delta_i - \frac{\sum Z_i \delta_i}{\sum Z_i} \right) \\ &= \sum Z_i \delta_i - \sum Z_i \delta_i \\ &= 0 \end{aligned}$$

Se observa que las diferencias que cumplen con la definición de $\bar{\delta}_i$ son $\bar{\beta}_i, U_i, V_i, u_i$ y v_i , por lo tanto $\sum Z_i \bar{\beta}_i = 0, \sum Z_i U_i = 0, \sum Z_i V_i = 0, \sum Z_i u_i = 0$ y $\sum Z_i v_i = 0$.

Demostración 7 $\sum Z_i \bar{\beta}_i = 0$

$$\begin{aligned} \sum Z_i \bar{\beta}_i (\bar{\beta}_i - \beta_i) &= \sum Z_i \bar{\beta}_i (\beta_i - \bar{\beta} - \beta_i) \\ &= -\bar{\beta} \sum Z_i \bar{\beta}_i \end{aligned}$$

pero $\sum Z_i \bar{\beta}_i = 0$ (demostración 6), por lo tanto $\sum Z_i \bar{\beta}_i (\bar{\beta}_i - \beta_i) = 0$

Demostración 8 $\frac{\sum_i Z_i (\beta_i - 2\bar{\beta}_i)}{\sum Z_j} = \bar{\beta}$

$$\begin{aligned} \frac{\sum_i Z_i (\beta_i - 2\bar{\beta}_i)}{\sum Z_j} &= \frac{\sum_i Z_i (\beta_i - 2\beta_i + 2\bar{\beta})}{\sum Z_j} \\ &= \frac{\sum_i Z_i (-\beta_i + 2\bar{\beta})}{\sum Z_j} \\ &= \frac{2\bar{\beta} \sum_i Z_i - \sum_i Z_i \beta_i}{\sum Z_j} \\ &= 2\bar{\beta} - \bar{\beta} \\ &= \bar{\beta} \end{aligned}$$

Demostración 9 $\frac{Z_i^2}{\tau_i} (bU_i - V_i) = Z_i (b\beta_i - V_i)$

Sustituyendo el valor de Z_i en β_i

$$\begin{aligned}\beta_i &= \frac{\omega_{x_i} \omega_{y_i}}{\omega_{x_i} + b^2 \omega_{y_i} - 2br_i \alpha_i} \left(\frac{U_i}{\omega_{y_i}} + \frac{bV_i}{\omega_{x_i}} - (bU_i + V_i) \frac{r_i}{\alpha_i} \right) \\ &= \frac{\omega_{x_i} U_i + b\omega_{y_i} V_i - br_i \alpha_i U_i - r_i \alpha_i V_i}{\omega_{x_i} + b^2 \omega_{y_i} - 2br_i \alpha_i}\end{aligned}$$

Evaluando $b\beta_i - V_i$ con β_i expresada de esta manera

$$\begin{aligned}b\beta_i - V_i &= \frac{b\omega_{x_i} U_i + b^2 \omega_{y_i} V_i - b^2 r_i \alpha_i U_i - br_i \alpha_i V_i - \omega_{x_i} V_i - b^2 \omega_{y_i} V_i + 2br_i \alpha_i V_i}{\omega_{x_i} + b^2 \omega_{y_i} - 2br_i \alpha_i} \\ &= \frac{\omega_{x_i} (bU_i - V_i) - br_i \alpha_i (bU_i - V_i)}{\omega_{x_i} + b^2 \omega_{y_i} - 2br_i \alpha_i} \\ &= \frac{(\omega_{x_i} - br_i \alpha_i)(bU_i - V_i)}{\omega_{x_i} + b^2 \omega_{y_i} - 2br_i \alpha_i} \\ &= Z_i \frac{(\omega_{x_i} - br_i \alpha_i)(bU_i - V_i)}{\omega_{x_i} \omega_{y_i}} \\ &= \frac{Z_i}{\tau_i} (bU_i - V_i)\end{aligned}$$

por lo tanto $\frac{Z_i^2}{\tau_i} (bU_i - V_i) = Z_i (b\beta_i - V_i)$

Demostración 10 $\frac{Z_i^2}{\epsilon_i} (bU_i - V_i) = Z_i (U_i - \beta_i)$

Utilizando un desarrollo similar al de la demostración 9, se evalúa $U_i - \beta_i$

$$\begin{aligned}U_i - \beta_i &= \frac{\omega_{x_i} U_i + b^2 \omega_{y_i} U_i - 2br_i \alpha_i U_i - \omega_{x_i} U_i - b\omega_{y_i} V_i + br_i \alpha_i U_i + r_i \alpha_i V_i}{\omega_{x_i} + b^2 \omega_{y_i} - 2br_i \alpha_i} \\ &= \frac{b\omega_{y_i} (bU_i - V_i) - r_i \alpha_i (bU_i - V_i)}{\omega_{x_i} + b^2 \omega_{y_i} - 2br_i \alpha_i} \\ &= Z_i \frac{(b\omega_{y_i} - r_i \alpha_i)(bU_i - V_i)}{\omega_{x_i} \omega_{y_i}} \\ &= \frac{Z_i}{\epsilon_i} (bU_i - V_i)\end{aligned}$$

por lo tanto $\frac{Z_i^2}{\epsilon_i} (bU_i - V_i) = Z_i (U_i - \beta_i)$

Apéndice C

Datos graficados

A continuación se muestran las gráficas de los datos utilizados como ejemplos en el capítulo 4.

Tabla CI: Resumen de resultados calculados para los datos que se presentan en este capítulo.

Datos generales

Fig.	Datos	Referencia	n	S	$MSWD$	Corr	r
C1	York-66	York (1966)	10	11.866	1.483	-0.9765	0
C2	Mahon-1	Mahon (1996)	8	10.382	1.730	-0.9740	0
C3	Mahon-2	Mahon (1996)	8	6.266	1.044	-0.9740	0.7071
C4	Reed-i-1	Reed (1989)	27	578.047	23.122	0.8202	0
C5	Reed-i-2	Reed (1989)	10	13.956	1.744	-0.1708	0

Pendiente, ordenada al origen y sus errores

Datos	a	b	MC		MV	
			$\hat{\sigma}_a$	$\hat{\sigma}_b$	$\bar{\sigma}_a$	$\bar{\sigma}_b$
York-66	5.47991	-0.480533	0.291934	0.0576167	0.294971	0.0579850
Mahon-1	0.00364111	-0.0763849	0.000349769	0.00952059	0.000320288	0.00871306
Mahon-1	0.00368416	-0.0775157	0.000375911	0.0102713	0.000364644	0.00996480
Reed-i-1	-0.365155	1.16684	0.0324699	0.0354464	0.0280406	0.0305674
Reed-i-2	-17.4835	4.54366	55.1930	10.9604	26.9173	5.33265

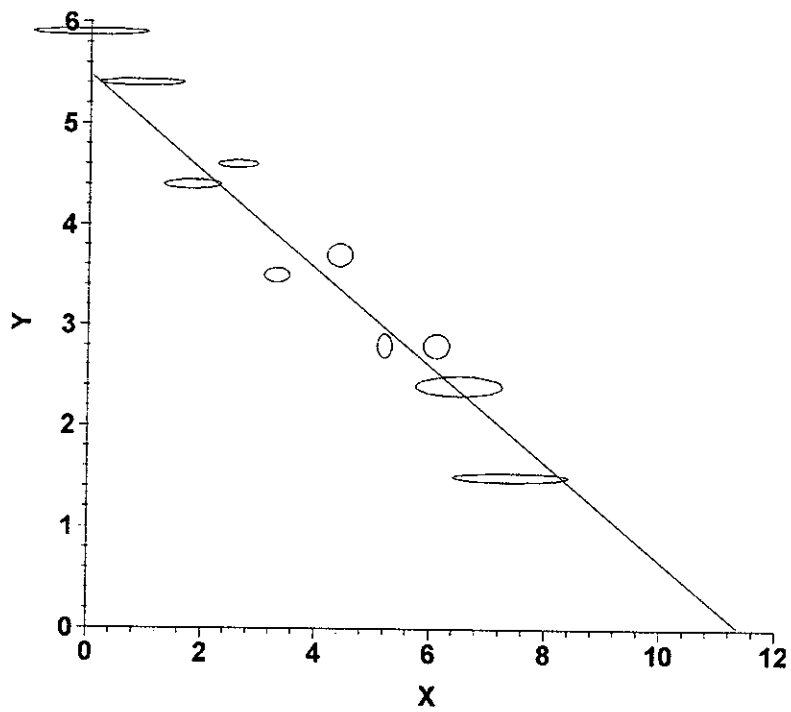


Figura C1: Datos York-66, tomados de York (1966).

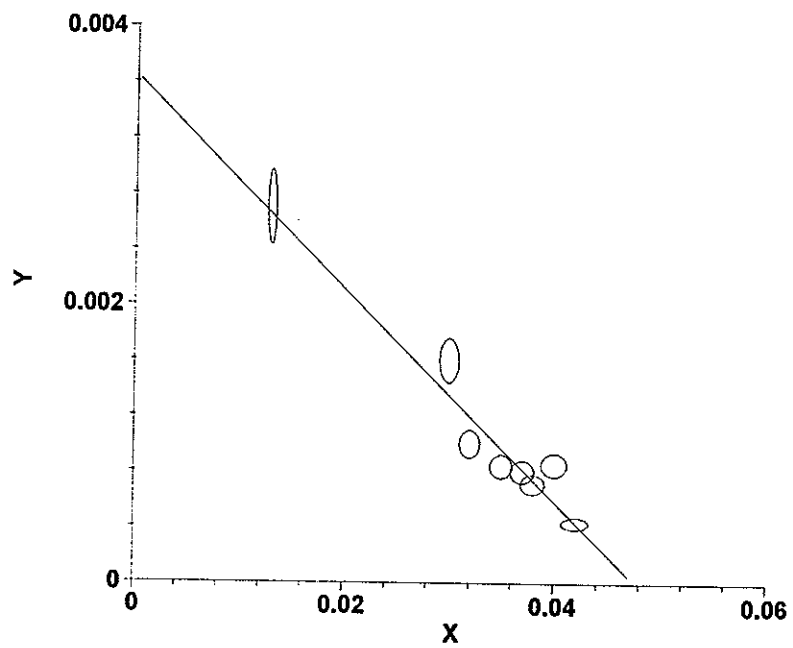


Figura C2: Datos Mahon-1, tomados de Mahon (1996).

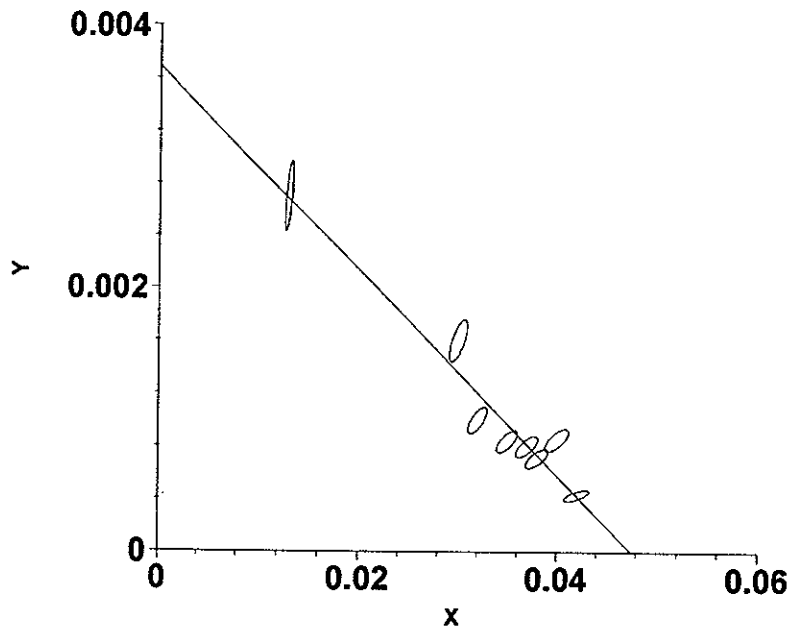


Figura C3: Datos Mahon-2, tomados de Mahon (1996).

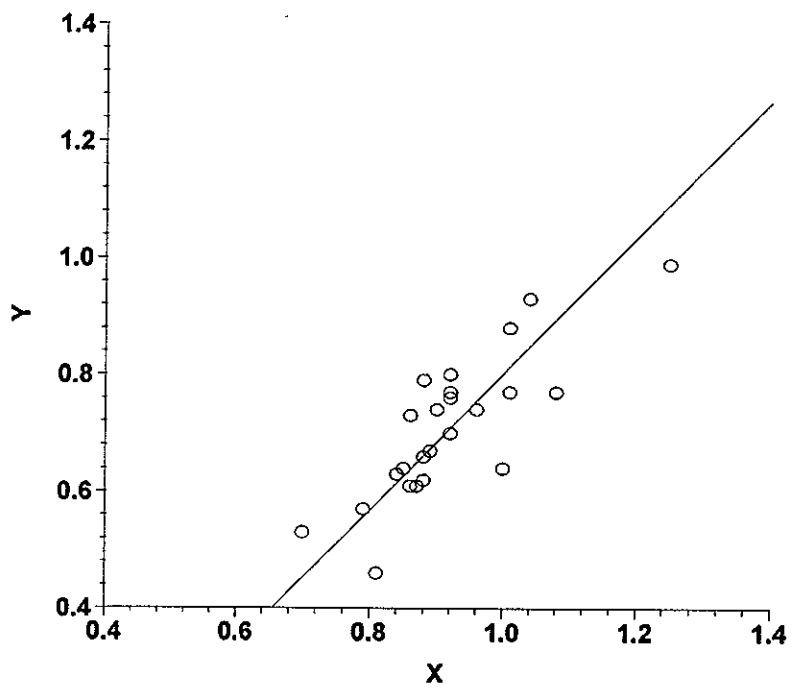


Figura C4: Datos Reed-i-1, tomados de Reed (1989).

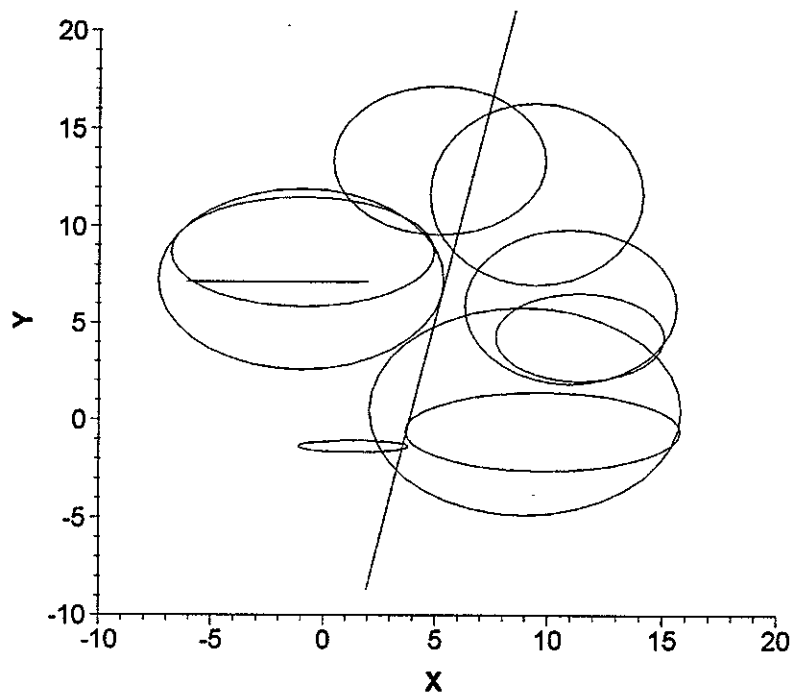


Figura C5: Datos Reed-i-2, tomados de Reed (1989).