

ACERCAMIENTO A LA ESTIMACIÓN NO LINEAL EN HIDROLOGÍA¹

M.I. Sergio Ignacio Martínez Martínez²

INTRODUCCIÓN

Este trabajo nace de la necesidad práctica de realizar mejores estimaciones de los parámetros de las ecuaciones que comúnmente se presentan en la hidrología. Se pretende generar un procedimiento sistemático para efectuar la estimación no lineal de dichos parámetros. Se supone que un refinamiento del proceso de estimación de los parámetros de ajuste de las ecuaciones que gobiernan diversos fenómenos hidrológicos puede llevar a obtener mejores estimaciones de dichos parámetros de ajuste.

MATERIALES Y MÉTODOS

La investigación se desarrolló en el Centro de Ciencias del Diseño y de la Construcción durante el año 2001; su marco teórico se inserta dentro del campo de la Estadística. Se recurrió al procesamiento numérico de datos encontrados en la literatura y en la práctica profesional de la Hidrología Superficial.

Ecuaciones lineales y ecuaciones no lineales

Importante para el fin de este artículo es la inclusión de la siguiente explicación que ilustra la hipótesis que guió la investigación. Se entiende por *ecuación lineal* aquella ecuación que es *lineal en sus parámetros* y como *no lineal* aquella que es *no lineal en sus parámetros*. Así, si se deriva parcialmente una ecuación con respecto a sus parámetros y se encuentra que ninguna de esas

derivadas parciales está en función de uno o más de los parámetros, entonces se dice que es una ecuación lineal; en caso contrario, es una ecuación no lineal. Muchos modelos matemáticos sencillos que aparecen en la hidrología son no lineales en los parámetros, por ejemplo, supóngase que la ecuación:

$$y = ax^b \quad (1)$$

se puede ajustar a una serie de n pares de datos (x_i, y_i) . El valor óptimo de los parámetros a y b se obtendrá encontrando el valor mínimo de la suma de cuadrados de los errores:

$$SSE_{nl} = \sum_{i=1}^n (y_i - ax_i^b)^2 \quad (2)$$

para calcular el mínimo de esta suma se tendrá que derivar con respecto a cada uno de los parámetros e igualar a cero. Si se toma en cuenta que:

$$\frac{\partial y}{\partial a} = x^b \quad y \quad \frac{\partial y}{\partial b} = ax^b \ln x \quad (3)$$

se formará un sistema de dos ecuaciones no lineales que deberá resolverse en términos de a y b . Este sistema es, relativamente, difícil de resolver, por lo que la ecuación 1 se suele transformar en la siguiente ecuación lineal:

$$\log y = \log a + b \log x \quad (4)$$

que toma la forma (en el plano $\log x - \log y$):

$$Y = A + BX \quad (5)$$

lo que sugiere, inmediatamente, hacer una regresión lineal, esto es, calcular el mínimo de la suma de cuadrados de los errores, pero ahora con:

$$SSE_l = \sum_{i=1}^n (Y_i - (A + BX_i))^2 \quad (6)$$

como:

$$\frac{\partial Y}{\partial A} = 1 \quad y \quad \frac{\partial Y}{\partial B} = X \quad (7)$$

¹ Investigación inscrita en el Programa de Investigaciones en Ingeniería. Se presentó una versión previa en el Tercer Seminario de Investigación del Centro de Ciencias del Diseño y de la Construcción (CCDC) de la Universidad Autónoma de Aguascalientes (UAA), H a 16 de enero de 2002.

² Profesor e Investigador del Departamento de Construcción y Estructuras, Centro de Ciencias del Diseño y de la Construcción. E-mail: simartin@correo.uaa.mx

se formará un sistema de dos ecuaciones lineales en A y B , llamadas *ecuaciones normales* [Walpole et al, 1996]. El sistema se resuelve fácilmente. También se puede calcular la suma total de cuadrados:

$$SST_i = \sum_{i=1}^n (Y_i - \bar{Y})^2 \quad (8)$$

donde \bar{Y} es la media de los valores $Y_i = (Y_i = \log y_i)$. Luego, se calcula el coeficiente de determinación r_i^2 (igual al cuadrado del coeficiente de correlación):

$$r_i^2 = \frac{SST_i - SSE_i}{SST_i} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2 - \sum_{i=1}^n (Y_i - (A + BX_i))^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (9)$$

r_i^2 puede tomar cualquier valor entre 0.0 y 1.0. Si su valor es cercano a uno se dice que existe una buena correlación (lineal) entre las variables X y Y . Finalmente, los parámetros a y b , se encuentran con:

$$a = 10^A \quad y \quad b = B \quad (10)$$

A veces, además de revisar el valor de r_i , se revisa gráficamente en un diagrama de dispersión si los datos tienen la tendencia de la curva ajustada; en ese caso se aceptan, sin ninguna otra prueba, las estimaciones de los parámetros a y b . El valor de r_i^2 es engañoso; es mejor calcular el coeficiente de determinación en función de las variables originales y de los parámetros a y b calculados antes; o sea:

$$r_{nl}^2 = \frac{SST_{nl} - SSE_{nl}}{SST_{nl}} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - ax_i^b)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

donde \bar{y} es la media de los valores y_i . Este nuevo coeficiente dará una medida más realista del ajuste, con un valor numérico usualmente menor que el calculado con la ecuación 9. En el caso de resolver el sistema de ecuaciones no lineales, o sea, de efectuar la regresión no lineal y calcular el coeficiente de correlación no lineal (ecuación 11), se tendrá un valor mayor o igual al obtenido con la regresión lineal y la misma ecuación 11. La simplicidad de la regresión lineal conlleva correr riesgos. Al hacer una estimación de parámetros efectuando una transformación de la ecuación original se introducen errores de escala que harán que las estimaciones no sean estadísticamente satisfactorias y, además, tendrán una consecuencia práctica para el analista, puesto que podrían llevarlo a cometer errores de consecuencias desastrosas, sobre todo si no echa mano de técnicas que verifiquen sus ajustes y de su buen juicio ingenieril (ver, por ejemplo: [Berezowsky et al, 1982], [Demetracopoulos, 1994] y [McCuen et al, 1990]).

Método de Marquardt

Enseguida se describe un camino para estimar los parámetros de ajuste y la mínima suma de cuadrados de los errores considerando una ecuación de ajuste no lineal. Sea el modelo siguiente [Constantinides, 1987]:

$$Y = f(X, b) \quad (12)$$

donde Y es el vector de los valores calculados de la variable dependiente, X es la matriz de valores de las variables independientes y b es el vector de parámetros de ajuste. Para el caso en que el modelo contiene sólo una variable dependiente, la suma de cuadrados de los errores o residuos está dada por:

$$SSE_{nl} = \Phi = (Y^* - Y)^T (Y^* - Y) \quad (13)$$

Y^* es el vector de observaciones experimentales de la variable dependiente. Dado que Y es no lineal con respecto a los parámetros, tomando la derivada parcial de Φ con respecto a b e igualándola a cero resultará una ecuación no lineal que será difícil de resolver para b . Gauss resolvió este problema, encontrando que el ajuste de funciones no lineales por mínimos cuadrados puede realizarse mediante un método iterativo que involucra una serie de aproximaciones lineales. En cada iteración, se puede utilizar la teoría de mínimos cuadrados para obtener la siguiente aproximación. Este método, conocido como el método de *Gauss-Newton*, convierte el problema no lineal en uno lineal aproximando la función Y mediante una expansión en serie de Taylor alrededor del valor estimado del vector de parámetros b :

$$Y(X, b + \Delta b) = Y(X, b) + \frac{\partial Y}{\partial b} \Delta b \quad (14)$$

En la ecuación 14 la serie de Taylor ha sido truncada después del segundo término. La ecuación 14 es lineal en Δb , por lo tanto, el problema se ha transformado de encontrar b a encontrar una corrección de b , esto es, Δb , que debe sumarse a la estimación de b para minimizar la suma de residuos al cuadrado. Para hacer esto se reemplaza Y en la ecuación 13 con el lado derecho de la ecuación 14 para obtener:

$$\Phi = (Y^* - Y - A\Delta b)^T (Y^* - Y - A\Delta b) \quad (15)$$

A es la matriz jacobiana de derivadas parciales de Y con respecto a b , evaluada en los n puntos donde se dispone de observaciones experimentales:

$$A = \begin{bmatrix} \frac{\partial Y_1}{\partial b_1} & \dots & \frac{\partial Y_1}{\partial b_k} \\ \dots & \dots & \dots \\ \frac{\partial Y_n}{\partial b_1} & \dots & \frac{\partial Y_n}{\partial b_k} \end{bmatrix} \quad (16)$$

Tomando la derivada parcial de Φ con respecto a $\Delta\mathbf{b}$, igualándola a cero, y resolviendo para $\Delta\mathbf{b}$, se obtiene:

$$\Delta\mathbf{b} = (\mathbf{A}^T\mathbf{A})^{-1} \mathbf{A}^T(\mathbf{Y}^* - \mathbf{Y}) \quad (17)$$

donde $\Delta\mathbf{b}$ es el vector de corrección que se sumará al vector de la actual iteración i , \mathbf{b}_i , para obtener una nueva estimación del vector de parámetros, \mathbf{b}_{i+1} :

$$\mathbf{b}_{i+1} = \mathbf{b}_i + \Delta\mathbf{b} \quad (18)$$

El método de Gauss-Newton se basa en la linealización del modelo no lineal; por lo tanto, es de esperarse que trabaje bien si el modelo es sólo un poco no lineal, o si la estimación inicial de los parámetros está cerca de la suma de cuadrados mínima. Los contornos de Φ constante, en el espacio de los parámetros de un modelo lineal, son elipsoides. Para un modelo no lineal, esos contornos están distorsionados, salvo en la vecindad del mínimo Φ , donde son cuasi-elípticos. Por tanto, el método de Gauss-Newton es muy efectivo si el punto de salida inicial para la búsqueda está en la región cuasi-elíptica. Por otro lado, este método puede divergir si el punto de comienzo está en una región muy distorsionada del hiperespacio de los parámetros. Otro método [Constantinides, 1987], que ha sido utilizado para encontrar el mínimo de la suma de cuadrados de un modelo no lineal, es el llamado del *descenso rápido*. En este método, el vector inicial de estimación de parámetros se corrige en la dirección del gradiente negativo de Φ :

$$\Delta\mathbf{b} = -K \left[\frac{\partial\Phi}{\partial\mathbf{b}} \right] \quad (19)$$

Donde K es un factor constante adecuado. De la ecuación 13, efectuando la derivada parcial de Φ con respecto a \mathbf{b} :

$$\left[\frac{\partial\Phi}{\partial\mathbf{b}} \right] = -2\mathbf{A}^T(\mathbf{Y}^* - \mathbf{Y}) \quad (20)$$

Y combinando con la ecuación 19 se obtiene:

$$\Delta\mathbf{b} = 2K\mathbf{A}^T(\mathbf{Y}^* - \mathbf{Y}) \quad (21)$$

Al comparar las ecuaciones 17 y 21 se ve que la única diferencia entre los vectores de corrección de los métodos de Gauss-Newton y del descenso rápido es la matriz $(\mathbf{A}^T\mathbf{A})^{-1}$ y el factor constante $2K$. El método del descenso rápido tiene la ventaja, sobre el de Gauss-Newton, de que no diverge, siempre que K , factor que determina el tamaño del paso, sea lo suficientemente pequeño. Sin embargo, la velocidad de convergencia hacia la suma de cuadrados mínima decrece conforme la búsqueda se

acerca al mínimo, por lo que este método pierde su atractivo. Marquardt publicó en 1963 una técnica de interpolación entre el método de Gauss-Newton y el del descenso rápido [Constantinides, 1987]. Esta interpolación se obtiene añadiendo la matriz diagonal $\lambda\mathbf{I}$ a $\mathbf{A}^T\mathbf{A}$ en la ecuación 17:

$$\Delta\mathbf{b} = (\mathbf{A}^T\mathbf{A} + \lambda\mathbf{I})^{-1} \mathbf{A}^T(\mathbf{Y}^* - \mathbf{Y}) \quad (22)$$

El valor de λ se escoge, en cada iteración, para que el vector de parámetros corregido dé una menor suma de cuadrados en la iteración siguiente. Cuando el valor de λ es pequeño en comparación con los elementos de $\mathbf{A}^T\mathbf{A}$, el método de Marquardt se acerca al de Gauss-Newton; cuando λ es muy grande, el método es idéntico al del descenso rápido, con la excepción del factor de escala, que no afecta la dirección del vector de corrección de los parámetros, pero proporciona un tamaño de paso muy pequeño. De acuerdo con Marquardt, es deseable minimizar Φ en la máxima vecindad sobre la cual la función linealizada da una adecuada representación de la función no lineal. Así, el procedimiento para escoger λ debe llevar a valores pequeños cuando el método de Gauss-Newton pueda converger eficientemente y valores grandes para cuando sea necesario el método del descenso rápido. El algoritmo del método de Marquardt involucra los siguientes pasos:

1. Hacer $i = 1$. Suponer valores iniciales para el vector de parámetros \mathbf{b} .
2. A partir de \mathbf{b} , calcular Φ con la ecuación 13, calculando previamente \mathbf{Y} con la ecuación 12: $\mathbf{Y} = f(\mathbf{X}, \mathbf{b})$.
3. Hacer $\lambda = 0.001$.
4. Evaluar la matriz \mathbf{A} con la ecuación 16, obtenida diferenciando \mathbf{Y} con respecto a \mathbf{b} .
5. Utilizar la ecuación 22 para calcular el vector corrección $\Delta\mathbf{b}$.
6. A partir de $\mathbf{b} + \Delta\mathbf{b}$, calcular Φ con la ecuación 13.
7. Si $\Phi(\mathbf{b} + \Delta\mathbf{b}) > \Phi(\mathbf{b})$ hacer $\lambda = 10\lambda$ y volver al paso 5.
8. Si no, hacer $\lambda = \lambda/10$.
9. Evaluar con la ecuación 18 la nueva estimación del vector de parámetros \mathbf{b} .
10. Evaluar $\Phi(\mathbf{b})$ con la ecuación 13.
11. Hacer $i = i + 1$. Terminar, si se ha agotado el número máximo de iteraciones. Si no, volver al paso 4 hasta que Φ no cambie o $\Delta\mathbf{b}$ se haga muy pequeño.

Diseño de la investigación

La investigación involucró la actualización y revisión de la literatura, la descripción general de la estimación no lineal, el análisis de la práctica común, la estimación lineal de los parámetros de siete casos de prueba, el

resumen de la teoría de la regresión lineal y no lineal, la proposición del procedimiento para hacer la estimación no lineal de los parámetros, la estimación no lineal de los parámetros de los siete casos de prueba, la contrastación de ambos tipos de estimación y la generación de conclusiones. De los siete casos de prueba estudiados se han escogido, para presentarlos en este trabajo, los cinco casos cuyas características son indicadas en el Cuadro I. El primer modelo puede

encontrarse, por ejemplo, en Miller *et al* (1974). El segundo modelo puede verse en muchos libros de Hidrología, uno de los cuales es el de Patra (2001). Los tres últimos pueden encontrarse, por ejemplo, en Martínez (2000). Los datos fueron recopilados en la literatura. Para el caso 1 (Miller *et al*, 1974), caso 2 (Patra, 2001), caso 3 (Linsley *et al*, 1977), caso 4 (Martínez, 1999) y caso 5 (Campos, 1984).

Cuadro I. Características de los cinco casos de prueba.

Caso	Modelo	Var(s) Indep(s)	Parámetros de ajuste	Descripción
1	$Q_{MMA} = \alpha A^\beta$	A	α, β	Gasto máximo medio anual en función del área de cuenca. Aplicable a diversas cuencas que están en una región homogénea.
2	$f = f_c + (f_0 - f_c)e^{-kt}$	t	f_0, f_c, k	Fórmula de la infiltración de Horton.
3	$Q = a(E - E_0)^b$	E	a, E ₀ , b	Curva de elevaciones-gastos de una corriente.
4	$V = KE^N$	E	K, N	Curva de elevaciones-capacidades de un embalse.
5	$i = \frac{aT_r^b}{d^c}$	T _r , d	a, b, c	Curva de intensidad de precipitación-duración-período de retorno.

Las variables elevadas en los casos prueba fueron varias, entre las que se encuentran: la suma de cuadrados, el coeficiente de determinación y la distribución de residuos.

Resultados

Después del estudio de la literatura se hicieron algunos ensayos hasta que se llegó a generar el siguiente procedimiento para efectuar la estimación no lineal de los parámetros de las ecuaciones:

1. De acuerdo con los datos que se quiere ajustar, identificar el tipo de ecuación que la experiencia previa ha encontrado que mejor se ajusta.
2. Estimar los valores iniciales de los parámetros.
3. Minimizar la suma de cuadrados de los errores. La minimización se puede hacer por uno de tres caminos: con un programa hecho a la medida, con una hoja de cálculo o con un programa orientado a la estadística. Es muy factible hacerlo por el segundo camino.
4. Calcular las desviaciones estándar de los parámetros. Investigar si éstos son estadísticamente significativos.

Construir los intervalos de confianza de los parámetros. Si algún parámetro no es significativo, podría ser necesario tratar de ajustar otro modelo o esperar a obtener más datos y volver a intentar el ajuste.

5. Medir la correlación entre los parámetros examinando la matriz de coeficientes de correlación. Si resulta que hay dos o más parámetros correlacionados entre sí, entonces, con los datos disponibles, será muy difícil establecer estimaciones separadas de ellos. O sea, es posible que los intervalos de confianza individuales no den la verdadera región de confianza conjunta de todos los parámetros, principalmente, en casos de alta no-linealidad.
6. Revisar si el modelo muestra evidencia de que hay relación entre la variable dependiente y, cuando menos, una de las variables independientes. Se puede hacer efectuando una prueba de análisis de varianza. Si, además, se quiere comparar los resultados de este ajuste con los obtenidos con regresión lineal, es conveniente utilizar el coeficiente de determinación r^2 .
7. Y, finalmente, realizar pruebas para verificar si los residuos entre los datos y el modelo se distribuyen aleatoriamente.

Cuadro 2. Datos y resultados de la curva de elevaciones-capacidades del Cedazo II (Caso 4)

i	E (m)	V (m ³)	V (m ³)	V (m ³)
		Original	Reg. lineal	Reg. no lineal
1	2.0	20041.0	13293.5	1706.9
2	4.0	57538.0	65154.4	18893.6
3	6.0	114504.0	165099.8	77104.9
4	8.0	215126.5	319336.9	209136.0
5	10.0	438769.0	532694.0	453480.7
6	12.0	838991.0	809192.6	853487.2
7	14.0	1466631.5	1152313.3	1456801.2
8	16.0	2314885.5	1565144.8	2314962.4

Ejemplo de aplicación

Un ejemplo que ilustra los resultados de ambos tipos de regresión es la curva de elevaciones capacidades (Caso 4) de la presa El Cedazo II (propuesta como anteproyecto, a varios kilómetros aguas arriba de la presa El Cedazo, localizada en el lado oriente de la mancha urbana de la Ciudad de Aguascalientes). Por razones de espacio sólo se mostrarán los pasos más importantes de los cálculos de los parámetros y de los coeficientes de determinación. Las elevaciones y los volúmenes originales se encuentran, respectivamente, en la segunda y tercera columna del cuadro 2.

Para efectuar la regresión lineal, se parte de la ecuación de elevaciones-capacidades (ver modelo 4 en el cuadro 1):

$$V = K E^N \quad (23)$$

que toma la forma lineal:

$$\log V = \log K + N \log E \quad (24)$$

Aplicando a los datos las transformaciones logarítmicas correspondientes y tratando a $\log V$ como y y a $\log E$ como x se obtienen: $\log K = 3.4333$, $N = 2.2931$ y $r_1^2 = 0.9618$. Al realizar la transformación de los resultados se encuentra que la relación entre las variables es:

$$V = 2712.27 E^{2.2931} \quad (25a)$$

Si con la ecuación 11 se calcula r_{nl}^2 en términos de la variable original, se encuentra que $r_{nl}^2 = 0.8555$, lo que, en principio, demuestra que la relación linealizada no es tan buena como se pudiera pensar al calcular r_1^2 con la variable transformada. La figura 1 muestra la curva ajustada. Se nota que los datos no están muy cerca de la curva, especialmente para valores grandes, y tampoco se distribuyen aleatoriamente por arriba y por debajo de ella.

La estimación de los parámetros con regresión no lineal es más complicada. Con el fin de ilustrar el procedimiento de obtención del valor mínimo de la suma

de cuadrados de los errores o residuos se utilizará el método de Marquardt, para el cual es necesario aplicar, una o más veces, en cada iteración, la ecuación matricial:

$$\Delta \mathbf{b} = (\mathbf{A}^T \mathbf{A} + \lambda \mathbf{D}^{-1})^{-1} \mathbf{A}^T (\mathbf{Y} - \mathbf{Y}) \quad (22)$$

Para formar la matriz A, de derivadas parciales, se obtienen las siguientes dos derivadas:

$$\frac{\partial V}{\partial K} = \frac{\partial (KE^N)}{\partial K} = E^N \quad \text{y} \quad \frac{\partial V}{\partial N} = \frac{\partial (KE^N)}{\partial N} = KE^N \ln E \quad (3')$$

Que se sustituirán en:

$$\mathbf{A} = \begin{bmatrix} \frac{\partial Y_1}{\partial b_1} & \dots & \frac{\partial Y_1}{\partial b_k} \\ \dots & \dots & \dots \\ \frac{\partial Y_n}{\partial b_1} & \dots & \frac{\partial Y_n}{\partial b_k} \end{bmatrix} \quad (16)$$

En este caso: $Y = V$, $b_1 = K$ y $b_2 = N$. Se forma el primer vector de parámetros, se elegirán los calculados con regresión lineal, o sea:

$$\mathbf{b} = \begin{bmatrix} 2712.27 \\ 2.2931 \end{bmatrix}$$

Con estos parámetros se calculan los valores de Y , necesarios para formar el vector de errores o residuos:

$$\mathbf{Y}' - \mathbf{Y} = \begin{bmatrix} 6747.9 \\ -7612.3 \\ -50582.5 \\ -104180.6 \\ -93869.9 \\ 29888.7 \\ 314454.7 \\ 749935.6 \end{bmatrix}$$

aplicando la ecuación matricial:

$$SSE_{nl} = \Phi = (\mathbf{Y}' - \mathbf{Y})^T (\mathbf{Y}' - \mathbf{Y}) \quad (13)$$

se encuentra que:

$$\Phi(\mathbf{b}) = 6.8451E + 11$$

Luego, sustituyendo los datos observados, se obtiene la matriz A de la primera iteración, formada por $n = 8$ renglones y $k = 2$ columnas (n juegos de datos y k parámetros):

$$\mathbf{A} = \begin{bmatrix} 4.9011 & 9214.0 \\ 24.0206 & 90317.5 \\ 60.8665 & 295795.3 \\ 117.7269 & 663980.4 \\ 196.3812 & 1226446.5 \\ 298.3119 & 2010543.7 \\ 424.8017 & 3040660.6 \\ 576.9890 & 4338962.6 \end{bmatrix}$$

Posteriormente se calcula la matriz $A^T A$:

$$A^T A = \begin{bmatrix} 659093.669 & 4734218446 \\ 4734218446 & 3.4155E + 13 \end{bmatrix}$$

La primera vez que se calcula el vector de corrección $\Delta \mathbf{b}$ se supone $\lambda = 0.001$. Si la nueva suma de los errores al cuadrado es mayor que la inicial, o sea, si $\Phi(\mathbf{b} + \Delta \mathbf{b}) > \Phi(\mathbf{b})$ se aumenta λ multiplicándola por 10; para este ejemplo, se encontró que era necesario llegar hasta $\lambda = 100000$ para que $\Phi(\mathbf{b} + \Delta \mathbf{b}) < \Phi(\mathbf{b})$. Por lo que la matriz $A^T A + \lambda I$ es:

$$A^T A + \lambda I = \begin{bmatrix} 659093.669 & 4734218446 \\ 4734218446 & 3.4155E + 13 \end{bmatrix} + 100000 \times$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 759093.669 & 4734218446 \\ 4734218446 & 3.4155E + 13 \end{bmatrix}$$

Enseguida, la matriz $A^T A + \lambda I$ se invierte, obteniéndose la matriz $(A^T A + \lambda I)^{-1}$:

$$(A^T A + \lambda I)^{-1} = \begin{bmatrix} 9.719E - 06 & -1.3472E - 09 \\ -1.3472E - 09 & 2.1601E - 13 \end{bmatrix}$$

Multiplicando la matriz A^T por el vector de residuos ($\mathbf{Y}^* - \mathbf{Y}$) se obtiene la matriz:

$$A^T (\mathbf{Y}^* - \mathbf{Y}) = \begin{bmatrix} 541273883.7 \\ 4.0703E + 12 \end{bmatrix}$$

Finalmente, el vector de corrección $\Delta \mathbf{b}$ se obtiene efectuando la operación: $\Delta \mathbf{b} = (A^T A + \lambda I)^{-1} A^T (\mathbf{Y}^* - \mathbf{Y})$. El resultado es:

$$\Delta \mathbf{b} = \begin{bmatrix} 9.719E - 06 & -1.3472E - 09 \\ -1.3472E - 09 & 2.1601E - 13 \end{bmatrix} \times \begin{bmatrix} 541273883.7 \\ 4.0703E + 12 \end{bmatrix} = \begin{bmatrix} -222.61 \\ 0.1500 \end{bmatrix}$$

El valor del vector de parámetros se expresará, provisionalmente, con

$$\mathbf{b} + \Delta \mathbf{b} = \begin{bmatrix} 2712.27 \\ 2.2931 \end{bmatrix} + \begin{bmatrix} -222.61 \\ 0.1500 \end{bmatrix} = \begin{bmatrix} 2489.66 \\ 2.4431 \end{bmatrix}$$

Antes de adoptarlo para el siguiente ciclo de iteraciones (en este ejemplo ya se sabe que éste será el nuevo vector \mathbf{b}), con este vector se calcula de nuevo, con la ecuación 13, el valor de la suma de cuadrados de los errores, o sea:

$$\Phi(\mathbf{b} + \Delta \mathbf{b}) = 1.9218E + 11$$

Como $\Phi(\mathbf{b} + \Delta \mathbf{b}) < \Phi(\mathbf{b})$ se puede adoptar como el nuevo vector de parámetros \mathbf{a} :

$$\mathbf{b} = \mathbf{b} + \Delta \mathbf{b} = \begin{bmatrix} 2489.66 \\ 2.4431 \end{bmatrix}$$

Se hace $\lambda = \lambda/10$ y se inicia un nuevo ciclo de corrección que debe comenzar en la determinación de la matriz A y en el que se repiten los cálculos que acaban de ser detallados. Recordar que dentro de un ciclo de corrección, cada vez que $\Phi(\mathbf{b} + \Delta \mathbf{b})$ es mayor que $\Phi(\mathbf{b})$ entonces se hace $\lambda = 10 \lambda$ y se regresa a calcular un nuevo valor del vector $\mathbf{b} + \Delta \mathbf{b}$ a partir de la matriz $A^T A + \lambda I$. El

proceso se da por terminado cuando Φ no cambia o $\Delta \mathbf{b}$ se hace muy pequeño. Así, después de varias decenas de iteraciones se encuentra que los parámetros que reportan un valor mínimo de la suma de errores al cuadrado son:

$$\mathbf{b} = \begin{bmatrix} 154.20 \\ 3.4685 \end{bmatrix}$$

O sea, $K = 154.20$, $N = 3.4685$ y la mínima suma de errores al cuadrado obtenida con regresión no lineal es:

$$\Phi(\mathbf{b}) = 3.7873E + 09$$

La ecuación de ajuste correspondiente es:

$$V = 154.20E^{3.4685} \quad (25b)$$

Finalmente, aplicando la ecuación 11 se obtiene el coeficiente de determinación $r_{nl}^2 = 0.9992$. Los volúmenes resultantes de este ajuste se pueden ver en la quinta columna del cuadro 2. La figura 1, a su vez, también muestra la curva ajustada. Los volúmenes observados, especialmente

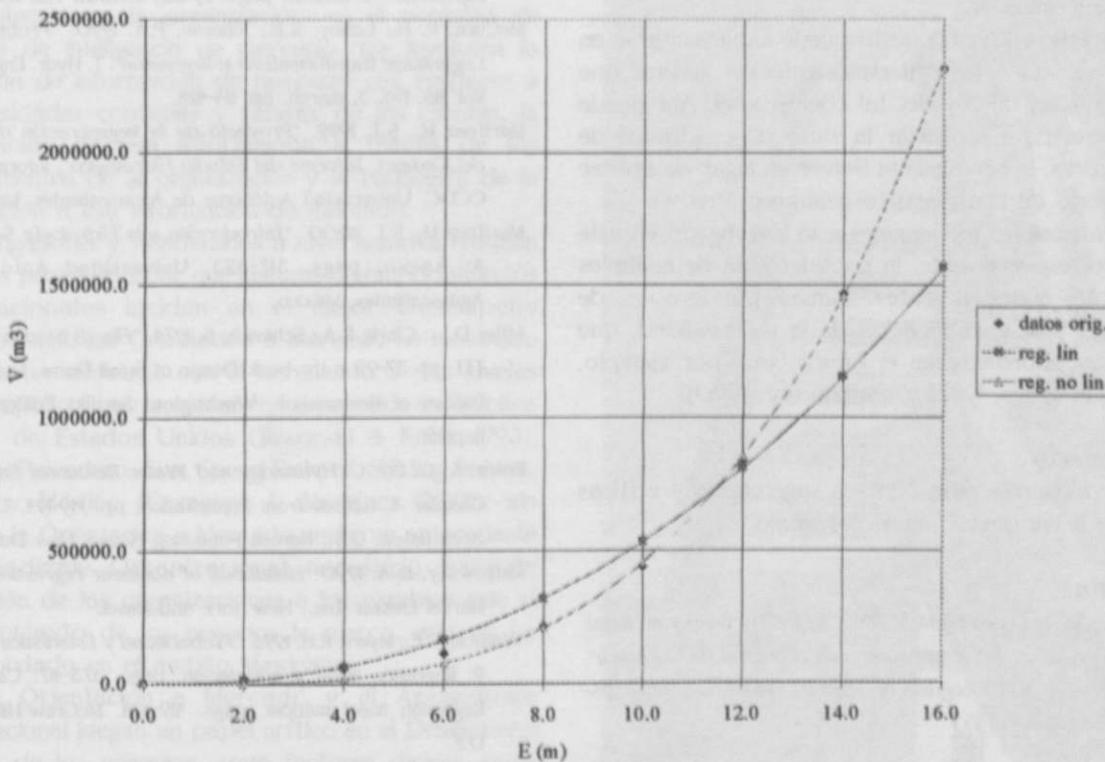
para valores grandes, se acercan más a la curva ajustada con regresión no lineal que a la curva obtenida con regresión lineal. La regresión no lineal es superior a la regresión lineal. El resultado del análisis de residuos se encuentra en el cuadro 3.

Resultados principales

En el cuadro 3 se indican los resultados principales de la contrastación de los análisis de regresión efectuados sobre los casos de prueba presentados. Los ajustes que mejor se comportaron fueron los realizados con la regresión no lineal. Se incluyen los coeficientes de determinación calculados con ambas regresiones.

Discusión

Los resultados confirman y extienden los obtenidos en estudios similares, como los de McCuen *et al* (1990) y Demetropoulos (1994). Aquí se presentaron cinco casos que incluyeron cuatro ecuaciones de forma diferente, se aplicaron pruebas estadísticas adicionales y se propuso un procedimiento sistemático de análisis. El procedimiento generado deja de lado la utilización de la regresión lineal pues ahora, con la computadora digital, se puede abordar la regresión no lineal que previamente no se utilizaba por requerir más trabajo computacional.



Cuadro 3. Resumen de la contrastación de ambas regresiones

Caso	r^2 lineal (var. orig.)	r^2 no lineal	Residuos	Observaciones
1	0.8182	0.8587	Aleatorios	
2	0.9818	0.9902	Aleatorios	
3	0.9971	0.9972	Aleatorios	Un par de datos produjo un residuo demasiado grande.
4	0.8555	0.9992	Aleatorios	El ajuste no lineal fue muy superior al lineal.
5	0.9979	0.9992	Aleatorios	Un parámetro no significativo.

Conclusiones

- La aportación más importante de este estudio es el procedimiento que se generó para estimar parámetros de ecuaciones hidrológicas; se requiere realizar un esfuerzo computacional mayor que el que comúnmente se hace pero, por utilizar la regresión no lineal en lugar de la regresión lineal, se obtienen mejores resultados.
- Se efectuó la estimación no lineal de los parámetros de cuatro tipos de ecuaciones muy comunes, comprobándose que los resultados arrojados son estadísticamente mejores que los obtenidos mediante regresión lineal.
- El procedimiento propuesto puede implementarse en un programa comercial de orientación general que maneje hojas de cálculo, tal como Excel. Ahí puede utilizarse, para minimizar la suma de cuadrados de los errores, la herramienta Solver en lugar de aplicar el método de Marquardt o cualquier otro.
- Existen posibles extensiones a lo investigado en este trabajo, especialmente, la consideración de múltiples variables independientes, sumas ponderadas de residuos y la cuantificación de la no-linealidad, que podrían abordarse en el futuro (ver, por ejemplo, [Constantinides, 1987] y [Ratkowsky, 1990]).

Agradecimiento

A los revisores, pues con sus sugerencias y críticas ayudaron a enriquecer nuestro trabajo.

Bibliografía

Berezowsky V., M.; Lara F., M.A. 1982. "Regresión lineal y no lineal: una breve discusión y aplicación a un problema de hidráulica". Memorias del VII Congreso Nacional de Hidráulica, págs. 95-107, AMH, México. D.F.

Campos A., D.F. 1984. "Procesos del ciclo hidrológico". Vol. II. Anexo C: Regresión y correlación lineales: conceptos y aplicaciones fundamentales, págs. C-1 -C-12, Universidad Autónoma de San Luis Potosí, México.

Constantinides, A., 1987. "Applied numerical methods with personal computers". Chapter 7: Linear and Nonlinear Regression Analysis, pp. 563-571, McGraw-Hill, New York, N.Y.

Demetracopoulus, A.C. 1994. "Nonlinear Regression Applied to Hydrologic Data". J. Irrig. and Drain. Engrg., ASCE, Vol. 120, No. 3, may-jun, pp. 652-659.

Linsley, R. K.; Kohler, M.A.; Paulhus, J.L.H. 1977. "Hidrología para ingenieros". 4: Caudal, págs. 89-122, McGraw Hill, México.

McCuen, R. H.; Leahy, R.B.; Johnson, P.A. 1990. "Problemas with Logarithmic transformations in Regression". J. Hydr. Engrg. ASCE, Vol. 116, No. 3, march, pp 414-428.

Martínez M., S.I. 1999. "Proyecto de la Regeneración del Arroyo del Cedazo. Informe del Estudio Hidrológico", informe interno, CCDC, Universidad Autónoma de Aguascalientes, México.

Martínez M., S.I. 2000. "Introducción a la Hidrología Superficial". A: Anexo, págs. 317-323, Universidad Autónoma de Aguascalientes, México.

Miller D.L.; Clark, R.A.; Schmach, S. 1974. "Floods Studies". Chapter III, pp 37-95 in the book Design of Small Dams, United States Bureau of Reclamation, Washington, Second Edition, Revised Reprint.

Patra, K. C. 2001. "Hydrology and Water Resources Engineering". Chapter 4. Losses from Precipitation, pp. 149-194, CRC Press, Boca Raton, U.S., Narosa Publishing House, New Delhi, India.

Ratkowsky, D.A. 1990 "Handbook of nonlinear regression models", Marcel Dekker Inc., New York and Basel.

Walpole, R.E.; Myers, R.H. 1996. "Probabilidad y Estadística". Capítulo 9: Regresión lineal y correlación, págs. 373-417; Capítulo 10: Regresión lineal múltiple, págs. 419-484. McGraw-Hill, México, D.F.