

Técnicas de Regresión e Interpolación Aplicadas a la Metrología

Kuruc, José

Centro Nacional de Metrología de Panamá (CENAMEP AIP)

Panamá, Panamá

ORCID: 0000-0001-5038-7766

Abstract

This work shows alternatives to calibration curves that are generally expressed as regressions [1], [2]. This is because when calibrating a measuring instrument, its results are expressed through a table that describes the errors and uncertainties [3]. However, many times the usage conditions of the equipment are slightly distant from the calibration conditions. This means it is very common to measure points that were not calibrated. This is where regressions become relevant, as they allow for interpolation at uncalibrated points within the calibration interval [3]. However, regressions do not always pass through the calibrated points, which in this case causes the exportation of errors through the regressed curve not to match the measured errors. Interpolation methods [4] are rarely used in metrology, but they are a potential alternative for exporting errors to unmeasured points without sacrificing the errors of the measured points. This is because the curves obtained through interpolation methods pass through all the measured points. Three interpolation methods are presented: the Lagrange method, the Newton method, and Splines [4]. It is proposed to compare these methods with regression curves through a case study of torque bench calibrations [5]. Here, the technical criteria that suggest which interpolation method is better and in which case it is convenient to decide between an interpolation case and a regression case are defined.

Keywords: Regressions, Lagrange Polynomial, Newton Polynomial, Splines, Calibration.

Resumen

Este trabajo muestra alternativas a las curvas de calibración que generalmente se expresan como regresiones [1], [2]. Esto es porque al calibrar un equipo de medición, sus resultados son expresados a través de una tabla en donde se describen los errores y las incertidumbres [3]. Sin embargo, muchas veces las condiciones de uso del equipamiento, distan un poco de las condiciones de calibración, esto quiere decir que es muy común que se midan puntos que no fueron calibrados. Allí es donde cobran relevancia las regresiones, que permiten

interpolación en puntos no calibrados, dentro del intervalo de calibración [3]. Sin embargo, las regresiones no siempre pasan por los puntos que fueron calibrados, lo que causa en este caso que la exportación de errores a través de la curva regresada, no coincida con los errores medidos. Los métodos de interpolación [4] son muy pocos utilizados en metrología, pero son una alternativa potencial para exportar los errores a puntos no medidos, sin sacrificar los errores de los puntos medidos. Esto es porque las curvas obtenidas a través de los métodos de interpolación pasan por todos los puntos medidos. Se presentan tres métodos de interpolación: el método de Lagrange, el método de Newton y los Splines [4]. Se propone comparar estos métodos con las curvas de regresión a través de un caso de calibraciones de bancadas de par torsional [5]. Aquí se definen los criterios técnicos que sugieren cuál método de interpolación es mejor y en que caso conviene decidir entre un caso de interpolación y uno de regresión.

Palabras claves: Regresiones, Polinomio de Lagrange, Polinomio de Newton, Splines, Calibración.

1. INTRODUCCIÓN

Se realizó un análisis comparativo entre las técnicas de regresión y las técnicas de interpolación para evaluar la capacidad de adaptación de estos recursos para expresar la mejor ecuación provenientes de un conjunto de datos de una calibración. Se desarrollaron los marcos teóricos de las regresiones simples, polinómicas, y los métodos de interpolación de Lagrange, Newton y Splines [4], [6]. Se abordó brevemente el teorema de unicidad y se esclareció porqué los métodos de Lagrange, Newton proporcionan el mismo polinomio para el mismo conjunto de datos. Se expuso además el nivel de dificultad y al mismo tiempo la eficacia del método de Splines para interpolar el conjunto de n datos con $n-1$ polinomios. Finalmente, se propusieron las mejores formas de evaluar la incertidumbre a través de estos métodos y que pudieran encajar lo mejor posible por lo propuesto en [3]. Se analizaron los resultados de una calibración de una bancada de par torsional y se sugirieron posibles desarrollos en trabajos futuros para la implementación de diferentes métodos para deducir curvas de calibración citados en diferentes contextos aplicables.

2. MÉTODO

A. Análisis de las Regresiones

Las regresiones son técnicas de modelado de datos cuyo objetivo es encontrar la mejor relación funcional entre una variable de interés que se quiere predecir (respuesta) [1], y una o más variables que influyen en la variable de interés (predictores) [1]. En este trabajo

se abordó el análisis a partir de regresiones cuyas variables de interés se ven afectadas por una variable influyente.

Tabla 1. Marco Teórico de las regresiones

• Regresión simple [2]	$y = b_0 + b_1x + \varepsilon$ (1)
	$b_0 = \frac{\sum_{i=1}^n y_i - b_1 \sum_{i=1}^n x_i}{n} = \bar{y} - b_1 \bar{x}, \quad b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$
	Dónde en (1): y: Es la variable respuesta. x: Es la variable predictor. b_0 : Es la constante de intercepto. b_1 : Es la constante de inclinación. ε : Es el error global de la regresión.
• Regresiones Polinómicas [2]	$y = b_0 + b_1x + b_2x^2 + \dots + b_kx^k + \varepsilon$ (2)
	$f(b_0, b_1, b_2, \dots, b_k) = \sum_{i=1}^n [y_i - (b_0 + b_1x_i + b_2x_i^2 + \dots + b_kx_i^k)]^2$ $b_0n + b_1 \sum x_i + b_2 \sum x_i^2 + \dots + b_k \sum x_i^k = \sum y_i$ $b_0 \sum x_i + b_1 \sum x_i^2 + b_2 \sum x_i^3 + \dots + b_k \sum x_i y_i = \sum x_i y_i$ $b_0 \sum x_i^k + b_1 \sum x_i^{k+1} + \dots + b_k \sum x_i^{2k} = \sum x_i^k y_i$
	Dónde en (2): y: Es la variable respuesta. x: Es la variable predictor. f: Es la función de regresión para estimación de regresores. $b_0, b_1, b_2, \dots, b_k$: es el grupo de constantes de la regresión.

B. Análisis de los métodos de interpolación

Los polinomios de interpolación son herramientas matemáticas que se utilizan para construir una función simple, en este contexto, un polinomio, que pase exactamente por un conjunto de puntos de datos discretos.

En el contexto de la calibración, si se tiene n puntos, la interpolación proporciona una ecuación $P(x)$, que permite estimar cualquier valor x entre los puntos medidos.

Los métodos de interpolación se rigen por el teorema de unicidad, que explica que para n puntos de datos, existe un único polinomio de grado n-1 que pasa por todos los puntos del conjunto de datos [4].

Tabla 2. Marco teórico de las interpolaciones.

<p>Polinomio de Lagrange: Como habíamos mencionado anteriormente, a partir del conjunto de datos, cada dato se define a través de la función f, entonces existe un polinomio $P(x)$ tal que $f(x_k) = P(x_k)$, para cada $k = 0, 1, \dots, n$</p> <p>Entonces el polinomio tendría la forma [4]:</p>	$P(x) = \sum_{k=0}^n f(x_k)L_{n,k}(x)$ <p>Dónde:</p> $L_{n,k} = \prod_{\substack{i=0 \\ i \neq k}}^n \frac{x-x_i}{x_k-x_i}$ <p>Es la base Lagrangiana del polinomio.</p>
<p>Polinomio de Newton [4]: El polinomio de Newton produce exactamente los mismos resultados que el polinomio de Lagrange, esto es posible por el teorema de unicidad mencionado anteriormente. Solo existe un único polinomio posible para un conjunto de datos que pueda describir la interpolación entre sus valores mínimos y máximos. Aplicando las diferencias divididas para deducir las constantes de interpolación del polinomio, obtenemos la forma del polinomio $P(x)$ como sigue:</p>	$a_0 = f[x_0] = P_n(x_0)$ $a_1 = f[x_i, x_{i+1}] = \frac{f[x_{i+1}] - f[x_i]}{x_{i+1} - x_i}$ $a_k = f[x_i, x_{i+1}, \dots, x_{i+k-1}, x_{i+k}] = \frac{f[x_i, x_{i+1}, \dots, x_{i+k}] - f[x_i, x_{i+1}, \dots, x_{i+k-1}]}{x_{i+k} - x_i}$ <p>Entonces:</p> $P_n(x) = f[x_0] + f[x_0, x_1](x - x_0) + \dots + f[x_i, x_{i+1}, \dots, x_{i+k-1}, x_{i+k}](x - x_0)(x - x_1) \dots (x - x_{k-1})$ $P_n(x) = a_0 + \sum_{k=1}^n a_k (x - x_0) \dots (x - x_{k-1})$
<p>Splines Cúbicos [4]: Un spline cúbico es una función que se construye a partir de una serie de polinomios de grado 3 unidos entre sí. Esto quiere decir que el spline rompe el problema en trozos más pequeños. Si se dispone de n puntos, el spline estaría compuesto por $n-1$ segmentos, donde cada segmento tiene su propia ecuación cúbica.</p> <p>La forma general de los splines sería la siguiente:</p>	$S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3$ $S(x) = \{S_1, S_2, \dots, S_{n-1}\}$

C. Análisis de la incertidumbre de las ecuaciones

Cuando se propone una curva de calibración como el mejor modelo de los datos entregados como resultados, esta curva de calibración aporta una componente de incertidumbre a la incertidumbre combinada de la calibración [3]. Las incertidumbres de los modelos que provienen de las regresiones ya cuentan con una forma estadística muy bien definida, basadas en el error global de la naturaleza de la regresión, sin embargo el tratamiento de las incertidumbres de las ecuaciones provenientes de algún método de interpolación, tiene la particularidad de que dependen de la dispersión individual de los datos medidos a la que hay que agregar el ruido de los efectos de interpolación. A continuación, se detallan los criterios a tomar en cuenta para expresar la incertidumbre de las ecuaciones (.

Tabla 3. Análisis de incertidumbre de las curvas regresadas o interpoladas.

<i>Incertidumbre de la regresión simple:</i> Depende de la dispersión entre el valor medido, y el predicho por la regresión con n-2 grados de libertad.	$u_{ecu} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - 2}}$
<i>Incertidumbre de la regresión polinómica:</i> Depende de la dispersión entre el valor medido y el predicho por la regresión con n-(k+1) grados de libertad.	$u_{ecu} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - (k + 1)}}$
<i>Incertidumbre del polinomio de Lagrange-Newton:</i> A diferencia de los métodos anteriores, las ecuaciones que se originan a partir de estos polinomios no tienen dispersiones en los puntos medidos debido a que la función del polinomio se define en el mismo punto de medición.	$u_{ecu} = \sqrt{\sum_{i=1}^n \left(\frac{\partial P_n}{\partial y_i} \right)^2 u_c^2}$ <p>Dónde:</p> $\frac{\partial P_n}{\partial y_i} = L_i(x)$ <p>Finalmente:</p> $u_{ecu} = \sqrt{\sum_{i=1}^n L_i^2 u_c^2}$
<i>Incertidumbre del polinomio Spline cúbico:</i> La expresión para la incertidumbre tiene un alto grado de complejidad. Un spline cúbico se compone por n-1 polinomios cúbicos $S_i(x)$, donde n es el número de puntos, cada segmento depende de un pequeño número de puntos vecinos y está acoplado al siguiente segmento en condiciones de continuidad. Conviene explorar métodos como Monte Carlo [7] para evitar ignorar analíticamente potenciales relaciones no lineales.	$u_{ecu} = \sqrt{\left(\frac{\partial S}{\partial y_i} \right)^2 u_c^2}$
<i>Incertidumbre de la calibración:</i> La incertidumbre de la calibración se expresa como la combinación de la incertidumbre de calibración, y la incertidumbre de la ecuación, culminando con su forma expandida (U) para un k=2 [3].	$u_{cal} = \sqrt{u_c^2 + u_{ecu}^2}$ $U = k u_{cal}$

D. Desarrollo de los algoritmos

El desarrollo de los algoritmos de regresión y de los métodos de interpolación se realizaron con el software R, por la cantidad de funciones disponible para análisis estadísticos y métodos numéricos.

3. RESULTADOS

El caso analizado proviene de los resultados obtenidos durante la validación del método de calibraciones de bancadas de par torsional en el CENAMEP AIP.



Fig.1 Bancada de par torsional.

Los resultados obtenidos durante la calibración se encuentran en la tabla 4. Estos resultados al ser analizados se compararon cual de las mejores curvas de calibración era la más adecuada entre una regresión simple y una regresión polinómica cúbica. El mejor ajuste se logró con una regresión cúbica centrada en el origen de la forma

$T = AT_m^3 + BT_m^2 + CT_m$, donde T es el par que se desea predecir en la interpolación y T_m es el par medido [5].

Tabla 4. Resultados de calibración de la bancada de par torsional del CENAMEP AIP hasta 1000 Nm.

Puntos de Calibración T (Nm)	Promedio T(Nm)	uc (Nm)
20	20.06	0.11
50	50.08	0.07
100	100.12	0.06
200	200.16	0.09
300	300.2	0.14
400	400.25	0.16
500	500.35	0.20
600	600.47	0.21
800	800.7	0.32
1000	1001.01	0.35

Posterior a la elección de la curva de ajuste de la calibración, los resultados de esta ecuación se compararon con con las ecuaciones obtenidas con el polinomio de Lagrange y Splines. En la figura 2, 3 y 4, se observan los contrastes de estos resultados.

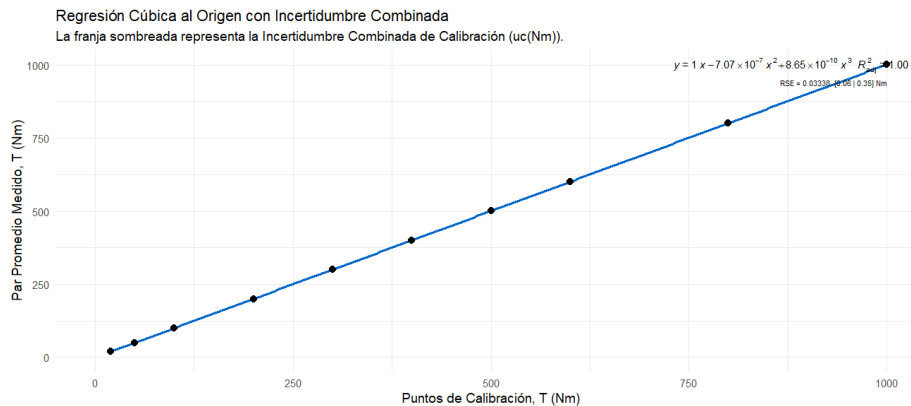


Fig.2 Regresión cúbica.

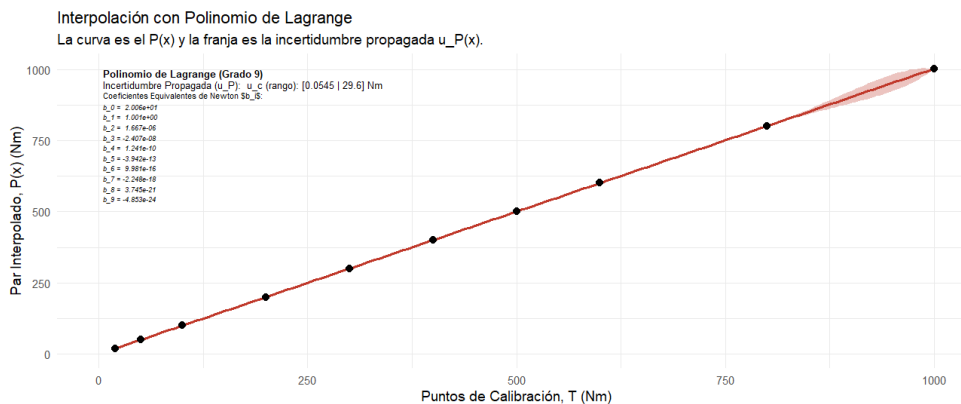


Fig.3 Interpolación de Lagrange [6].

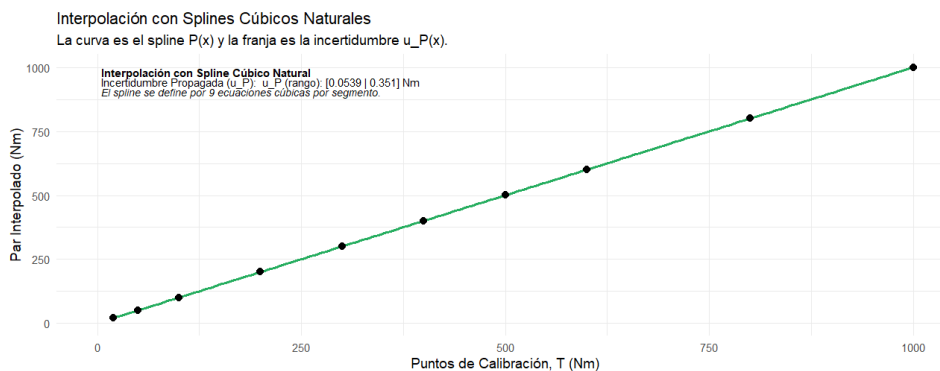


Fig.4 Interpolación por Spline

Nota: La interpolación por el método de spline requirió de 9 polinomios para llevarse a cabo.

4.CONCLUSIONES

El polinomio de Lagrange muestra su típica inestabilidad en los últimos puntos aumentando la incertidumbre radicalmente. Este fenómeno se da en la medida en que existan más datos por analizar.

La interpolación por spline en este contexto no tiene mucha significancia con respecto a la regresión, computacionalmente no es eficiente, sin embargo alcanza buenos resultados. Se propone adaptar este método de interpolación para un conjunto de datos que muestren mayor dispersión.

Las regresiones son muy confiables para bajas dispersiones y también cuando se disponen de más datos para describir las relaciones.

Los métodos de interpolación son muy poderosos para cuando se tienen resultados con muy pocos datos, como se pudo observar, no hay mucha diferencia entre un método y otro en los primeros puntos interpolados. Se propone evaluar estos resultados en calibraciones donde se apliquen curvas de calibración a través de regresiones con 5 datos o menos. En estos casos es común encontrar altas dispersiones.

Referencias

- [1] Jay L. Devore, *PROBABILIDAD Y ESTADÍSTICA PARA INGENIERÍA Y CIENCIAS*, 7a ed. CENGAGE Learning, 2008.
- [2] Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, y Keying Ye, *Probabilidad y estadística para ingeniería y ciencias*, 8a ed. PEARSON, 2007. [En línea]. Disponible en: www.pearsoneducacion.net/walpole
- [3] JCGM, "Evaluation of measurement data-Guide to the expression of uncertainty in measurement Évaluation des données de mesure-Guide pour l'expression de l'incertitude de mesure", 2008. [En línea]. Disponible en: www.bipm.org
- [4] Richard L. Burden, Douglas J. Faires, y Annette M. Durden, *ANÁLISIS NUMÉRICO*, 10a ed. CENGAGE Learning, 2016. Consultado: el 16 de noviembre de 2025. [En línea]. Disponible en: [//latinoamerica.cengage.com](http://latinoamerica.cengage.com)
- [5] D. Peschel, D. Röske, y en nombre de muchos otros colegas no mencionados aquí, "Physikalisch-Technische Bundesanstalt Directriz DKD-R 10-8 Calibración estática de instalaciones de calibración para llaves dinamométricas", 2020, doi: 10.7795/550.20200825B.
- [6] J. Committee for Guides in Metrology, "First edition 2008 Evaluation of measurement data-Supplement 1 to the 'Guide to the expression of uncertainty in measurement'-Propagation of distributions using a Monte Carlo method Évaluation des données de mesure-Supplément 1 du 'Guide pour l'expression de l'incertitude de mesure'-Propagation de distributions par une méthode de Monte Carlo", 2008.
- [7] W. Mora, *INTERPOLACIÓN POLINOMIAL*. con OOoBasic Y OOoCalc (VERSIÓN 1.0-AGOSTO 16, 2011.). *Revista Digital Matemática, Educación e Internet*, 2011. [En línea]. Disponible en: www.cidse.itcr.ac.cr/revistamate/

Autorización y Licencia CC

Los autores autorizan a APANAC XX a publicar el artículo en las actas de la conferencia en Acceso Abierto (Open Access) en diversos formatos digitales (PDF, HTML, EPUB) e integrarlos en diversas plataformas online como repositorios y bases de datos bajo la licencia CC:

Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

Ni APANAC XX ni los editores son responsables ni del contenido ni de las implicaciones de lo expresado en el artículo.