Formulas for Obtaining Intermediate Values of Biological Activity in QSAR using Lagrangian polynomial and Newton’s method

Nizam Uddin
M. B. Khalsa College, Rajmohalla, Indore - 452 002 (M.P.) India.
Email: nizamuddin4research@gmail.com, Contact Number: +919630775404

Received 28 September 2012; accepted 17 October 2012

Abstract
This research paper is based on obtaining intermediate values of the Biological Activity in QSAR. Biological Activity is depended on many parameters and coefficients. Increasing the values of parameters and coefficients in interval are increased the values of Biological Activity. Lagrangian Polynomial and Newton’s method for interpolation are applied on this relationship for obtaining intermediate values.

Keywords: Lagrangian polynomial, QSAR, Intermediate values, Biological Activity, Newton’s method.

1. Introduction:
Quantitative structure-activity relationships (QSAR) represent an attempt to correlate structural or property descriptors of compounds with activities. These physicochemical descriptors, which include parameters to account for hydrophobicity, topology, electronic properties, and steric effects, are determined empirically or, more recently, by computational methods. Activities used in QSAR include chemical measurements and biological assays [1]. Interpolation is the technique of obtaining the value of a function for any intermediate value of independent variable [2]. Mathematically, we can say that given a set of value of the function f(x) for certain value of the independent variable “x”, the method of finding the value f(x) for any given value of “x” is known as Interpolation thus the value of f(x) determined is known as interpolated[3]. Lagrangian polynomial is used for polynomial interpolation. For a given set of distinct points x_i and numbers y_j, the Lagrange polynomial is the polynomial of the least degree that at each point x_i assumes the corresponding value y_j [2,3,4]. This is called Lagrangian polynomial [5,6]. Newton’s method of interpolation are Newton’s forward interpolation formula, Newton’s backward interpolation formula, Newton’s divide difference interpolation formula [2,4].

2. Statistical Concepts in QSAR:
A QSAR attempts to find consistent relationships between the variations in the values of molecular properties and the biological activity for a series of compounds so that these "rules" can be used to evaluate new chemical entities.
A QSAR generally takes the form of a linear equation

Biological Activity = Constant + (C_1 \bullet P_1) + (C_2 \bullet P_2) + ... \quad (I)

Where the parameters P_i through P_n are computed for each molecule in the series and the coefficients C_i through C_n are calculated by fitting variations in the parameters and the biological activity [7].

If \( f(\text{CP}) = \text{Const} + (C_1 \bullet P_1) + (C_2 \bullet P_2) + (C_3 \bullet P_3) + ... \) \quad (suppose \( \text{CP}=X \))

Where “BA” is biological activity and X is variable from above function.

3. Lagrangian Polynomial:
Let \( y=f(x) \) be a function which assume “n” values \((x_1, y_1), (x_2, y_2), (x_3, y_3), ..., (x_n, y_n)\) corresponding to the argument \(x_1, x_2, x_3, ..., x_n\) not necessarily equally spaced [2,3,6]:

\[
Y(x) = \sum_{i=1}^{n} Y_i \prod_{j=1, j\neq i}^{n} \frac{(x-x_j)}{(x_i-x_j)}
\]

\[
(2)
\]

4. Formula for obtaining intermediate values:
When the values of “X” are change then the values of “BA” change, Increasing the values of “X” are increased the values of “BA”.
So we can write:

\((X_1, BA_1), (X_2, BA_2), (X_3, BA_3), .............(X_n, BA_n)\)
From equation (2), we got:

$$BA(X) = \sum_{i=1}^{n} BA_i \prod_{j=1}^{n} \frac{(X - X_j)}{(X_i - X_j)} \quad (3)$$

5. Newton’s method:
Newton’s method of interpolation are Newton’s forward interpolation formula, Newton’s backward interpolation formula, Newton’s divide difference interpolation formula [2,3,4].

5.1 Formula (i):
If the point lies in the upper half then we used Newton’s forward interpolation formula [2,3,4]. When $X_1, X_2, X_3, \ldots, X_n$ are equally spaced with interval “h”:

$$BA(X) = \frac{BA_0}{h} + \frac{\Delta BA_0}{h!} (X-X_1) + \frac{\Delta^2 BA_0}{2!} (X-X_1)(X-X_2) + \cdots + \frac{\Delta^n BA_0}{n!} (X-X_1)(X-X_2)\cdots(X-X_{n-1}) \quad (4)$$

[Where $\Delta$ is forward difference operator]

If we use this relationship

$$X - X_1 = u$$

[Where “u” is variable], Then

$X = u + X_1$.

$X - X_1 = h (u + 1)$.

Substituting these values ($X-X_1, X-X_2, \ldots, X-X_n$) in the equation (4), We got:

$$BA(X) = BA_0 + \frac{\Delta BA_0}{h} (u + 1) + \frac{\Delta^2 BA_0}{2!} (u + 1)(u + 2) + \cdots + \frac{\Delta^n BA_0}{n!} (u + 1)(u + 2)\cdots(u + n) \quad (5)$$

Simplifying, we got:

$$BA_0 (X) = BA_0 + h \Delta BA_0 \frac{X-X_1}{h} + \Delta^2 BA_0 \frac{(X-X_1)(X-X_2)}{2h} + \cdots + \frac{\Delta^n BA_0}{n!} (X-X_1)(X-X_2)\cdots(X-X_{n-1}) \quad (6)$$

[Where $\Delta_d$ is divide difference operator]

5.2 Formula (ii):
If the point lies in the lower half then we used Newton’s backward interpolation formula [2,3,4]. When $T_1, T_2, T_3, \ldots, T_n$ are equally spaced with interval “b”·

$$BA(X) = \frac{BA_n}{b} - \frac{\Delta BA_n}{b!} (X-X_n) + \frac{\Delta^2 BA_n}{2!} (X-X_n)(X-X_{n+1}) + \cdots + \frac{\Delta^n BA_n}{n!} (X-X_n)(X-X_{n+1})\cdots(X-X_{n-(n-1)}) \quad (7)$$

[Where $\nabla$ is backward difference operator]

If we use this relationship:

$$X - X_n = z$$

Then

$X = z + X_n$.

$X = X_n + h$.

$X = X_n + h$.

$X = X_n + h$.

$X = X_n + h$.

$X = X_n + h$.

$$X = X_n + h \quad (8)$$

Substituting these value ($X-X_n, X-X_{n+1}, \ldots, X-X_1$) in the equation (5), We got:

$$BA(X) = BA_n + \frac{\Delta BA_n}{b} (z) + \frac{\Delta^2 BA_n}{2!} (z)(z+1) + \cdots + \frac{\Delta^n BA_n}{n!} (z)(z+1)\cdots(z+(n-1)) \quad (9)$$

Simplifying, we got:

$$BA_0 (X) = BA_n + b \Delta BA_n \frac{X-X_n}{b} - \Delta^2 BA_n \frac{(X-X_n)(X-X_{n+1})}{2b} + \cdots - \frac{\Delta^n BA_n}{n!} (X-X_n)(X-X_{n+1})\cdots(X-X_{n-(n-1)}) \quad (10)$$

5.3 Formula (iii):
If $X_1, X_2, X_3, \ldots, X_n$ are not be equally spaced. We use Newton’s divide difference interpolation formula [2,3,4].

$$BA_0 (X) = BA_n + b \Delta BA_n \frac{X-X_n}{b} - \Delta^2 BA_n \frac{(X-X_n)(X-X_{n+1})}{2b} + \cdots - \frac{\Delta^n BA_n}{n!} (X-X_n)(X-X_{n+1})\cdots(X-X_{n-(n-1)}) \quad (11)$$

[Where $\Delta_d$ is divide difference operator]

6. Conclusion:
We obtained intermediate values of biological activity using formulas in the interval. When the values of $f(X)$ increased in interval then the values of biological activity are increased in interval. “n” is the last point and it is depend on number of interval. When we study of Biological Activity of chemicals or molecules in interval, these formulas is usefull to obtaining intermediate values.

7. Acknowledgments:
I would like to gratefully and sincerely thank Sayyad Maksud Ali, Mansur Ali, Abdul Ali and Isahaq Uddin Sheikh for continuous their support. I would like to sincerely thank Soumya V G. I would like to acknowledge Munavar Ali and Dr. Amir Ali (SAIMS- Indore).

References:
About the Author:
Nizam Uddin s/o Isahaq Uddin Sheikh is post graduate student of M. B. Khalsa college Indore, India. He passed higher secondary with Biology and Mathematics (additional subject) from Govt. Excellence higher secondary school Depalpur, Indore. He has passed B.Sc. in Bioinformatics from M. B. Khalsa College (Devi Ahilya Vishwavidyalaya) Indore, India. He is currently the student of M.Sc. in Bioinformatics from M. B. Khalsa college (Punjab Technical University) Indore, India. He is influence of Indian Islamic Scholar Muhammad Ilyas Kandhalvi's thought. He has command in Mathematics & statistics and their applications in Biological sciences. He is the author of many research papers in International Journals. His interests are Biostatistics, Biomathematics and solving complex problem in Biological Sciences using Mathematics.

Source of support: Nil; Conflict of interest: None declared