

BAB II

TINJAUAN PUSTAKA

Bab ini menguraikan teori-teori yang relevan sebagai landasan penelitian.

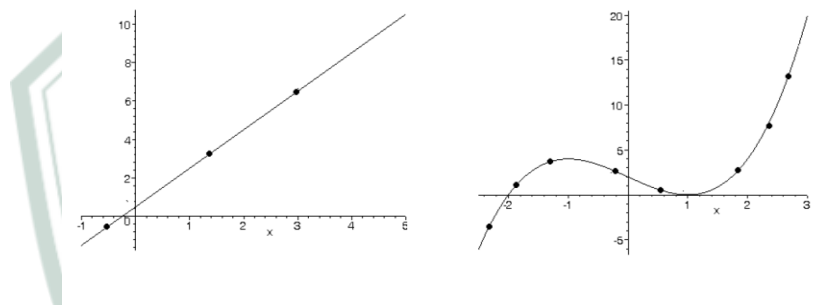
2.1 Imputasi

Missing data adalah kondisi ketika sebagian nilai dalam dataset tidak tersedia, tidak tercatat, atau hilang. *Missing data* dapat terjadi karena permasalahan sistem, seperti sensor yang tidak dapat menerima input. Kesalahan karena manusia juga dapat jadi penyebabnya, seperti kesalahan saat pengumpulan data atau jawaban responden yang tidak akurat. Apabila tidak ditangani dengan tepat, data yang hilang dapat menurunkan kualitas analisis dan menyebabkan hasil penelitian menjadi bias (Mukarromah et al., 2015).

Salah satu cara untuk menangani *missing data* adalah dengan imputasi. Imputasi adalah teknik statistik untuk menggantikan nilai data yang hilang dengan nilai estimasi tertentu pada data yang tidak lengkap atau tidak konsisten, baik dalam survei maupun studi longitudinal. Teknik ini bertujuan untuk mengurangi bias dan meningkatkan efisiensi analisis. Namun, imputasi juga memiliki keterbatasan, seperti kompleksitas analisis, potensi melebihi ketepatan estimasi, serta ketidakkonsistenan dalam penanganan nilai hilang pada subset data yang berbeda. Secara umum, teknik imputasi diklasifikasikan menjadi dua jenis, yaitu imputasi tunggal, dan imputasi berganda (Gladius Jennifer, 2016). Dalam konteks data runtun waktu atau data berurutan, pendekatan yang dapat digunakan dalam imputasi tunggal adalah interpolasi, yaitu metode yang memperkirakan nilai data hilang berdasarkan pola nilai data yang tersedia di sekitarnya. Oleh karena itu, pada subbab berikutnya akan dibahas interpolasi polinomial sebagai salah satu metode numerik yang digunakan dalam proses imputasi data hilang.

2.2 Interpolasi Polinomial

Interpolasi merupakan proses untuk menentukan atau menghitung nilai suatu fungsi berdasarkan grafik yang dibentuk dari sejumlah titik yang diketahui. Grafik tersebut harus melewati seluruh titik dengan tingkat ketelitian yang tinggi. Nilai-nilai fungsi tersebut biasanya diperoleh dari hasil percobaan atau pengamatan yang tidak dapat dilakukan setiap saat, melainkan hanya pada waktu-waktu tertentu. Oleh karena itu, untuk mengetahui nilai fungsi pada waktu selain saat pengamatan dilakukan, digunakan metode interpolasi. Teknik ini diterapkan ketika data memiliki tingkat ketelitian tinggi sehingga kurva yang dibentuk wajib melalui setiap titik data yang ada, seperti yang ditunjukkan pada grafik berikut yang menggambarkan dua bentuk kurva interpolasi polinomial dari dua jenis data yang berbeda.



Gambar 2.1 Grafik Interpolasi

Fungsi interpolasi yang paling umum digunakan adalah fungsi polinomial karena perhitungannya relatif sederhana. Sebuah polinomial dikatakan melakukan interpolasi terhadap sejumlah nilai apabila polinomial tersebut dapat digunakan untuk memperkirakan nilai y yang berhubungan dengan suatu x , di mana nilai x tersebut tidak terdapat dalam data hasil pengamatan, tetapi berada di antara nilai-nilai x yang telah diamati.

Berdasarkan Munir (2003), jika terdapat data sebanyak $n + 1$ titik (x_i, y_i) dengan $i = 0, 1, 2, \dots, n$, maka bentuk polinomialnya dapat dinyatakan sebagai berikut.

$$p_n(x) = a_0 + a_1x + \dots + a_nx^n, \quad a_n \neq 0 \quad (2.1)$$

dengan demikian, $p_n(x)$ merupakan polinomial dengan koefisien $a_0, a_1, \dots, a_n$, di mana suku berderajat tertinggi memiliki koefisien tidak sama dengan nol.

Dengan mensubstitusikan sebanyak $n + 1$ titik (x_i, y_i) ke dalam polinomial tersebut, akan didapat $n + 1$ persamaan dengan $n + 1$ variabel yang belum diketahui, yaitu sebagai berikut.

$$a_0 + a_1x_0 + \dots + a_nx_0^n = y_0$$

$$a_0 + a_1x_1 + \dots + a_nx_1^n = y_1$$

$$a_0 + a_1x_2 + \dots + a_nx_2^n = y_2$$

$$\vdots$$

$$a_0 + a_1x_n + \dots + a_nx_n^n = y_n$$

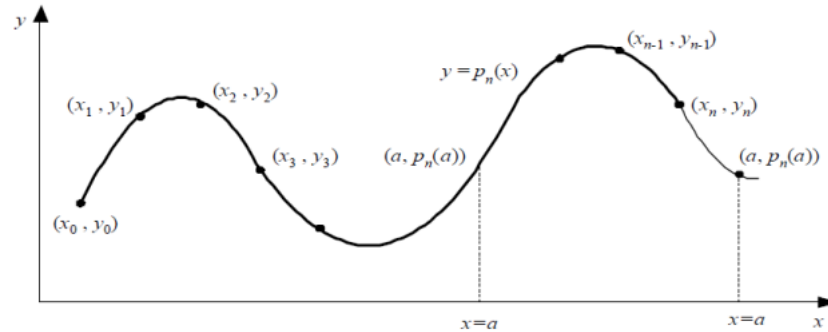
dengan syarat $a_n \neq 0$.

Fungsi $p_n(x)$ merupakan polinomial berderajat n yang menghasilkan $n + 1$ persamaan dengan $n + 1$ koefisien yang masing-masing perlu ditentukan.

Hasil penyelesaian dari sistem persamaan di atas adalah nilai-nilai a_i dengan $i = 0, 1, \dots, n$ yaitu koefisien-koefisien dari fungsi polinomial yang digunakan sebagai fungsi pendekatan. Dengan memasukkan nilai x sebagai variabel bebas ke dalam fungsi polinomial tersebut, akan diperoleh nilai y sebagai variabel terikat dari titik yang ingin ditentukan.

Setelah polinomial interpolasi $p_n(x_i)$ diperoleh, polinomial tersebut dapat digunakan untuk memperkirakan nilai y pada $x = a$, dengan rumus $y = p_n(a)$. Terdapat dua kemungkinan posisi nilai $x = a$, yaitu:

- (i) Jika $x = a$ berada di dalam rentang titik data ($x_0 < a < x_n$) maka hasil $y = p_n(a)$ disebut nilai interpolasi
- (ii) Jika $x = a$ berada di luar rentang titik data ($a < x_0$ atau $a > x_n$) maka hasil $y = p_n(a)$ disebut sebagai nilai ekstrapolasi.



Gambar 2.2 Grafik Interpolasi dan Ekstrapolasi

Grafik tersebut menggambarkan proses interpolasi dan ekstrapolasi. Nilai $y_i = p_n(x_i)$ disebut sebagai nilai interpolasi apabila titik $(a, p_n(a))$ berada didalam rentang titik-titik data, yaitu $x_0 < a < x_n$. Sedangkan $y_i = p_n(x_i)$ disebut sebagai nilai ekstrapolasi apabila titik $(a, p_n(a))$ yang terletak di luar rentang titik-titik data, yaitu ketika $a < x_0$ atau $a > x_n$.

Definisi 2.2.1. Sebuah polinomial $p_n(x)$ yang memiliki derajat kurang dari atau sama dengan n dalam bentuk baku merupakan fungsi yang dapat dinyatakan sebagai

$$p_n(x) = a_0 + a_1x + \dots + a_nx^n, \quad a_n \neq 0 \quad (2.2)$$

dengan $a_0, a_1, \dots, a_n$ merupakan koefisien-koefisien bilangan real. Apabila seluruh koefisien tersebut tidak bernilai nol, maka polinomial tersebut disebut memiliki derajat tepat n .

Teorema 2.2.1 Jika terdapat $n + 1$ pasangan titik

$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ dengan semua x_i berbeda, maka terdapat tepat satu polinomial $p_n(x)$ berderajat $\leq n$ yang melewati semua titik tersebut:

$$p_n(x_i) = y_i \text{ untuk } i = 0, 1, 2, \dots, n.$$

2.3 Interpolasi Lagrange

Seperti yang tertulis pada persamaan (2.2), bentuk persamaan polinomial berderajat satu, yaitu interpolasi dengan garis yang menghubungkan dua buah titik, disebut juga interpolasi linier, dinyatakan sebagai:

$$P_1(x) = a_0 + a_1x \quad (2.3)$$

Jika diberikan dua buah titik (x_0, y_0) dan (x_1, y_1) , maka terdapat fungsi polinomial berderajat satu $P_1(x)$ dengan koefisien a_0 dan a_1 . Kedua titik tersebut disubstitusikan ke dalam persamaan (2.3) sehingga diperoleh dua persamaan linear, yaitu:

$$y_0 = a_0 + a_1x_0,$$

$$y_1 = a_0 + a_1x_1.$$

Dari kedua persamaan tersebut diperoleh

$$a_1 = \frac{y_1 - y_0}{x_1 - x_0} \text{ dan } a_0 = \frac{x_1y_0 - x_0y_1}{x_1 - x_0}, x_0 \neq x_1.$$

Selanjutnya, nilai a_0 dan a_1 disubstitusikan ke persamaan (2.3), sehingga diperoleh:

$$P_1(x) = \frac{x_1y_0 - x_0y_1}{x_1 - x_0} + \frac{(y_1 - y_0)x}{x_1 - x_0}. \quad (2.4)$$

Atau dapat ditulis dalam bentuk lain:

$$P_1(x) = y_0 \frac{y_1 - y_0}{x_1 - x_0} (x - x_0). \quad (2.5)$$

Bentuk ini dapat dimanipulasi menjadi:

$$P_1(x) = y_0 \frac{x - x_1}{x_0 - x_1} + y_1 \frac{x - x_0}{x_1 - x_0}. \quad (2.6)$$

Jika didefinisikan:

$$L_0(x) = \frac{x-x_1}{x_0-x_1}, \quad L_1(x) = \frac{x-x_0}{x_1-x_0},$$

maka persamaan (2.6) dapat ditulis sebagai:

$$P_1(x) = y_0L_0(x) + y_1L_1(x) \quad (2.7)$$

yang disebut sebagai polinom Lagrange berderajat satu.

Secara umum, polinom Lagrange berderajat n dapat dinyatakan sebagai:

$$P_n(x) = \sum_{i=0}^n y_i L_i(x) = y_0L_0(x) + y_1L_1(x) + \cdots + y_nL_n(x), y_n \neq 0.$$

Jika $y_i = f(x_i)$, maka diperoleh:

$$\begin{aligned} P_n(x) &= \sum_{i=0}^n f(x_i)L_i(x) \\ &= f(x_0)L_0(x) + f(x_1)L_1(x) + \cdots + f(x_n)L_n(x), f(x_n) \neq 0 \end{aligned} \quad (2.8)$$

dengan

$$L_i(x) = \prod_{j=0, j \neq i}^n \frac{x-x_j}{x_i-x_j}. \quad (2.9)$$

Lebih lanjut, $P_n(x)$ menyatakan fungsi interpolasi polinomial Lagrange berderajat n , di mana $f(x_i)$ merupakan koefisien fungsi dan $L_i(x)$ disebut sebagai polinomial dasar Lagrange (*Lagrange basis polynomial*).

2.4 Interpolasi Newton

Interpolasi Newton merupakan salah satu metode interpolasi polinomial yang umum digunakan, terdiri atas dua bentuk utama, yaitu metode Newton beda terbagi dan metode Newton-Gregory. Metode Newton-Gregory merupakan kasus khusus dari metode Newton beda terbagi, di mana titik-titik data yang digunakan memiliki jarak yang sama (Aulia et al., 2020). Namun, dalam penelitian ini metode yang digunakan adalah metode Newton beda terbagi, yang selanjutnya disebut sebagai metode Newton.

Ide dasar interpolasi Newton sama dengan interpolasi Lagrange, yaitu mencari polinomial yang melalui sejumlah titik tertentu dengan meninjau kembali polinomial linear pada persamaan (2.6) yang dinyatakan dalam bentuk persamaan (2.8) yaitu

$$P_1(x) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0), x_0 \neq x_1.$$

Pada metode Newton, bentuk persamaan ini dapat ditulis sebagai:

$$P_1(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0). \quad (2.10)$$

Jika diberikan dua buah titik yaitu (x_0, y_0) dan (x_1, y_1) , maka terdapat fungsi polinomial berderajat satu yaitu $P_1(x)$ dengan $x_0 \neq x_1$.

Sama seperti pada interpolasi Lagrange, namakan:

$$a_0 = y_0 = f(x_0) \text{ dan } a_1 = \frac{y_1 - y_0}{x_1 - x_0} = \frac{f(x_1) - f(x_0)}{x_1 - x_0},$$

maka persamaan (2.10) dapat ditulis kembali menjadi:

$$P_1(x) = a_0 + a_1(x - x_0), \quad (2.11)$$

koefisien a_0 dan a_1 disebut sebagai koefisien polinomial Newton, di mana $a_0 = f(x_0)$ dan $a_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$ yang dalam penulisannya dapat ditulis menjadi $a_1 = f[x_1, x_0]$ karena persamaan tersebut merupakan persamaan dalam bentuk *divided difference*.

Selanjutnya, bentuk polinomial derajat dua, yaitu interpolasi dengan garis yang menghubungkan tiga buah titik atau dapat disebut dengan interpolasi kuadratik dapat ditulis sebagai:

$$P_2(x) = a_0 + a_1x + a_2x^2 \quad (2.12)$$

diberikan tiga buah titik (x_0, y_0) , (x_1, y_1) , dan (x_2, y_2) setelah disubstitusikan ke dalam persamaan (2.12) diperoleh sistem persamaan:

$$y_0 = a_0 + a_1x_0 + a_2x_0^2,$$

$$y_1 = a_0 + a_1x_1 + a_2x_1^2,$$

$$y_2 = a_0 + a_1x_2 + a_2x_2^2.$$

Dari ketiga persamaan tersebut diperoleh bentuk umum polinom Newton berderajat dua:

$$P_2(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) \quad (2.13)$$

atau dapat juga ditulis sebagai:

$$P_2(x) = P_1(x) + a_2(x - x_0)(x - x_1). \quad (2.14)$$

Dengan mensubstitusikan $x = x_2$, maka nilai a_2 dapat diperoleh:

$$a_2 = \frac{f(x_2) - a_0 - a_1(x_2 - x_0)}{(x_2 - x_0)(x_2 - x_1)}, \quad x_0 \neq x_1 \neq x_2. \quad (2.15)$$

Selanjutnya, dengan mensubstitusikan a_0 dan a_1 ke dalam persamaan (2.15), diperoleh bentuk beda terbagi tingkat dua sebagai berikut:

$$\begin{aligned} a_2 &= \frac{\frac{f(x_2) - f(x_1)}{x_2 - x_1} - \frac{f(x_1) - f(x_0)}{x_1 - x_0}}{x_2 - x_0} \\ &= \frac{f[x_2, x_1] - f[x_1, x_0]}{x_2 - x_0}. \end{aligned}$$

Berdasarkan bentuk tersebut, beda terbagi orde yang lebih tinggi didefinisikan secara rekursif. Untuk $n \geq 2$, beda terbagi orde ke- n dinyatakan sebagai:

$$f[x_n, x_{n-1}, \dots, x_1, x_0] = \frac{f[x_n, x_{n-1}, \dots, x_1] - f[x_{n-1}, x_{n-2}, \dots, x_0]}{x_n - x_0}.$$

Persamaan (2.14) menunjukkan bahwa $P_2(x)$ dapat dibentuk dari polinom sebelumnya yaitu $P_1(x)$. Dengan cara serupa, polinom Newton berderajat n dapat dibentuk secara rekursif dari polinom berderajat $n - 1$.

Seperti yang telah dijelaskan sebelumnya, a_1 merupakan koefisien yang dinyatakan dalam bentuk beda terbagi. Hal yang sama juga berlaku untuk

koefisien yang lainnya, yaitu $a_2, a_3, \dots, a_n$. Dengan demikian, seluruh koefisien a_i pada polinom Newton dapat dituliskan sebagai berikut:

$$\begin{aligned} a_0 &= f[x_0] \\ a_1 &= f[x_1, x_0] \\ a_2 &= f[x_2, x_1, x_0] \\ &\dots \\ a_n &= f[x_n, x_{n-1}, \dots, x_1, x_0]. \end{aligned}$$

Sehingga bentuk umum polinomial Newton berderajat n dapat dinyatakan secara rekursif:

$$P_n(x) = P_{n-1}(x) + a_n(x - x_0)(x - x_1) \dots (x - x_{n-1}), \quad (2.16)$$

dengan basis:

$$P_0(x) = f(x_0) = a_0,$$

maka, bentuk umum polinomial Newton beda terbagi dapat dinyatakan sebagai:

$$\begin{aligned} P_n(x) &= f(x_0) + (x - x_0)f[x_1, x_0] + \dots + (x - x_0)(x - x_1) \dots \\ &\quad (x - x_{n-1})f[x_n, x_{n-1}, \dots, x_1, x_0]. \end{aligned} \quad (2.17)$$