# ESTIMASI MODEL REGRESI NONPARAMETRIK DENGAN METODE B-SPLINE

**Oleh**
**Nur Salam[1)], Yuana Sukmawaty[2)], Annisa Halida[3)]**
**[1,2]Dosen Program Studi Statistika, FMIPA, Universitas Lambung Mangkurat.**
**[3]Mahasiswa Program Studi Statistika, FMIPA, Universitas Lambung Mangkurat.**
**Email : [1]n_salam@ulm.ac.id, [2]yuana_s@ulm.ac.id, [3] annisahalidaaa@gmail.com**

**Abstract**

Analisis regresi merupakan teknik statistik yang digunakan untuk menjelaskan hubungan antara variabel bebas (X) dengan variabel tak bebas (Y). Salah satu model dalam analisis regresi adalah model regresi nonparametrik dengan asumsi bentuk fungsi regresinya tidak diketahui. Untuk mengestimasi fungsi regresi yang tidak diketahui tersebut dapat dilakukan melalui metode pendekatan spline. Spline atau B-Spline adalah potongan-potongan polinomial, yang polinomial memiliki sifat tersegmen. Sifat tersegmen ini memberikan fleksibilitas lebih dari polinomial biasa, sehingga memungkinkan untuk menyesuaikan diri secara lebih efektif terhadap karakteristik dari suatu fungsi atau data. Penelitian ini bertujuan untuk mengestimasi model regresi nonparametrik dengan metode B-Spline. Metode penelitian ini menggunakan studi literatur dengan mengumpulkan semua bahan, baik itu buku, jurnal atau referensi lain yang menunjang dan relevan dengan materi yang akan dibahas dan diteliti. Hasil penelitian menunjukkan bahwa estimator dari fungsi regresi pada model regresi nonparametrik yang diperoleh dengan menggunakan metode pendekatan B-spline adalah : y ̂= S_λ y dengan S_λ = 〖〖B(B〗^T B+n λK^T)〗^(-1) B^T.

**Kata kunci : Analisis Regresi, Regresi Nonparametrik, B-Spline**

## PENDAHULUAN

Analisis regresi merupakan salah satu teknik statistik yang digunakan secara luas dalam ilmu pengetahuan terapan. Disamping digunakan untuk mengetahui bentuk hubungan antara dua variabel atau lebih, regresi juga dapat dipergunakan untuk peramalan. Variabel-variabel dalam regresi ada dua jenis yaitu variabel independen (variabel prediktor) dan variabel dependen ( variabel respon). (Suparti et al. 2018).

Ada beberapa pendekatan dalam analisis regresi yaitu pendekatan parametrik dan pendekatan nonparametrik. Pendekatan parametrik adalah pendekatan yang dilakukan jika asumsi bentuk fungsi regresinya diketahui dan bergantung pada parameter, sehingga mengestimasi fungsi regresinya sama dengan mengestimasi parameternya. Sedangkan pendekatan nonparametrik dilakukan jika asumsi bentuk fungsi regresinya tidak diketahui (Sudjana,2000).

Terdapat beberapa metode pendekatan pada regresi nonparametrik, salah satunya adalah metode B-Spline. B-Spline merupakan model regresi yang mempunyai interpretasi statistik dan visual sangat khusus dan sangat baik disamping itu juga mampu menangani karakter data yang mulus (smooth). B-Spline dalam regresi nonparametrik mempunyai sifat fleksibel dan mempunyai kemampuan mengestimasi perilaku data yang cenderung berbeda pada interval yang berlainan (Eubank, 1999 dan Budiantara, 2005). B-Spline merupakan potongan (truncated) polynomial tersegmen yang kontinu, sehingga memiliki kemampuan menyesuaikan diri lebih efektif dari karakteristik data. Metode B-Spline mempunyai keunggulan dalam mengatasi pola

ISSN No. 1978-3787(Cetak)
ISSN 2615-3505 (Online)

data yang menunjukkan naik atau turun yang tajam.

Berdasarkan uraian di atas, peneliti tertarik untuk mengkaji tentang estimasi model regresi nonparametrik, sehingga penelitian ini diberi judul Estimasi Model Regresi Nonparametrik dengan Metode B-Spline.

## LANDASAN TEORI

Sebelum membahas konsep estimasi dan normal asimtotik terlebih dahulu dibicarakan beberapa pengertian dasar yang merupakan konsep awal yang harus dipahami agar mudah mengikuti pembahasan yang dibicarakan.

**Teorema 1.2.1 (Barnes, 2006)**

Misalkan

$$\mathbf{Y = Ax}$$

dimana y adalah vektor m x 1, x adalah n x 1, A adalah matriks m x n dan A tidak bergantung pada x, maka

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = \boldsymbol{A}$$

**Teorema 1.2.2 (Barnes, 2006)**

Untuk kasus khusus dimana A merupakan matriks simetris dan $\alpha = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}$

Dimana x berukuran n x 1, A matriks berukuran n x n dan A tidak bergantung pada x, maka

$$\frac{\partial \alpha}{\partial \boldsymbol{x}} = \boldsymbol{2x^T A}$$

**Analisis Regresi**

Analisis regresi merupakan studi mengenai ketergantungan satu variabel tak bebas terhadap satu atau lebih variabel bebas dengan tujuan untuk mengestimasi atau memperkirakan nilai variabel tak bebas dari nilai variabel bebas yang diketahui.

Selain dapat menganalisis hubungan antar variabel, analisis regresi dapat dipergunakan dalam hal-hal sebagai berikut:

Menduga variabel yang tak bebas dengan menentukan berapa banyak variasi di dalam variabel tidak bebas yang dapat dijelaskan dengan variabel bebas.

Menentukan berapa banyak variasi di dalam variabel tidak bebas yang dapat dijelaskan dengan variabel bebas.

Menentukan apakah variabel bebas menjelaskan sebuah variasi signifikan di dalam variabel tidak bebas.

Menentukan struktur atau bentuk hubungan berupa persamaan matematika yang menghubungkan variabel–variabel bebas dan tidak bebas.(Gujarati, 2010)

1.2.2. Regresi Parametrik

Regresi parametrik merupakan metode statistik yang digunakan untuk mengetahui pola hubungan antara variabel prediktor dengan variabel respons, dengan asumsi bahwa telah diketahui bentuk fungsi regresinya. Hubungan antara variabel respons dan variabel prediktor dalam model dapat terjadi dengan fungsi linier maupun nonlinier dalam parameter (Draper dan Smith, 1996). Sedangkan secara umum bentuk model regresi linear berganda dengan k variabel prediktor diberikan oleh persamaan berikut:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \varepsilon_i \quad \text{atau}$$

$$y_i = \beta_0 + \textstyle\sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i \qquad \ldots(1.2.2.1)$$

dengan j = 1, 2, . . ., k, xij adalah nilai dari variabel prediktor ke-j untuk pengamatan ke-i, untuk i = 1, 2, . . ., n, atau dapat ditulis dalam bentuk matriks sebagai berikut:

$$y = XT\,\beta + \varepsilon \qquad \ldots(1.2.2.2)$$

Dengan, y adalah vektor dari variabel respons berukuran n × 1, XT merupakan matriks berukuran n × p , dengan p = k +1 dan β adalah vektor parameter yang akan diestimasi berukuran p × 1, ε adalah vektor error random berukuran n × 1 berdistribusi normal, independen dengan mean nol dan varians σ2. Secara lengkap matriks dan vektor–vektor tersebut diberikan oleh y, ε, β dan matriks X yaitu:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\mathbf{X}^T = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ 1 & x_{31} & x_{32} & \cdots & x_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \text{ dan } \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Untuk baris ke-i dari matriks XT dapat dinotasikan oleh XiT yang memuat angka 1 diikuti oleh nilai variabel prediktor dari pengamatan ke-j, yaitu : $\boldsymbol{X}_i^T = [1 \quad x_{i1} \quad \cdots \quad x_{ik}]$.

**Regresi Nonparametrik**

Regresi nonparametrik digunakan apabila bentuk pola hubungan antara variabel respons dengan variabel prediktor tidak diketahui bentuk fungsi regresinya. Dalam regresi nonparametrik kurva regresi hanya diasumsikan mulus (smooth) dalam arti termuat dalam suatu ruang fungsi tertentu sehingga mempunyai sifat fleksibilitas yang tinggi. Model regresi nonparametrik secara umum adalah sebagai berikut:

$y_i = f(z_i) + \varepsilon_i$, i = 1, 2, . . .,n. (2.3.2.1)

dengan, yi adalah variabel respons, zi merupakan variabel prediktor, f (zi) adalah fungsi regresi dengan bentuk kurva regresinya tidak diketahui dan εi adalah error random berdistribusi normal, dengan mean nol dan varians σ2.

Jika diberikan matriks berikut :

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad f(\mathbf{Z}) = \begin{bmatrix} f(z_1) \\ f(z_2) \\ \vdots \\ f(z_n) \end{bmatrix} \text{ dan}$$

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

maka model regresi pada persamaan (2.3.2.1) dapat dituliskan dalam bentuk matriks sebagai berikut:

y = f (Z ) + ε (Hardle, 1994) …(2.3.2.2)

**Estimasi**

Estimasi merupakan pendugaan terhadap suatu parameter populasi dimana parameter tersebut terdapat dalam suatu model percobaan (Wibisono,2005).

Estimasi parameter

Parameter adalah sembarang nilai yang menjelaskan ciri populasi. Estimasi parameter adalah estimasi yang digunakan untuk suatu populasi dari sampel (Wibisono,2005).

**Penalized Least Square**

Penalized Least Square adalah cara lain untuk mendapatkan estimasi model regresi. Secara umum Penalized Least Square diberikan oleh:

Iλ(f) = R(f) + λ J(f) …(2.4.2.1)

dimana fungsional Iλ(f) memuat tiga komponen, yaitu komponen Least Square R(f), Roughness Penalty J(f) yakni ukuran kemulusan dan kekasaran kurva dan memetakan data dan parameter penghalus λ, dimana:

$R(f) = n^{-1}\sum_{i=1}^{n}(y_i - f(z_i))^2$ …(2.4.2.2)

$J(f) = \int_a^b [f''(z)]^2 dt$ ….(2.4.2.3)

Sehingga, jika persamaan (2.4.2.2) dan persamaan (2.4.2.3) disubtitusikan dalam persamaan (2.4), maka persamaan untuk Penalized Least Square dapat ditulis menjadi:

ISSN No. 1978-3787(Cetak)
ISSN 2615-3505 (Online)

$$I_\lambda(f) = n^{-1}\sum_{i=1}^{n}\left(y_i - f(z_i)\right)^2 + \lambda\int_a^b [f''(z)]^2 dt$$

. …(2.4.2.4)

Parameter penghalus λ merupakan konstanta positif yang bernilai 0 < λ < 1, yang berfungsi untuk mengontrol keseimbangan data dan kemulusan kurva, oleh karena itu pemilihan nilai λ merupakan suatu hal yang sangat penting. Ada beberapa metode untuk menentukan nilai optimal dari λ salah satunya adalah metode Generalized Cross Validation (GCV), dengan persamaan:

$$GCV(\lambda) = \frac{MSE\ (\lambda)}{\{1-n^{-1}tr(S_\lambda)\}^2}$$

…(2.4.2.5)

Dimana tr(Sλ) adalah trace matriks Sλ = (I + λK)-1 , K matriks yang bergantung pada X1,X2,…,Xn dan I matriks identitas (Tripena,2011).

## METODE PENELITIAN

Adapun prosedur-prosedur yang dilakukan dalam penelitian ini adalah sebagai berikut :

1. Mengumpulkan bahan-bahan penelitian yang berhubungan dengan regresi nonparametrik.
2. Mempelajari bahan-bahan yang telah dikumpulkan pada point (1) di atas.
3. Menjelaskan model regresi nonparametrik dan bentuk modelnya dalam bentuk matriks.
4. Menentukan metode estimasi estimasi dan fungsi regresi nonparametrik dengan metode B-spline.
5. Membuat simulasi model estimasi estimasi dan fungsi regresi nonparametrik dengan metode B-Spline.
6. Mengaplikasikan hasil estimasi dalam suatu contoh soal ataupun contoh kasus yang ada wilayah di Kalimantan Selatan.
7. Menarik kesimpulan dari hasil pembahasan.

## HASIL DAN PEMBAHASAN

Mengestimasi fungsi regresi nonparametrik dilakukan agar memperoleh model yang sesuai dan mendekati model sebenarnya. Ada beberapa tekknik estimasi dalam regresi nonparametrik antara lain pendekatan histogram, estimasi kernel, estimasi deret orthogonal, analisis *wavelet* dan estimasi *spline*.

### Fungsi – Fungsi Pendekatan

Pada regresi nonparametrik, bentuk kurva regresinya tidak diketahui, untuk itu dilakukan dengan pendekatan spline. *Spline* adalah potongan-potongan polinomial, yang polinomial memiliki sifat tersegmen. Sifat tersegmen ini memberikan fleksibilitas lebih dari polinomial biasa, sehingga memungkinkan untuk menyesuaikan diri secara lebih efektif terhadap karakteristik dari suatu fungsi atau data.

Beberapa jenis *Spline* yang dapat digunakan dalam regresi nonparametrik, yang dikenal sebagai fungsi basis, salah satunya adalah basis *B-Spline* (Suhartono,2009). Jika fungsi regresi didekati dengan *B-Spline* maka dapat ditulis menjadi (Budiantara dkk, 2006).

$$h(z) = \sum_{l=1}^{m+k} \alpha_l\, B_{l-m,m}\,(z)$$

dimana $B_{l-m,m}(z)$ adalah basis dengan orde ke- (l-m) dari titik knot a < u1, u2, . . ., u, K < b, dengan derajat m, dan 1,2, . . .,$\alpha, \alpha_{m+k}$ adalah parameter.

Fungsi B-Spline dengan derajat m secara rekursif didefinisikan sebagai berikut:

untuk m > 1:

$$B_{j,m}(z) = \frac{z-u_j}{u_{j+m}-u_j}B_{j,m-1}(z) + \frac{u_{j+m}-z}{u_{j+m}-u_{j+1}}B_{j+1,m-1}(z)$$

untuk m =1

$$B_{j,1}(z) = \begin{cases} 1, & \text{jika } u \le z \le u_{j+1} \\ 0, & \text{lainnya} \end{cases}$$

Untuk mengestimasi parameter α pada persamaan (3.1.1), didefinisikan matriks

B(λ) = (Bj,m(zi)) … (3.1.2)

dengan i = 1, 2, …, n dan j =(1-m), …, K, atau dapat ditulis dalam bentuk matriks berukuran n x (m + K).

$$\boldsymbol{B}(\boldsymbol{\lambda}) = \begin{bmatrix} \boldsymbol{B}_{(1-m),m(z_1)} & \cdots & \boldsymbol{B}_{k,m}(z_1) \\ \vdots & \ddots & \vdots \\ \boldsymbol{B}_{(1-m),m(z_n)} & \cdots & \boldsymbol{B}_{k,m}(z_n) \end{bmatrix}$$

**Model Regresi Nonparametrik**

Regresi nonparametrik adalah metode yang digunakan untuk mengetahui pola hubungan antara variabel bebas yang tidak diketahui dengan mengasumsikan tentang bentuk fungsinya.

Model regresi nonparametrik sebagai berikut :

$$y_i = f(z_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad \text{...(3.2.1)}$$

dan $\varepsilon = y - f(Z)$ ... (3.2.2)

Least Square merupakan suatu metode yang meminimukan jumlah kuadrat error. Least Square dapat dinyatakan pada persamaan (2.4.2.2). Berdasarkan persamaan (3.2.1), maka persamaan (2.4.2.2) dapat dinyatakan dengan :

$$\boldsymbol{R}(\boldsymbol{f}) = n^{-1}\sum_{i=1}^{n} \varepsilon_i^{\,2}$$

$$\boldsymbol{R}(\boldsymbol{f}) = n^{-1}\varepsilon_1^{\,2} + n^{-1}\varepsilon_2^{\,2} + \cdots + n^{-1}\varepsilon_n^{\,2}$$

$$\boldsymbol{R}(\boldsymbol{f}) = n^{-1}\varepsilon_1\varepsilon_1 + n^{-1}\varepsilon_2\varepsilon_2 + \cdots + n^{-1}\varepsilon_n\varepsilon_n$$

$$\boldsymbol{R}(\boldsymbol{f}) = n^{-1}(\varepsilon_1\varepsilon_1 + \varepsilon_2\varepsilon_2 + \cdots + \varepsilon_n\varepsilon_n)$$

(3.2.3)

Jika diberikan matriks $\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$ dan $\boldsymbol{\varepsilon}^{\boldsymbol{T}} = [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n]$ maka persamaan (3.2.2) di atas dapat dinyatakan dalam bentuk :

$$\boldsymbol{R}(\boldsymbol{f}) = n^{-1}\left( [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n] \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \right)$$

Sehingga, persamaan (3.2.3) dapat dituliskan dalam bentuk :

$$\boldsymbol{R}(\boldsymbol{f}) = n^{-1}\boldsymbol{\varepsilon}^{\boldsymbol{T}}\boldsymbol{\varepsilon} \quad \dots(3.2.3)$$

Substitusi persamaan (3.2.2) ke dalam persamaan (3.2.3) di peroleh (Eubank, 1999)

$$\boldsymbol{R}(\boldsymbol{f}) = n^{-1}(\boldsymbol{y} - \boldsymbol{f})^{\boldsymbol{T}}(\boldsymbol{y} - \boldsymbol{f})$$

$$\boldsymbol{R}(\boldsymbol{f}) = n^{-1}(\boldsymbol{y}^{\boldsymbol{T}} - \boldsymbol{f}^{\boldsymbol{T}})(\boldsymbol{y} - \boldsymbol{f})$$

$$\boldsymbol{R}(\boldsymbol{f}) = n^{-1}(\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{y} - \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{f} + -\boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{y} + \boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{f})$$

…(3.2.4)

Hasil $\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{f}$ akan berbentuk skalar atau matriks 1x1. Berdasarkan sifat-sifat dari transpose suatu matriks, maka $(\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{f})^{\boldsymbol{T}} = \boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{y}$. Karena $\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{f}$ adalah skalar atau maka transpose dari skalar atau matriks 1x1 akan menghasilkan skalar. Hal ini berakibat $\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{f} = \boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{y}$ bernilai sama, maka dengan mengambil salah satu bentuknya, persamaan (3.2.4) dapat dinyatakan menjadi:

$$\boldsymbol{R}(\boldsymbol{f}) = n^{-1}(\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{y} - \boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{f} - \boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{y} + +\boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{f})$$

$$\boldsymbol{R}(\boldsymbol{f}) = n^{-1}(\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{y} - 2\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{f} + \boldsymbol{f}^{\boldsymbol{T}}\boldsymbol{f}) \quad \text{... (3.2.5)}$$

**Roughness Penalty**

Misalkan diberikan sebuah fungsi f(z), yang bentuk kurva dari fungsi tersebut tidak diketahui. Karena fungsinya tidak diketahui, maka dilakukan suatu pendekatan fungsi. Salah satu fungsi pendekatannya adalah basis Spline yaitu B-Spline. Jadi, jika fungsi regresi f didekati dengan B-Spline, maka f dapat di tulis dalam bentuk persamaan (3.1.1). Jika terdapat K titik knot dengan B-Spline berderajat m, maka persamaan (3.1.1) dapat dijabarkan menjadi:

$$f(z) = \alpha_1 B_{1\text{-m},m}(z) + \alpha_2 B_{2\text{-m},m}(z) + \cdots + \alpha_{m+k} B_{m+k\text{-m},m}(z)$$

$$= \alpha_1 B_{(1\text{-m}),m}(z) + \alpha_2 B_{(2\text{-m}),m}(z) + \cdots + \alpha_{m+k} B_{k,m}(z)$$

$$= [\, B_{(1\text{-m}),m}(z) \quad B_{(2\text{-m}),m}(z) \quad \cdots \quad B_{k,m}(z)] \begin{bmatrix} \alpha 1 \\ \alpha 2 \\ \vdots \\ \alpha_{m+k} \end{bmatrix}$$

$$f(z) = \boldsymbol{B\alpha} \quad \dots(3.3.1)$$

Untuk pengamatan sebanyak i = 1, 2, …, n maka

$$f(z_1) = \alpha_1 B_{(1-m),m}(z_1) + \alpha_2 B_{(2-m),,m}(z_1) + \cdots + \alpha_{m+K} B_{K,m}(z_1)$$

$$f(z_2) = \alpha_1 B_{(2-m),m}(z_2) + \alpha_2 B_{(2-m),,m}(z_2) + \cdots + \alpha_{m+K} B_{K,m}(z_2)$$

$$\vdots$$

$$f(z_n) = \alpha_1 B_{(1-m),m}(z_n) + \alpha_2 B_{(2-m),m}(z_n) + \cdots + \alpha_{m+K} B_{K,m}(z_n)$$

…(3.3.2)

Dengan mendefinisikan bahwa $B_{j,m}(z_i) = B_{ij}$ dan $f(z_i) = f_i$, untuk

ISSN No. 1978-3787(Cetak)
ISSN 2615-3505 (Online)

pengamatan sebanyak i = 1, 2, …, n dan j = (1-m), (2-m), …, K maka persamaan (3.3.2) di atas dapat dinyatakan sebagai :

$$f_1 = \alpha_1 B_{11} + \alpha_2 B_{12} + \cdots + \alpha_{m+K} B_{1v}$$
$$f_2 = \alpha_1 B_{21} + \alpha_2 B_{22} + \cdots + \alpha_{m+K} B_{2v}$$
$$\vdots$$
$$f_n = \alpha_1 B_{n1} + \alpha_2 B_{n2} + \cdots + \alpha_{m+K} B_{nv} \quad \ldots(3.3.3)$$

dengan

$B_{i1} = B_{(1-m),m}(z_i), B_{i2} = B_{(2-m),m}(z_i), \ldots, B_{iv} = B_{K,m}(z_i)$

Jika diberikan matriks-matriks berikut

$$\boldsymbol{f} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} \quad \boldsymbol{B} = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1v} \\ B_{21} & B_{22} & \cdots & B_{2v} \\ \vdots & \vdots & \ddots & \vdots \\ B_{n1} & B_{n2} & \cdots & B_{nv} \end{bmatrix}$$
$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{m+K} \end{bmatrix}$$

maka persamaan (3.3.3) di atas dapat dinyatakan dalam bentuk matriks :

$$\boldsymbol{f} = \boldsymbol{B}\ \boldsymbol{\alpha} \quad \ldots(3.3.4)$$

Menurut Gyorfy(2002), Roughness Penalty dapat dinyatakan sebagai integral dari kuadrat kedua suatu fungsi, atau dapat ditulis dengan $\alpha$ :

$$= \int_a^b [f''(z)]^2 dz \quad \ldots(3.3.5)$$

dengan mensubstitusi persamaan (3.3.1) ke (5.13), maka diperoleh :

$$J(z) = \int_a^b [f''(z)]^2 dz$$
$$= \int_a^b [D^2 f(z)]^2 dz$$
$$= \int_a^b [D^2 f(z)][D^2 f(z)] dz$$
$$= \int_a^b [D^2 \boldsymbol{B\alpha}][D^2 \boldsymbol{B\alpha}] dz \quad \ldots(3.3.6)$$

Karena hasil dari $\boldsymbol{B\alpha}$ adalah skalar atau matriks 1x1, maka berdasarkan sifat-sifat transpose suatu matriks maka $(\boldsymbol{B\alpha})^T = \boldsymbol{\alpha^T B^T}$ akan menghasilkan hasil yang sama, yakni berupa scalar atau matriks 1x1. Akibatnya persamaan (3.3.6) di atas dapat dituliskan dengan

$$J(f) = \int_a^b [D^2 \boldsymbol{\alpha^T B^T}][D^2 \boldsymbol{B\alpha}] dz$$
$$= \alpha^T \int_a^b [D^2 \boldsymbol{B^T}][D^2 \boldsymbol{B}] dz\ \boldsymbol{\alpha}$$
$$= \boldsymbol{\alpha^T K \alpha}$$

...(3.3.7)

dengan

$$K = \int_a^b [D^2 \boldsymbol{B}][D^2 \boldsymbol{B^T}] dz$$

**Penalized Least Square**

Penalized Least Square adalah cara lain untuk mendapatkan estimasi model regresi. Secara umum Penalized Least Square diberikan oleh:

Iλ(f) = R(f) + λ J(f) … (3.4.1)

dimana R(f) merupakan komponen Least Square, J(f) merupakan Roughness Penalty dan $\lambda \geq 0$ merupakan parameter smoothing. Dengan mensubtitusi persamaan (3.2.3) dan (3.3.5) ke dalam persamaan (3.4.1) di atas maka fungsi Penalized Least Square dapat dinyatakan sebagai berikut:

$$I_\lambda(f) = n^{-1} \sum_{i=1}^{n} (y_i - f(z_i))^2 + \lambda \int_a^b [f''(z)]^2 dz \quad \ldots(3.4.2)$$

Kemudian dengan mensubtitusi persamaan (3.4.1) dan (3.3.7) ke persamaan (3.4.2), maka diperoleh:

$$I_\lambda(f) = n^{-1}(\boldsymbol{y^T y} - 2\boldsymbol{y^T f} + \boldsymbol{f^T f}) + \lambda \boldsymbol{\alpha^T K \alpha}$$

...(3.4.3)

Karena fungsi regresi didekatkan menggunakan fungsi pendekatan B-Spline, dimana vektor fungsi yang dinyatakan pada persamaan (3.3.4), maka subtitusi persamaan (3.3.5) tersebut dalam persamaan (3.4.3) di atas menjadi:

$$I_\lambda(f) = n^{-1}(\boldsymbol{y}^T\boldsymbol{y} - 2\boldsymbol{y}^T\boldsymbol{B\alpha} + (\boldsymbol{B\alpha})^T\boldsymbol{B\alpha}) + \lambda\boldsymbol{\alpha}^T\boldsymbol{K\alpha}$$

.

$$I_\lambda(f) = n^{-1}(\boldsymbol{y}^T\boldsymbol{y} - 2\boldsymbol{y}^T\boldsymbol{B\alpha} + \boldsymbol{\alpha}^T\boldsymbol{B}^T\boldsymbol{B\alpha}) + \lambda\boldsymbol{\alpha}^T\boldsymbol{K\alpha}. \quad \dots(3.4.4)$$

**Estimasi Model Regresi Nonparametrik**

Untuk mendapatkan estimasi dari model regresi nonparametrik dengan cara mengestimasi fungsi pada komponen nonparametrik. Fungsi regresi f(t) dinyatakan dengan pendekatan B-Spline, maka untuk meminimumkan fungsi Penalized Least Square pada persamaan (3.4.4) sama halnya dengan menurunkan $I_\lambda$(f) terhadap $\boldsymbol{\alpha}$. Kemudian menyamakan hasilnya dengan nol.

$$\frac{\partial I_\lambda(f)}{\partial \boldsymbol{\alpha}} = 0$$

$$\frac{\partial(n^{-1}(\boldsymbol{y}^T\boldsymbol{y} - 2\boldsymbol{y}^T\boldsymbol{B\alpha} + \boldsymbol{\alpha}^T\boldsymbol{B}^T\boldsymbol{B\alpha}) + \lambda\boldsymbol{\alpha}^T\boldsymbol{K\alpha})}{\partial\boldsymbol{\alpha}} = 0$$

$$n^{-1}\frac{\partial(y^Ty)}{\partial\alpha} - 2n^{-1}\frac{\partial(y^TB\alpha)}{\partial\alpha} + n^{-1}\frac{\partial(\alpha^TB^TB\alpha)}{\partial\alpha} + \lambda\frac{\partial(\alpha^TK\alpha)}{\partial\alpha} = 0$$

$$0 - 2n^{-1}\boldsymbol{y}^T\boldsymbol{B} + 2n^{-1}\boldsymbol{\alpha}^T\boldsymbol{B}^T\boldsymbol{B} + 2\lambda\boldsymbol{\alpha}^T\boldsymbol{K} = 0 \quad \dots(3.5.1)$$

Secara berturut-turut, berdasarkan Teorema 1.2.1 dan Teorema 1.2.2 maka dari persamaan (3.5.1) di atas diperoleh:

$$-n^{-1}\boldsymbol{y}^T\boldsymbol{B} + n^{-1}\boldsymbol{\alpha}^T\boldsymbol{B}^T\boldsymbol{B} + \lambda\boldsymbol{\alpha}^T\boldsymbol{K} = 0$$
$$-n^{-1}(y^TB)^T + n^{-1}(\alpha^TB^TB)^T + \lambda(\alpha^TK)^T = 0$$
$$-n^{-1}\boldsymbol{B}^T\boldsymbol{y} + n^{-1}\boldsymbol{B}^T\boldsymbol{B\alpha} + \lambda\boldsymbol{K}^T\boldsymbol{\alpha} = 0 \quad \dots(3.5.2)$$

Selanjutnya, semua ruas pada persamaan (3.5.2) di atas dikalikan dengan n menjadi:

$$-\boldsymbol{B}^T\boldsymbol{y} + \boldsymbol{B}^T\boldsymbol{B}\ \boldsymbol{\alpha} + n\ \lambda\boldsymbol{K}^T\boldsymbol{\alpha} = 0$$
$$-\boldsymbol{B}^T\boldsymbol{y} + (\boldsymbol{B}^T\boldsymbol{B} + n\ \lambda\boldsymbol{K}^T)\alpha = 0$$
$$(\boldsymbol{B}^T\boldsymbol{B} + n\ \lambda\boldsymbol{K}^T)\boldsymbol{\alpha} = \boldsymbol{B}^T\boldsymbol{y}$$
$$\boldsymbol{\alpha} = (\boldsymbol{B}^T + n\ \lambda\ \boldsymbol{K}^T)^{-1}\ \boldsymbol{B}^T\boldsymbol{y} \quad \dots(3.5.3)$$

Selanjutnya, karena parameter yang diestimasi hanya $\boldsymbol{\alpha}$ maka dapat diperoleh estimator sebagai berikut ini :

$$= (\boldsymbol{B}^T + n\ \lambda\ \boldsymbol{K}^T)^{-1}\ \boldsymbol{B}^T\boldsymbol{y} \quad \dots(3.5.4)$$

Karena fungsi regresi dinyatakan sebagai kombinasi linier dari fungsi yang diketahui, maka estimasi fungsi regresi diperoleh dengan mensubtitusi persamaan (3.5.4) kedalam persamaan berikut ini :

$$\boldsymbol{f} = \boldsymbol{B\alpha}$$
$$\hat{\boldsymbol{f}} = \boldsymbol{B}\hat{\boldsymbol{\alpha}}$$
$$= \boldsymbol{B}\ (\boldsymbol{B}^T + n\ \lambda\ \boldsymbol{K}^T)^{-1}\ \boldsymbol{B}^T\boldsymbol{y}$$
$$= \boldsymbol{B}\ (\boldsymbol{B}^T + n\ \lambda\ \boldsymbol{K}^T)^{-1}\ \boldsymbol{B}^T\boldsymbol{y} \quad \dots(3.5.5)$$

Estimasi model regresi diperoleh dengan mensubtitusi persamaan (3.5.5) sebagai berikut :

$$\hat{\boldsymbol{y}} = \hat{\boldsymbol{f}}$$
$$= \boldsymbol{B}(\boldsymbol{B}^T\boldsymbol{B} + n\ \lambda\boldsymbol{K}^T)^{-1}\ \boldsymbol{B}^T\boldsymbol{y}$$
$$= \boldsymbol{S}_\lambda\ \boldsymbol{y} \quad \dots(3.5.6)$$

dimana $\boldsymbol{S}_\lambda = \boldsymbol{B}(\boldsymbol{B}^T\boldsymbol{B} + n\ \lambda\boldsymbol{K}^T)^{-1}\boldsymbol{B}^T$.

Untuk mendapatkan estimator model regresi nonparametrik, maka harus dilakukan pengestimasian fungsi parameter pada komponen nonparametrik. Estimasi fungsi pada komponen nonparametrik dilakukan dengan mendekatkan fungsi nonparametriknya dengan suatu fungsi pendekatan, yaitu fungsi B-Spline tersebut dengan menurunkan fungsi Penalized Least Square terhadap $\boldsymbol{\alpha}$ dan menyamakan hasilnya dengan nol. Sehingga diperoleh persamaan (3.5.6) diatas sebagai estimator model regresi nonparametrik.

## PENUTUP

### Kesimpulan

Berdasarkan pembahasan yang telah dilakukan diperoleh kesimpulan bahwa:

1. Estimasi fungsi pada komponen nonparametrik dilakukan dengan mendekatkan fungsi nonparametriknya dengan suatu fungsi pendekatan, yaitu fungsi B-Spline tersebut dengan menurunkan fungsi Penalized Least Square terhadap a dan menyamakan hasilnya dengan nol. Sehingga

ISSN No. 1978-3787(Cetak)
ISSN 2615-3505 (Online)

diperoleh persamaan $\hat{f} = B\left[\left(B^T + n\lambda K^T\right)\right]^{-1} B^T y$ di atas sebagai estimator model regresi nonparametrik.

2. Estimasi model regresi nonparametrik diperoleh : $\hat{y} = \hat{f} = S_\lambda y$ dengan $S_\lambda = \left[\left[B(B\right]^T B + n\lambda K^T)\right]^{-1} B^T$.

**Saran**

Penelitian yang lebih lanjut dapat dilakukan dengan meneliti estimasi model regresi nonparametrik menggunakan metode lain, seperti metode kernel, K- nearest neighborhood, deret ortogonal, analisis wavelet dan lain-lain. Serta dapat menerapkannya pada contoh kasus.

**Ucapan Terima Kasih**

Ucapan terima kasih penulis haturkan kepada Universitas Lambung Mangkurat yang telah mendanai penelitian ini melalui dana DIPA Universitas Lambung Mangkurat tahun anggaran 2021.

**DAFTAR PUSTAKA**

[1] Draper, N. R. And Smith, H. 1966. Applied Regression Analysis. Second Edition, John Wiley & Sons, Inc, New York.
[2] Eubank,R. L. 1999. Spline Smoothing and Nonparametric Regression. Second edition. Marcel Dekker, New York.
[3] Gujrati, D. N. And Porter, D. C. 2010. Dasar-dasar Ekonometrika. Salemba Empat. Jakarta.
[4] Hardle, W. 1990, Applied Nonparametric Regression, The Syndicate of the Universitas of Cambridge.
[5] Louis, L . 1986. Kalkulus dan Ilmu Ukur Analitik. Edisi ke-5 jilid 1. Erlangga. Jakarta.
[6] Nizamitdinov, A and Memmedli, M. 2012. An Application of Various Nonparametric Techniques by Nonparametric Regression Splines.
[7] http://www.naun.org/multimedia/NAUN/ . Diakses 3 Juli 2012.
[8] Rodriguez, G.2001.Smoothing and Nonparametric Regression. Diakses 23 Juli 2012.
[9] Suparti et al. 2018. Regresi Nonparametrik. Cetakan Pertama, WADE Group. Ponorogo Jawa Timur.
[10] Sudjana, 2000. Metode Statistika, Edisi ke-6, PT. Tarsito Bandung.
[11] Tripena, A. 2011. Penentuan Model Regresi Spline Terbaik. Universitas Diponegoro.Semarang.
[12] Walpole, R. E. 1994. Pengantar Statistik. Gramedia Pustaka Utama. Jakarta.
[13] Wibisono, Y. 2005. Metode Statistik. Gadjah Mada University Press. Yogyakarta.