J M S K
Jurnal Matematika, Statistika, & Komputasi
http://journal.unhas.ac.id/index.php/jmsk
Vol. 6, No.1, 29-43, Juli 2009

ISBN:1858-1382

Model B-Spline dalam Menaksir Kurva Regresi Nonparametrik

Raupong*

Abstrak

Penaksir *B-Spline* adalah salah satu metode yang digunakan untuk menaksir kurva regresi nonparametrik. Model *B-Spline* dengan titik knots digunakan untuk menyelesaikan kelemahan model *spline* pada saat orde yang tinggi, titik knot yang banyak atau knots yang terlalu dekat yang akan membentuk matriks dalam perhitungan yang hampir singular sehingga persamaan normal sulit diselesaikan. Metode yang digunakan dalam menaksir parameter *B-Spline* adalah Metode Kuadrat Terkecil (*Ordinary Least Square)* yang didefinisikan secara rekursif. Aplikasi *B-Spline* yang diberikan adalah ratio berat dan tinggi badan bayi setiap bulan. Hasil yang diperoleh adalah *B-Spline* linier dengan menggunakan satu titik knot merupakan model yang terbaik untuk data tersebut dalam menjelaskan pengaruh umur terhadap ratio berat dan tinggi badan bayi, kriteria pemilihan model terbaik yaitu R^2 (Koefisien Determinasi) dan MSE (*Means Square Error)*.

Kata Kunci: *B-Spline, Spline, titik knot, metode kuadrat terkecil, rekursif,* R^2 *(Kofisien Determinasi), Means Square Error.*

1. Pendahuluan

Analisis regresi merupakan suatu metode dalam analisis statistik yang mengalami perkembangan pesat dan banyak digunakan dalam berbagai bidang kehidupan. Analisis regresi menggambarkan sekumpulan teknik statistika yang menjadi dasar pengambilan kesimpulan (*inferensia*) tentang hubungan fungsional antara variabel prediktor (bebas) dengan varibel respon (tidak bebas). Hal ini mengakibatkan analisis regresi telah menjadi sesuatu yang pokok dalam menyelesaikan berbagai permasalahan sehingga banyak penelitian dan tulisan mengenai analisis ini (Ishaq, 2007).

Hubungan antara variabel tersebut diasumsikan mengikuti model regresi tertentu. Persoalan utama dalam analisis regresi adalah bagaimana mendapatkan taksiran bentuk kurva regresi. Untuk hal itu dikenal beberapa pendekatan yaitu: regresi parametrik dan nonparametrik. Dalam regresi parametrik bentuk kurva regresi diasumsikan diketahui baik itu linier, kuadratik, eksponensial maupun yang lainnya. Untuk dapat menggunakan pendekatan regresi parametrik ini, diperlukan pengetahuan awal tentang karakteristik data yang diselidiki. Berbeda dengan regresi nonparametrik bentuk kurva regresi diasumsikan tidak diketahui. Kurva regresi hanya diasumsikan mulus *(smooth)* yaitu termuat di dalam suatu ruang fungsi tertentu.

Selama ini banyak penelitian yang telah menggunakan pendekatan regresi parametrik dalam menaksir parameternya. Sebagian dapat memenuhi tujuan dari penelitian tersebut, namun banyak penelitian yang tidak mencapai tujuan yang diinginkan oleh peneliti itu sendiri. Hal ini disebabkan karena tujuan penelitian yang tidak sesuai dengan model parametrik serta asumsi dari bentuk kurva parametrik yang tidak terpenuhi. Untuk menyelesaikan masalah tersebut sebaiknya menggunakan regresi nonparametrik . Hal ini disebabkan karena regresi nonparametrik bersifat

* *Jurusan Matematika FMIPA Universitas Hasanuddin Makassar*

lebih fleksibel terhadap kondisi data yang diperoleh di lapangan sehingga data dapat diharapkan mencari sendiri menaksir kurvanya tanpa ada pengaruh faktor subjektifitas dari perancang penelitian.

Regresi nonparametrik menyediakan metode pendekatan untuk menaksir kurva regresi salah satunya adalah *spline*. *Spline* merupakan fungsi yang diperoleh dengan meminimumkan *Penalized Least Square* (PLS) yaitu kriteria taksiran yang menggabungkan *goodness of fit* dengan kemulusan (*smooth*) kurva.

Pendekatan *spline* optimal dapat berdasarkan parameter pemulus atau penggunaan titik-titik knots. Namun penaksir spline dengan menggunakan parameter pemulus optimal yang dikembangkan oleh peneliti-peneliti secara visual sulit membedakan perubahan perilaku dari suatu pola data pada interval-interval yang berlainan. Ini menunjukkan bahwa secara visual perannya sulit dilihat dalam membedakan perubahan perilaku dari suatu pola data pada interval-interval yang berlainan. Sehingga ciri khas dari pendekatan *spline* yang berupa *segmented* tidak terlihat secara signifikan (Budiantara, 2004).

Spline mempunyai kelemahan pada saat orde spline yang tinggi, titik knots yang banyak dan knots yang terlalu dekat akan membentuk matriks dalam perhitungan yang hampir singular, sehingga persamaan normal sulit diselesaikan. basis yang dapat mengatasi kelemahan ini adalah basis *B-Spline* (Budiantara *et al.*, 2006).

Dalam tulisan ini akan dibahas mengenai penentuan penaksir *B-Spline* pada model regresi nonparametrik dan membandingkan antara model *B-Spline* linier dengan model *B-Spline* kuadratik.

2. Modifikasi Penaksiran Robust

2.1 Regresi Nonparametrik dan *Smoothing Spline*

Secara umum persamaan regresi nonparametrik dapat dituliskan sebagai berikut:

$$y_i = g(t_i) + \varepsilon_i \,, \qquad i = 1,2,....,n \,. \tag{1}$$

Ada beberapa pendekatan untuk menaksir fungsi $g(t_i)$ dalam regresi nonparametrik antara lain pendekatan *kernel*, *Spline*, dan *K-Nearest Neigbor* (Hardle, 1990). *Estimator Deret Fourier* (Eubank, 1988), dan *Histogram*. Pendekatan regresi nonparametrik yang cukup populer adalah *Spline* (Ishaq, 2007).

Spline ditemukan pertama kali oleh Whitaker pada tahun 1923. Pada mulanya Whitaker menggunakan fungsi polinomial untuk menghampiri fungsi-fungsi kontinu dan diferesiabel. Saat itu fungsi polinomial sangat terkenal karena mudah dan sederhana, mempunyai interpolasi tunggal dan selalu dapat diaplikasikan untuk fungsi yang diferensiabel dengan memanfaatkan deret *Taylor* atau *Maclaurin.*

Walaupun demikian, bukan berarti fungsi polinomial merupakan fungsi pendekatan yang tidak mempunyai kelemahan sama sekali. Dalam beberapa hal khusus, perlu dikaji ulang penggunaan polinomial untuk fungsi-fungsi yang bersifat lokal, *smooth*, dan fungsi yang mempunyai loncatan naik/turun yang tajam. Dalam kondisi seperti ini, polinomial sulit diandalkan sebagai penaksir fungsi, karena polinomial mempunyai osilasi yang rendah dan sangat kaku dalam domain yang ditentukan. Akibatnya harus dicari suatu fungsi yang dapat mengatasi kelemahan polinomial tersebut. Fungsi yang dimaksud adalah *Spline* (Budiantara, 2004).

Spline merupakan potongan-potongan polinomial yang memiliki sifat tersegmen kontinu. Sifat inilah yang memberikan fleksibilitas yang lebih dari pada polinomial biasa, sehingga memungkinkan untuk menyesuaikan diri secara efektif terhadap karakteristik lokal dari suatu fungsi. Pada pendekatan spline fungsi y_i diasumsikan termuat dalam suatu *ruang Sobolev*

$$W_2^m = \left\{ g; \int_0^1 \left[g^{(m)}(x) \right]^2 dx < \infty \right\}. \tag{2}$$

Pemilihan ruang fungsi *sobolev* dimotivasi oleh sifat *smoothness* (Eubank, 1988).

Fungsi *Spline* pada mulanya hanya digunakan untuk hampiran numerik fungsi-fungsi *smooth* dan lokal. Fungsi *Spline* sama sekali belum diaplikasikan dalam analisis regresi, sebagai suatu hasil proses optimasi dalam regresi nonparametrik. Setelah dua sampai tiga dasawarsa terakhir, Schoenberg memperkenalkan fungsi *Spline* dari hasil optimasi bersyarat dalam analisis regresi nonparametrik. Selanjutnya bermunculan pula tulisan-tulisan tentang *Spline* pada regresi nonparametrik dalam (Budiantara, 2004) seperti Reinch (1971), Wahba (1971;1973;1975;1990), Speckman (1985;1988), Budiantara *et al*. (1997), Cox (1983), Cox dan O'Sullivan (1996), Koenker *et al.* (1994) dan puluhan tulisan lain yang serupa.
Estimator spline diperoleh dengan meminimumkan PLS:

$$\left\{ n^{-1} \sum_{i=1}^{n} (y_i - g(t_i)^2) + \lambda \int \left[g^m(t) \right]^2 dt \right\}, \tag{3}$$

yaitu kriteria optimasi yang menggabungkan antara kecocokan terhadap data (*goodness of fit*), kemulusan kurva (*penalty*), dan parameter pemulus λ (Ishaq, 2007).

Penaksir *Spline* yang diperoleh dari optimasi *PLS* merupakan penaksir linear dalam observasi (Budiantara, 2004) yang dapat ditulis sebagai berikut:

$$\hat{h}_\lambda(x) = G(\lambda). \tag{4}$$

Untuk suatu matriks $G(\lambda)$ yang berukuran $n \times n$ tergantung kepada parameter pemulus λ. Jika parameter pemulus sangat kecil maka penaksir spline sangat kasar, sedangkan jika parameter pemulus sangat besar maka penaksir spline sangat mulus (Hardle, 1990).

Jika ingin menaksir kurva regresi nonparametrik dengan pendekatan *Spline*, maka dapat dilakukan dengan salah satu dari dua cara berikut:

1. Mencari model *Spline* optimal dengan cara memilih parameter pemulus λ optimal.
1. Mencari model *Spline* optimal dengan cara memilih titik knot optimal.

Titik knots merupakan titik perpaduan bersama yang memperlihatkan terjadinya perubahan perilaku dari fungsi spline pada interval-interval yang berbeda. Peran titik knots dan parameter pemulus λ dalam model *Spline* adalah serupa tetapi kemampuan menaksir spline dengan memilih parameter penghalus optimal tidak terlihat secara eksplisit dan visual. Dalam persoalan praktis, bentuk visual sangat membantu dalam pemodelan awal maka pemilihan titik knots adalah optimal dalam regresi nonparametrik *Spline* karena cenderung lebih mudah dan lebih disukai dibandingkan dengan pemilihan parameter penghalus optimal. Hal ini disebabkan karena:

1. Titik knots mudah divisualisasikan dan sifatnya mudah dipahami.
2. Dengan titik knot memiliki interprestasi Matematika dan Statistika yang lebih baik dan sederhana karena optimasi yang digunakan cukup hanya dengan optimasi *least square* biasa (Islamiyati, 2006).

Spline mempunyai kelemahan pada saat orde *Spline* yang tinggi, titik knot yang banyak atau knot yang terlalu dekat akan membentuk matriks dalam perhitungan yang hampir singular, sehingga persamaan normal sulit diselesaikan. Basis yang dapat mengatasi kelemahan ini adalah basis *B-Spline* (Budiantara *et al.*, 2006).

2.2. Model *B*-Spline

Model regresi $y_i = g(t_i) + \varepsilon_i,\ i = 1,2,....,n$, dengan ε_i merupakan residual dan $g(t_i)$ adalah kurva regresi. Jika kurva regresi *g* didekati dengan fungsi *B-Spline* maka *g* dapat ditulis menjadi

$$g(t) = \sum_{j=1}^{m+K} \gamma_j B_{j-m,m}(t), \tag{5}$$

dengan $B_{j-m,m}$ merupakan basis *B-Spline* (Eubank, 1988). Cara membangun fungsi *B-Spline* orde *m* dengan titik knots $a < u_1 < ... < u_K < b$, yaitu dengan cara terlebih dahulu mendefinisikan knot tambahan sebanyak *2m*, yaitu

$$u_{-(m-1)},...,u_{-1},u_0,...,u_{K+m},$$

dimana

$$u_{-(m-1)} = \ldots = u_0 = a \text{ dan } u_{K+1} = \ldots = u_{K+m} = b.$$

Fungsi *B-Spline* didefinisikan secara rekursif sebagai berikut:

$$B_{i,m}(t) = \frac{t-u_i}{u_{i+m-1}-u_i} B_{i,m-1}(t) + \frac{u_{i+m}-t}{u_{i+m}-u_{i+1}} B_{i+1.m-1}(t), \tag{6}$$

dimana

$$B_{j.1}(t) = \begin{cases} 1, & \text{jika } \mathrm{u_j} < t < u_{j+1} \\ 0, & jika\ t < u_j \quad atau \quad t \geq u_j \end{cases} \text{(Eubank,1988)}, \tag{7}$$

dengan *m* adalah derajat *B-Spline* , untuk m=2 memberikan fungsi *B-Spline* linier, m=3 memberikan fungsi *B-Spline* kuadratik dan m=4 memberikan fungsi *B-Spline* kubik (Budiantara *et al.*, 2006).

Untuk menaksir koefisien γ pada persamaan (5), didefinisikan matriks:

$$B(\lambda) = \left(B_{j,m}(t_i)\right)_{\substack{i=1,...,n \\ j=-(m-1),...,K}}. \tag{8}$$

Atau dapat ditulis sebagai berikut:

$$B(\lambda) = \begin{bmatrix} B_{-(m-1),m}(t_1), B_{-(m-2),m}(t_1),..., B_K, m(t_1) \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ B_{-1(m-1),m}(t_n), B_{-(m-2),m}(t_n),..., B_{K,m}(t_n) \end{bmatrix}. \tag{9}$$

Jadi $B(\lambda)$ menurut Budiantara *et al.* (2006) adalah sebuah matriks berukuran $n \times (m+K)$.

2.3 Penaksir Parameter B-Spline Linier

Adapun persamaan umum regresi nonparametrik dengan menggunakan pendekatan kurva B-Spline:

$$y_i = g(t_i) + \varepsilon_i \quad = \sum_{j=1}^{m+K} \gamma_j B_{j-m,m}(t_i) + \varepsilon_i \text{ , } i = 1,2,...,n.$$

Selanjutnya akan ditunjukkan cara memperoleh penaksir parameter *B-Spline* linier dengan menggunakan Metode Kuadrat Terkecil (*Ordinarry Least Square)* secara rekursif. Fungsi B-Spline linier (m=2) dengan titik knots adalah K dapat ditulis

$$\begin{aligned} y_i &= \sum_{j=1}^{2+K} \gamma_j B_{j-m,m}(t_i) + \varepsilon_i \text{ , } i = 1,2,...,n \\ &= \gamma_1 B_{1-2,2}(t_i) + \gamma_2 B_{2-2,2}(t_i) + \gamma_3 B_{3-2,2}(t_i) + ... + \gamma_n B_{(2+K)-2,2} + \varepsilon_i \\ &= \gamma_1 B_{-1,2}(t_i) + \gamma_2 B_{0,2}(t_i) + \gamma_3 B_{1,2}(t_i) + ... + \gamma_n B_{K,2} + \varepsilon_i \text{ .} \end{aligned}$$

Untuk menaksir koefisien γ pada fungsi *B-Spline* linier dengan titik knots K adalah:

$$B(\lambda) = \left(B_{j,m}(t_i)\right)_{\substack{i=1,...,n \\ j=-1,...,n}} \text{ ,}$$

$$B(\lambda) = \left(B_{j,2}(t_i)\right)_{\substack{i=1,...,n \\ j=-1,...,n}} \text{ ,}$$

atau

$$B(\lambda) = \left(B_{-(m-1),m}(t_i), B_{-(m-2),m}(t_i), ..., B_{K,m}(t_i)\right), \ i = 1,2,...,\text{n ,}$$

$$B(\lambda) = \left(B_{-(2-1),2}(t_i) \quad B_{-(2-2),2}(t_i) \quad B_{-(3-2),2}(t_i) \quad ... \quad B_{-(2+K-2),2}(t_i)\right),$$

$i = 1,2,...,n,$

$$B(\lambda) = \left(B_{-1,2}(t_i) \quad B_{0,2}(t_i) \quad B_{1,2}(t_i) \quad ... \quad B_{K,2}(t_i)\right).$$

Jadi $y_i = \left(B_{-1,2}(t_i) \quad B_{0,2}(t_i) \quad B_{1,2}(t_i) \quad ... \quad B_{K,2}(t_i)\right) \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ ... \\ \gamma_n \end{bmatrix} + \varepsilon_i \text{ .}$

Misalkan

$$B(\lambda) = \left(B_{-1,2}(t_i) \quad B_{0,2}(t_i) \quad B_{1,2}(t_i) \quad ... \quad B_{K,2}(t_i)\right), \ \gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ ... \\ \gamma_n \end{bmatrix}, \quad \varepsilon_i = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ ... \\ \varepsilon_n \end{bmatrix},$$

maka $y_i = B(\lambda)\gamma + \varepsilon$.

Taksiran parameter $\gamma = (\gamma_1, \gamma_2, ..., \gamma_n)^T$ diperoleh dengan menggunakan Metode Kuadrat Terkecil. Penaksir $\hat{\gamma}$ diperoleh dengan menyelesaikan optimasi:

$$Min\{\varepsilon^T \varepsilon\} = Min\{(y - B(\lambda)\gamma)^T (y - B(\lambda)\gamma)\}.$$

Dalam lambang matriks

$$Min\left\{(\varepsilon_1,\varepsilon_2,...,\varepsilon_n)\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ ...\\ \varepsilon_n\end{pmatrix}\right\}=$$

$$Min\left\{\begin{array}{l}\left[\begin{pmatrix}y_1\\ y_2\\ ...\\ y_n\end{pmatrix}-\begin{pmatrix}B_{-1,2}(t_1) & B_{0,2}(t_1) & B_{1,2}(t_1) & ... & B_{n,2}(t_1)\\ B_{-1,2}(t_2) & B_{0,2}(t_2) & B_{1,2}(t_2) & ... & B_{n,2}(t_2)\\ ... \\ B_{-1,2}(t_n) & B_{0,2}(t_n) & B_{1,2}(t_n) & ... & B_{n,2}(t_n)\end{pmatrix}\begin{pmatrix}\gamma_1\\ \gamma_2\\ ...\\ \gamma_n\end{pmatrix}\right]^T\\ \left[\begin{pmatrix}y_1\\ y_2\\ ...\\ y_n\end{pmatrix}-\begin{pmatrix}B_{-1,2}(t_1) & B_{0,2}(t_1) & B_{1,2}(t_1) & ... & B_{n,2}(t_1)\\ B_{-1,2}(t_2) & B_{0,2}(t_2) & B_{1,2}(t_2) & ... & B_{n,2}(t_2)\\ ... \\ B_{-1,2}(t_n) & B_{0,2}(t_n) & B_{1,2}(t_n) & ... & B_{n,2}(t_n)\end{pmatrix}\begin{pmatrix}\gamma_1\\ \gamma_2\\ ...\\ \gamma_n\end{pmatrix}\right]\end{array}\right\},$$

dengan $y=(y_1, y_2,..., y_n)$ dan $B(\lambda)$ adalah matriks yang berukuran $n\times(2+K)$,

$$J=\varepsilon^T\varepsilon \;=y^T y-2y^T B(\lambda)\gamma+\gamma B(\lambda)^T B(\lambda)\gamma\ .$$

Jika $J=\sum\varepsilon_i^2=\sum\left(y_i-\gamma_1 B_{-1,2}(t_i)-\gamma_2 B_{0,2}(t_i)-\gamma_3 B_{1,2}(t_i)\ \ ...\ \ \gamma_n B_{n,2}(t_i)\right)^2$, maka

$$\frac{\partial J}{\partial\gamma_1}=-2\sum\left(y_i-\gamma_1 B_{-1,2}(t_i)-\gamma_2 B_{0,2}(t_i)-\gamma_3 B_{1,2}(t_i)\ \ ...\ \ \gamma_n B_{n,2}(t_i)\right)=0$$

$$\frac{\partial J}{\partial\gamma_2}=-2\sum\left(y_i-\gamma_1 B_{-1,2}(t_i)-\gamma_2 B_{0,2}(t_i)-\gamma_3 B_{1,2}(t_i)\ \ ...\ \ \gamma_n B_{n,2}(t_i)\right)=0$$

$$\frac{\partial J}{\partial\gamma_3}=-2\sum\left(y_i-\gamma_1 B_{-1,2}(t_i)-\gamma_2 B_{0,2}(t_i)-\gamma_3 B_{1,2}(t_i)\ \ ...\ \ \gamma_n B_{n,2}(t_i)\right)=0$$

$$\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots$$

$$\frac{\partial J}{\partial\gamma_n}=-2\sum\left(y_i-\gamma_1 B_{-1,2}(t_i)-\gamma_2 B_{0,2}(t_i)-\gamma_3 B_{1,2}(t_i)\ \ ...\ \ \gamma_n B_{n,2}(t_i)\right)=0$$

sehingga

$$\sum y_i=\gamma_1 B_{-1,2}(t_i)+\gamma_2\sum B_{0,2}(t_i)+...+\gamma_n\sum B_{n,2}(t_i)$$

$$\sum y_i=\gamma_1\sum B_{-(m-1),2}(t_i)+\gamma_2\sum B_{-(m-2),2}(t_i)^2+...+\gamma_n\sum B_{K,2}(t_i)$$

dengan

$$B(\lambda)=\begin{bmatrix} B_{-(2-1),2}(t_1) & B_{-(2-2),2}(t_1) & B_{-(3-2),2}(t_1) & B_{K,2}(t_1) \\ B_{-(2-1),2}(t_2) & B_{-(2-2),2}(t_2) & B_{-(3-2),2}(t_2) & B_{K,2}(t_2) \\ \ldots & & & \\ B_{-(2-1),2}(t_n) & B_{-(2-2),2}(t_n) & B_{-(3-2),2}(t_n) & B_{K,2}(t_n) \end{bmatrix},$$

$$B(\lambda)=\begin{bmatrix} B_{-1,2}(t_1) & B_{0,2}(t_1) & B_{1,2}(t_1) & \ldots & B_{K,2}(t_1) \\ B_{-1,2}(t_2) & B_{0,2}(t_2) & B_{1,2}(t_2) & \ldots & B_{K,2}(t_2) \\ \ldots & & & & \\ B_{-1,2}(t_n) & B_{0,2}(t_n) & B_{1,2}(t_n) & \ldots & B_{K,2}(t_n) \end{bmatrix}, \quad b=\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \ldots \\ \gamma_n \end{bmatrix},$$

sehingga persamaan di atas dapat dituliskan menjadi

$$B^T(\lambda)Y = \left[(B^T(\lambda))(B(\lambda)\right]b$$

$$b = (B^T(\lambda)B(\lambda))^{-1}(B^T(\lambda)Y)$$

$$b=\hat{\gamma} \text{ , vektor taksiran dari } \gamma \text{ .}$$

Jadi diperoleh taksiran

$$\hat{\gamma} = (B^T(\lambda)B(\lambda))^{-1}(B^T(\lambda)Y).$$

Hal ini berakibat untuk estimasi kurva regresi $g(t_i)$ yang diberikan oleh

$$\hat{g}(t_i) = B(\lambda)\hat{\gamma} \;= B(\lambda)(B^T(\lambda)B(\lambda))^{-1}(B(\lambda)Y) = H(\lambda)Y,$$

dengan

$$H(\lambda)Y = B(\lambda)(B^T(\lambda)B(\lambda))^{-1}B^T(\lambda).$$

Terlihat bahwa $\hat{g}(t_i)$ merupakan penaksir linier dalam observasi $Y=(y_1,y_2,\ldots,y_n)^T$ dan sangat bergantung pada titik knots $u_1,u_2,\ldots,u_K$.

3. Metodologi Penelitian

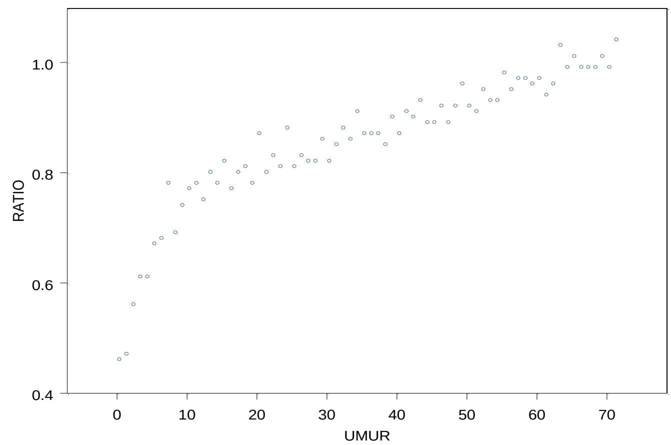
Data yang digunakan data sekunder, yang diambil dari data penelitian pada suatu kota untuk menaksir pola hubungan antara umur dan ratio berat dan tinggi bayi (Budiantara, 2004). Adapun langkah-langkah yang dilakukan adalah sebagai berikut:

1. Membuat basis dari fungsi.
2. Mencari penaksir *B-Spline* linier dan *B-Spline* kuadratik yaitu koefisien γ menggunakan *software S-Plus* 2000.
3. Menentukan model terbaik dengan menggunakan R^2 (koefisien determinasi) dan *MSE* (*Means Square Error)* sebagai indikator pemilihan.
4. Membandingkan antara model *B-Spline* linier dan *B-Spline* kuadratik dengan kriteria yang mempunyai MSE (*Means Square Error)* kecil dan koefisien determinasi ($R^{2)}$ terbesar, yang akan digunakan sebagai model pendekatan untuk menentukan pola hubungan antara umur dengan ratio berat dan tinggi badan bayi.

4. Aplikasi

Aplikasi *B-Spline* linier dan *B-Spline* kuadratik adalah menggunakan data umur dengan ratio berat dan tinggi badan bayi yang diambil dari data penelitian Budiantara (2004). Salah satu indikator yang menentukan seorang bayi itu mengalami pertumbuhan adalah ratio antara berat badan dan tinggi badan bayi. Hal ini dipakai karena diharapkan ada pola hubungan antara umur bayi (x) terhadap ratio berat badan dan tinggi badan bayi (y).

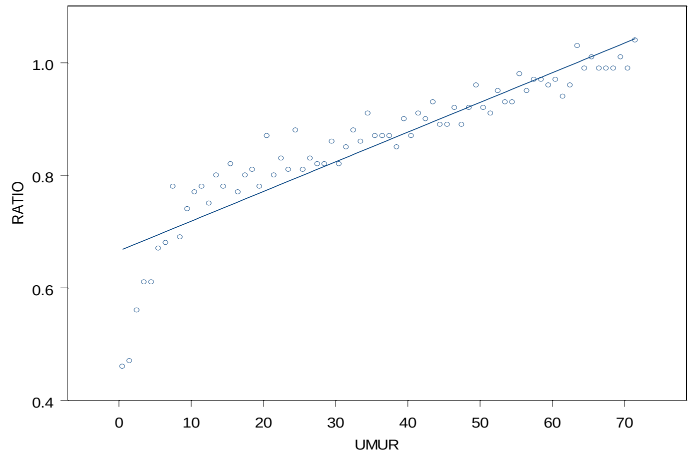
Di bawah ini akan diberikan ilustrasi visual pendekatan regresi parametrik linier, kuadratik, *B-spline* linier dan *B-Spline* kuadratik menggunakan titik knot untuk menaksir pola hubungan antara umur terhadap ratio berat dan tinggi badan bayi pada suatu kota. Dengan sampel sebanyak 72 bayi diambil secara acak setelah itu umur dan ratio berat dengan tinggi badan bayi dicatat.



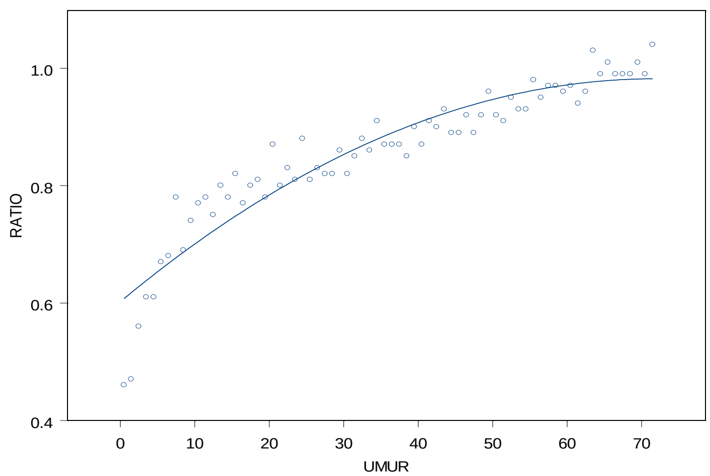
Gambar 6. Plot hubungan antara umur bayi terhadap ratio berat dan tinggi badan bayi.

4.1. Regresi Parametrik

Pola hubungan antara umur bayi (x) terhadap ratio berat badan dan tinggi badan bayi (y) dengan menggunakan pendekatan regresi parametrik linier dan kuadratik dapat dilihat dalam Gambar 7 dan 8.



Gambar 7. Plot regresi parametrik linier.



Gambar 8. Plot regresi parametrik kuadratik.

Berdasarkan Gambar 7 dan Gambar 8, terlihat bahwa dengan bertambahnya umur bayi (*x*) dapat menyebabkan kenaikan pada ratio berat badan dan tinggi badan bayi (*y*), sedangkan pengaruh umur bayi (*x*) terhadap ratio berat dan tinggi badan bayi (*y*) untuk model regresi parametrik linier dan kuadratik, disajikan pada Tabel 1 berikut.

Tabel 1. Pengaruh umur bayi terhadap ratio berat dan tinggi badan bayi dengan menggunakan model regresi linier dan kuadratik.

Model Regresi	**Parameter**	**Koefisien Regresi**	***F*-Hitung**	R^2	**MSE**
Linier	β_0	0,6656	324,44	0,823	0,002668
	β_1	0,005276			
Kuadratik	β_0	0,6022	247,06	0,877	0,001869
	β_1	0,01056			

β_2	- 0,000073

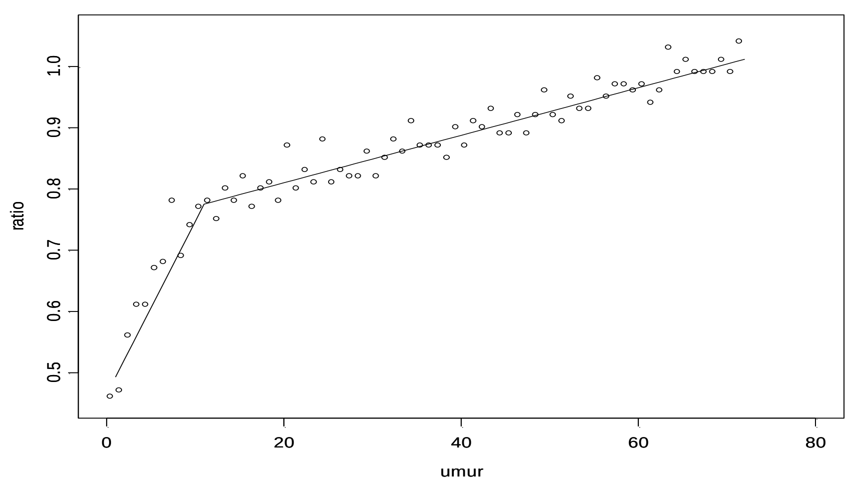
Tabel 1, menunjukkan bahwa umur bayi (x) berpengaruh terhadap ratio berat dan tinggi badan bayi (y) baik pada regresi parametrik linier maupun kuadratik. Nilai R^2 (koefisien determinasi) pada model regresi parametrik kuadratik adalah 0,877 sedangkan nilai R^2 (koefisien determinasi) regresi parametrik linier adalah 0,823. Nilai *Mean Square Error* (MSE) model regresi parametrik kuadratik lebih kecil yaitu 0,001869 jika dibandingkan dengan parametrik linier yaitu 0,002668. Hal ini menunjukkan bahwa model terbaik pada pendekatan regresi parametrik adalah model regresi kuadratik.

Namun secara visual terlihat ketidaksempurnaan pada model parametrik linier dan kuadratik dalam menaksir pola data yang *over estimate* pada bayi umur 0 bulan (bayi baru lahir) sampai umur 9 bulan. Sedangkan pada umur bayi sekitar 10 bulan sampai 26 bulan kedua model regresi ini menaksir pola data dengan *under estimate*.

Oleh karena itu, diperlukan suatu model taksiran lain yang lebih mampu menjelaskan hubungan antara umur bayi (x) terhadap ratio berat badan dan tinggi badan bayi (y). Dalam hal ini sebaiknya menggunakan pendekatan regresi nonparametrik B-Spline linier dan kuadratik dengan menggunakan titik knots. Ukuran yang digunakan untuk melihat ketepatan model maka dipergunakan R^2 dan MSE.

4.2. Penaksir *B-Spline* Linier dan *B-Spline* Kuadratik

Plot penaksir *B-Spline* linier dengan menggunakan 1 titik knots dapat ditunjukkan pada Gambar 9. Secara visual dapat dilihat bahwa kenaikan ratio berat badan dan tinggi badan badan bayi yang berumur nol bulan (bayi baru lahir) sampai dengan 10,5 bulan mengalami kenaikan yang sangat signifikan pada bayi yang berumur 10,5 bulan sampai dengan bayi yang berumur sekitar 71,5 bulan kenaikan ratio berat dan tinggi badan bayi mengalami kenaikan yang tidak terlalu signifikan, sehingga pola hubungan antara umur dengan ratio berat dan tinggi badan bayi pada setiap interval bulan dapat teridentifikasi.



Gambar 9. Plot *B-Spline* linier dengan satu titik knot (k=10,5).

Taksiran model *B-Spline* linier dengan menggunakan satu titik knots yaitu pada titik k_1 =10,5 dapat ditunjukkan pada Tabel 2. Dari tabel tersebut dapat diketahui bahwa nilai R^2 (Koefisien Determinasi) dari model *B-Spline* linier dengan 1 titik knot yaitu pada k=10,5 adalah 0,958117 lebih besar jika dibandingkan dengan koefisien determinasi dari regresi parametrik linier dan kuadratik yaitu 0,823 dan 0,877. Hal ini mengindikasikan besarnya pengaruh umur bayi terhadap ratio berat badan dan tinggi badan bayi. Nilai MSE (*Mean Square Error*) dari model *B-Spline* Linier dengan 1 titik knot lebih rendah yaitu 0,00064819 jika dibandingkan dengan MSE dari model regresi parametrik linier dan kuadratik yaitu 0,002668 dan 0,001869.

Tabel 2. Pengaruh umur bayi terhadap ratio berat dan tinggi badan bayi dengan menggunakan model B-Spline Linier dengan 1 titik knot.

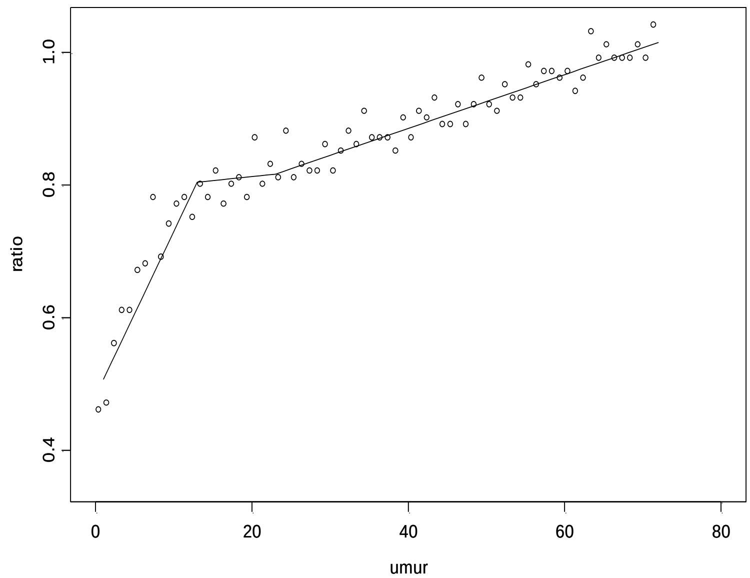
Model Regresi	**Parameter**	**Koefisien Regresi**	***F*-Hitung**	R^2	**MSE**
B-Spline Linier	γ_1	0,4928998	518,5237	0, 117	0,00064819
	γ_2	0,7753550			
	γ_3	1,0117403			

Pada Tabel 3 ditunjukkan taksiran model *B-Spline* linier dengan menggunakan 2 titik knot yaitu pada k_1=12,5 dan k_2 = 22,5. Nilai R^2 dari model tersebut adalah 0,9518657 lebih besar dari pada nilai R^2 pada model regresi parametrik linier, kuadratik dan *B-Spline* linier. Begitu pula dengan nilai MSE (*Means Square Error)* dari model *B-Spline* linier dengan 2 titik knot pada yaitu 0,00062483 lebih rendah dibandingkan dengan nilai MSE dari model regresi parametrik linier, kuadratik, dan *B-Spline* linier yaitu 0,002668, 0,001869 dan 0,00064819.

Tabel 3. Pengaruh umur bayi terhadap ratio berat dan tinggi badan bayi dengan menggunakan model B-Spline Kuadratik dengan 1 titik knot.

Model Regresi	**Parameter**	**Koefisien Regresi**	***F*-Hitung**	R^2	**MSE**
B-Spline kuadratik	γ_1	0,5070078	448,2377	0,9518657	0,00074494
	γ_2	0,8042153			
	γ_3	0,8163859			
	γ_4	1,0149879			

Pendekatan *B-Spline* linier dengan 2 titik knot pada pada k_1=12,5 dan k_2= 22,5 secara visual ditunjukkan pada Gambar 10. Bayi yang berumur 0 bulan (baru lahir) sampai dengan bayi yang berumur sekitar 12,5 bulan mengalami kenaikan ratio berat badan dan tinggi badan yang sangat signifikan, sedangkan bayi yang berumur sekitar 12,5 bulan sampai dengan bayi yang berumur sekitar 22,5 bulan mengalami penurunan ratio berat dan tinggi badan. Untuk bayi yang berumur 23,5 sampai dengan 75,5 bulan ratio berat badan dan tinggi badan bayi perlahan mengalami peningkatan sama seperti halnya yang terjadi pada bayi umur 0 bulan (baru lahir) samapai dengan bayi yang berumur 12,5 bulan.



Gambar 10. Plot *B-Spline* linier dengan 2 titik knot (k_1=12,5 dan k_2 = 22,5).

4.3. Model *B-Spline* Linier Dan *B-Spline* Kuadratik

Model *B-Spline* linier dengan menggunakan 1 titik knot pada k =10,5, diberikan oleh

$$\hat{y}(x) = 0{,}4928998\ B_{-1,2}(t) + 0{,}7753550\ B_{0,2}(t) + 1{,}0117403\ B_{1,2}(t),$$

dengan

$$B_{-1,2}(t) = \begin{cases} \dfrac{10.5 - t}{10}, & 0.5 < \mathrm{t} \le 10.5 \\ 0, & 10.5 < t \le 71.5 \end{cases}, \quad B_{0,2}(t) = \begin{cases} \dfrac{t - 0.5}{10}, & 0.5 < \mathrm{t} \le 10.5 \\ \dfrac{71.5 - t}{61}, & 10.5 < t \le 71.5 \end{cases},$$

$$B_{1,2}(t) = \begin{cases} 0, & 0.5 < \mathrm{t} \le 10.5 \\ \dfrac{t - 10.5}{61}, & 10.5 < t \le 71.5 \end{cases},$$

$$B_{2,2}(t) = \begin{cases} 0, & 0.5 < \mathrm{t} \le 12.5 \\ 0, & 12.5 < t \le 22.5 \\ \dfrac{t - 22.5}{49}, & 22.5 < t \le 71.5 \end{cases},$$

adalah basis dari model *B-Spline* linier dengan 1 titik knot pada k = 10,5.

Model *B-Spline* linier dengan menggunakan 2 titik knot pada k_1=12,5 dan k_2 = 22,5 diberikan oleh

$$\hat{y}(x) = -\ 0.44764169\, B_{-1,2}(t) + 0.93151595\, B_{0,2}(t)$$

$$- 0.02933401\, B_{1,2}(t) + 1.01387715\; B_{2,2}(t)$$

dengan

$$B_{-1,2}(t) = \begin{cases} \dfrac{(12.5 - t\ )}{12}, & 0.5 < t \le 12.5 \\ 0\ , & 12.5 < t \le 22.5 \\ 0\ , & 22.5 < t \le 71.5 \end{cases} ,$$

$$B_{0,2}(t) = \begin{cases} \dfrac{t - 0.5}{12}, & 0.5 < \mathrm{t} \le 12.5 \\ \dfrac{22.5\text{-}\mathrm{t}}{10}, & 12.5 < t \le 22.5 \\ 0, & 22.5 < t \le 71.5 \end{cases}$$

$$B_{1,2}(t) = \begin{cases} 0, & 0.5 < \mathrm{t} \le 12.5 \\ \dfrac{\mathrm{t}\text{-}12.5}{10}, & 12.5 < t \le 22.5 \\ \dfrac{71.5 - t}{49}, & 22.5 < t \le 71.5 \end{cases} ,$$

$$B_{2,2}(t) = \begin{cases} 0, & 0.5 < \mathrm{t} \le 12.5 \\ 0, & 12.5 < t \le 22.5 \\ \dfrac{t - 22.5}{49}, & 22.5 < t \le 71.5 \end{cases}$$

adalah basis dari model *B-Spline* linier dengan 2 titik knots pada k_1 = 12,5 dan k_2 = 22,5.

Tabel 4. Nilai R^2 (Koefisien Determinasi) dan MSE (*Means Square Error)* dari masing-masing model.

Model Regresi	$\boldsymbol{R^2}$	**MSE**
Parametrik Linier	0,823	0,002668
Parametrik Kuadratik	0,877	0,001869
B-Spline Linier dengan k= 10,5	0,958	0,00064819
B-Spline Linier dengan k_1=12,5 dan k_2 = 22,5	0,951	0,00074494

Dari Tabel 4 di atas dapat dikatakan bahwa model *B-Spline* linier dengan *k* = 10,5 mempunyai nilai R^2 yaitu 0,958 dan MSE yang terkecil yaitu 0,00064819 merupakan model yang terbaik dalam menaksir hubungan antara umur bayi terhadap ratio berat dan tinggi badan bayi.

Secara visual terlihat bahwa hubungan antara umur bayi dengan ratio berat badan dan tinggi badan bayi mempunyai pola yang berbeda pada setiap interval umur yang berbeda. Pada bayi yang berumur 0 bulan (baru lahir) memiliki kecenderungan kenaikan ratio berat dan tinggi badan yang cukup besar hal ini disebabkan oleh beberapa faktor diantaranya yaitu asupan gizi yang baik melalui makan bayi yang teratur tanpa disertai oleh aktivitas lain misalnya bermain. Sedangkan untuk bayi yang berumur 10,5 tahun mengalami kenaikan ratio berat dan tinggi badan yang tidak signifikan hal ini disebabkan oleh munculnya aktivitas yang dilakukan oleh bayi misalnya bermain, merangkak, dan latihan berjalan yang intensitasnya cukup tinggi yang menyebabkan pola makan yang tidak teratur sehingga pertumbuhan badan bayi sedikit. Bayi yang berumur sekitar 15,5 bulan sampai dengan 75,5 bulan kembali mengalami peningkatan ratio berat badan dan tinggi badan bayi yang cukup signifikan. Ini berati terdapat indikasi perubahan prilaku pola pertumbuhan bayi pada sub-sub interval umur bayi tertentu.

5. Kesimpulan dan Saran

5.1. Kesimpulan

Berdasarkan uraian pada bagian sebelumnya maka dapat disimpulkan sebagai berikut:

1. Salah satu model yang digunakan untuk menaksir kurva regresi nonparametrik adalah *B-Spline* dengan penaksiran parameternya menggunakan Metode Kuadrat Terkecil (*Ordinarry Least Square)*.
2. Untuk membandingkan model *B-Spline* linier dan *B-Spline* kuadratik dengan menggunakan 1 titik knot didasarkan pada nilai R^2 terbesar dan nilai MSE terkecil. Hasil aplikasi dengan menggunakan data umur dengan ratio berat badan dan tinggi badan bayi menunjukkan bahwa model *B-Spline* linier dengan 1 titik knot pada k=10,5 merupakan model yang terbaik untuk menjelaskan pengaruh umur tehadap ratio berat badan dan tinggi badan bayi.

5.2. Saran

Penulis menyadari bahwa permasalahan yang dikaji masih sangat terbatas, sehingga disarankan sebagai berikut:

1. Mengembangkan bidang kajian seperti model *B-Spline* polinomial dengan menggunakan lebih dari 1 titik knot dan melakukan pemilihan titik knot optimal dengan menggunakan metode GCV (*Generalized Cross Validation)* atau metode lainnya.
2. Aplikasi kajian penaksir *B-Spline* pada bidang-bidang lainnya.

Daftar Pustaka

[1] Budiantara, I.N., 2004, Spline: Historis, Motivasi dan Perannya dalam Regresi Nonparametrik, *Makalah Pembicara Utama pada Konferensi nasional Matematika XII*, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Udayana, Denpasar- Bali.

[2] Budiantara, I.N., 2005, Model Keluarga Spline Polinomial Truncated dalam Regresi Semiparametrik, *Jurnal ITS*, Surabaya.

[3] Budiantara, I.N. *et.al.*, 2006, Pemodelan *B-Spline* dan Mars Pada Nilai Ujian Masuk Terhadap IPK Mahasiswa Jurusan Disain Komunikasi Visual UK. Petra Surabaya, *Jurnal UKM Petra*, Surabaya.

[4] Giulio, C. and Valori, G., *An Inductive proof of the Derivative B-Spline Recursion Formula*, Departement of Mathematics, University of Bologna P.zza di Porta S, Donato 5 4027 Bologna, Italy.

[5] Eubank, R.L., 1988, *Spline Smoothing and Nonparametric Regression*, Marcel Dekker Inc., New York.

[6] Hardle, W., 1990, *Applied Nonparametric Regression*, Cambridge University Press, New York.

[7] Islamiyati, Anna, 2006, Spline Kubik dengan Knots Optimal, *Jurnal Penelitian*, Makassar.

[8] Ishaq, Sifriyani, 2007, Pendekatan Spline Untuk Fungsi Variansi, *Jurnal ITS*, Surabaya.