BAB III

REGRESI NONPARAMETRIK BIRESPON SPLINE

3.1. Regresi Nonparametrik Spline

Regresi spline adalah regresi dimana kurva regresinya didekati oleh fungsi spline. Fungsi spline adalah fungsi polinomial yang lebih khusus dan memberikan fleksibilitas yang lebih tinggi daripada fungsi polinomial pada umumnya. Fungsi spline yang berhubungan dengan regresi spline terkait adalah fungsi spline linear, kuadratik, dan kubik.

Secara umum, model regresi nonparametrik spline dapat ditulliskan sebagai berikut :

$$y_i = \sum_{i=0}^{p} \alpha_i \, x^i + \sum_{j=1}^{K} \beta_j \left(x - k_j\right)_+^p + \varepsilon_i \, , i = 1,2, \dots , n \qquad (3.1)$$

Dimana, p adalah derajat polinomial dan K adalah banyak titik knot pada fungsi *truncated*, serta ε_i merupakan error acak independen dengan mean nol dan varians σ^2.

3.2. Regresi Nonparametrik Birespon Spline

Dalam regresi nonparametrik spline jika terdapat satu variabel respon dan satu variabel prediktor maka diperoleh regresi nonparametrik spline univariabel. Jika dalam analisis regresi terdapat satu variabel respon dengan variabel prediktor lebih dari satu, maka regresi tersebut dinamakan regresi nonparametrik spline multivariabel. Sedangkan regresi birespon didefinisikan sebagai salah satu model regresi yang memiliki variabel respon yang lebih dari satu buah dan diantara variabel-variabel tersebut terdapat korelasi atau hubungan yang kuat, baik secara logika maupun matematis (Simillia, 2007).

Jika regresi birespon memiliki bentuk kurva regresi yang tidak diketahui, maka pendekatan yang digunakan adalah nonparametrik sehingga dikatakan regresi nonparametrik birespon. Model regresi nonparametrik birespon spline dapat dituliskan sebagai berikut :

$$y_{1j} = f_1(x_{1j}) + g_1(x_{2j}) + \varepsilon_{1j}$$
$$y_{2j} = f_2(x_{1j}) + g_2(x_{2j}) + \varepsilon_{2j}$$

Bentuk kurva regresi $f_1(x_{1j})$, $g_1(x_{2j})$, $f_2(x_{1j})$, dan $g_2(x_{2j})$ diasumsikan tidak diketahui. Error random ε_{1j} dan ε_{2j} saling berkorelasi. Dimana error random $\varepsilon_{1j}, j = 1,2,\dots,n$ saling independen dengan mean nol dan variansi σ_1^2, dan $\varepsilon_{2j}, j = 1,2,\dots,n$ saling independen dengan mean nol dan variansi σ_2^2.

Berdasarkan persamaan (2.7) yang telah diberikan pada bab sebelumnya, bentuk umum fungsi spline truncated $f(x)$ derajat p dengan K titik knot $k_1, k_2, \dots, k_K$ dan dua pediktor diberikan sebagai berikut :

$$f(x) = \sum_{i=1}^{p} \alpha_i x_1^i + \sum_{j=1}^{K} \beta_j (x_1 - k_j)_+^p + \sum_{i=1}^{p} \alpha_i{}^* x_2^i + + \sum_{j=1}^{K} \beta_j{}^* (x_2 - k_j)_+^p \quad (3.2)$$

Secara umum model regresi nonparametrik spline dengan derajat p dan K titik knot dapat ditulis sebagai berikut :

$$y_i = \sum_{i=1}^{p} \alpha_i x_1^i + \sum_{j=1}^{K} \beta_j (x_1 - k_j)_+^p + \sum_{i=1}^{p} \alpha_i{}^* x_2^i \quad + + \sum_{j=1}^{K} \beta_j{}^* (x_2 - k_j)_+^p + \varepsilon_i\ ,\ i = 1,2,\dots,n \quad (3.3)$$

Berdasarkan bentuk umum model birespon Spline linear dengan satu titik knot dari persamaan (3.3) adalah sebagai berikut :

$$y_{1j} = \alpha_1 x_{1j} + \beta_1 (x_{1j} - k_1) + \alpha_1{}^* x_{2j} + \beta_1{}^* (x_{2j} - k_2) + \varepsilon_{1j}$$
$$y_{2j} = \gamma_1 x_{1j} + \delta_1 (x_{1j} - \lambda_1) + \gamma_1{}^* x_{2j} + \delta_1{}^* (x_{2j} - \lambda_2) + \varepsilon_{2j}$$

Bentuk umum model birespon Spline linear kuadratik dengan satu titik knot dari persamaan (3.3) adalah sebagai berikut :

$$y_{1j} = \alpha_1 x_{1j} + \alpha_2 x_{1j}^2 + \beta_1 (x_{1j} - k_1)_+^2 + \alpha_1{}^* x_{2j} + \alpha_2{}^* x_{2j}^2 + \beta_1{}^* (x_{2j} - k_1{}^*)_+^2 + \varepsilon_{1j}$$
$$y_{2j} = \gamma_1 x_{1j} + \gamma_2 x_{1j}^2 + \delta_1 (x_{1j} - \lambda_1)_+^2 + \gamma_1{}^* x_{2j} + \gamma_2{}^* x_{2j}^2 + \delta_1{}^* (x_{2j} - \lambda_1^*)_+^2 + \varepsilon_{2j}$$

Untuk memudahkan, fungsi spline truncated $f(x)$ yang diasumsikan berderajat 2 dengan 2 titik, maka persamaan (3.2) menjadi sebagai berikut :

Untuk respon 1,

$$f_1(x) = \sum_{i=1}^{2} \alpha_i x_1^1 + \sum_{j=1}^{2} \beta_j (x_1 - k_j)_+^2 + \sum_{i=1}^{2} \alpha_i{}^* x_2^i + + \sum_{j=1}^{2} \beta_j{}^* (x_2 - k_j)_+^2$$

Untuk respon 2,

$$f_2(x) = \sum_{i=1}^{2} \gamma_i x_1^1 + \sum_{j=1}^{2} \delta_j \left(x_1 - \lambda_j\right)_+^2 + \sum_{i=1}^{2} \gamma_i{}^* x_2^i + + \sum_{j=1}^{2} \delta_j{}^* \left(x_2 - \lambda_j\right)_+^2$$

Jika dalam regresi birespon spline kurva regresi $f(x)$ didekati dengan fungsi spline truncated kuadratik 2 knot, maka persamaan (3.3) modelnya sebagai berikut :

$$y_{1j} = \alpha_1 x_{1j} + \alpha_2 x_{1j}^2 + \beta_1 \left(x_{1j} - k_1\right)_+^2 + \beta_2 \left(x_{1j} - k_2\right)_+^2 + \alpha_1{}^* x_{2j} + \alpha_2{}^* x_{2j}^2$$
$$+ \beta_1{}^* \left(x_{2j} - k_1{}^*\right)_+^2 + + \beta_2{}^* \left(x_{2j} - k_2{}^*\right)_+^2 + \varepsilon_{1j}$$
$$y_{2j} = \gamma_1 x_{1j} + \gamma_2 x_{1j}^2 + \delta_1 \left(x_{1j} - \lambda\right)_+^2 + \delta_2 \left(x_{1j} - \lambda_2\right)_+^2 + \gamma_1{}^* x_{2j} + \gamma_2{}^* x_{2j}^2$$
$$+ \delta_1{}^* \left(x_{2j} - \lambda_1^*\right)_+^2 + + \delta_2{}^* \left(x_{2j} - \lambda_2^*\right)_+^2 + \varepsilon_{2j}$$

Apabila diuraikan untuk setiap j=1,2,...,n, maka model Spline dalam regresi nonparametrik birepon menjadi sebagai berikut.

Untuk j=1; $y_{11} = \alpha_1 x_{11} + \alpha_2 x_{11}^2 + \beta_1 (x_{11} - k_1)_+^2 + \beta_2 (x_{11} - k_2)_+^2 + \alpha_1{}^* x_{21} + \beta_2{}^* x_{21}^2 + \beta_1{}^* (x_{21} - k_1^*)_+^2 + + \beta_2{}^* (x_{21} - k_2)_+^2 + \varepsilon_{11}$

Untuk j=2; $y_{12} = \alpha_1 x_{12} + \alpha_2 x_{12}^2 + \beta_1 (x_{12} - k_1)_+^2 + \beta_2 (x_{12} - k_2)_+^2 + \alpha_1{}^* x_{22} + \beta_2{}^* x_{22}^2 + \beta_1{}^* (x_{22} - k_1^*)_+^2 + + \beta_2{}^* (x_{22} - k_2)_+^2 + \varepsilon_{12}$

Untuk j=n; $y_{1n} = \alpha_1 x_{1n} + \alpha_2 x_{1n}^2 + \beta_1 (x_{1n} - k_1)_+^2 + \beta_2 (x_{1n} - k_2)_+^2 + \alpha_1{}^* x_{2n} + \beta_2{}^* x_{2n}^2 + \beta_1{}^* (x_{2n} - k_1^*)_+^2 + + \beta_2{}^* (x_{2n} - k_2)_+^2 + \varepsilon_{1n}$

Model di atas merupakan model pada respon 1. Berikut disajikan model Spline pada respon 2.

Untuk j=1; $y_{21} = \gamma_1 x_{11} + \gamma_2 x_{11}^2 + \delta_1 (x_{11} - \lambda_1)_+^2 + \delta_2 (x_{11} - \lambda_2)_+^2 + \gamma_1{}^* x_{21} + \gamma_2{}^* x_{21}^2 + \delta_1{}^* (x_{21} - \lambda_1^*)_+^2 + + \delta_2{}^* (x_{21} - \lambda_2^*)_+^2 + \varepsilon_{21}$

Untuk j=2; $y_{22} = \gamma_1 x_{12} + \gamma_2 x_{12}^2 + \delta_1 (x_{12} - \lambda_1)_+^2 + \delta_2 (x_{12} - \lambda_2)_+^2 + \gamma_1{}^* x_{22} + \gamma_2{}^* x_{22}^2 + \delta_1{}^* (x_{22} - \lambda_1^*)_+^2 + + \delta_2{}^* (x_{22} - \lambda_2^*)_+^2 + \varepsilon_{22}$

Untuk j=n; $y_{2n} = \gamma_1 x_{1n} + \gamma_2 x_{1n}^2 + \delta_1 (x_{1n} - \lambda_1)_+^2 + \delta_2 (x_{1n} - \lambda_2)_+^2 + \gamma_1{}^* x_{2n} + \gamma_2{}^* x_{2n}^2 + \delta_1{}^* (x_{2n} - \lambda_1^*)_+^2 + +\delta_2{}^* (x_{2n} - \lambda_2^*)_+^2 + \varepsilon_{2n}$

Berikut model regresi nonparametrik birespon spline jika disajikan dalam bentuk matriks:

$$\begin{pmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n} \\ -- \\ y_{21} \\ y_{22} \\ \vdots \\ y_{2n} \end{pmatrix} = \begin{pmatrix} f_1(x_{11}) + f_1(x_{21}) \\ f_1(x_{12}) + f_1(x_{22}) \\ \vdots \\ f_1(x_{1n}) + f_1(x_{2n}) \\ ---------- \\ f_2(x_{11}) + f_2(x_{21}) \\ f_2(x_{12}) + f_2(x_{22}) \\ \vdots \\ f_2(x_{1n}) + f_2(x_{2n}) \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1n} \\ -- \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2n} \end{pmatrix}$$

3.3. Taksiran Regresi Nonparametrik Birespon Spline

Dengan menguraikan fungsi f dan memisahkan antara parameter dengan variabel, maka persamaan tersebut dalam bentuk matrik dapat ditulis sebagai berikut :

$$y_\sim = x\beta_\sim + \varepsilon_\sim \tag{3.4}$$

Dimana:

$$y_\sim = \begin{pmatrix} y_{1\sim} \\ -- \\ y_{2\sim} \end{pmatrix} = \begin{pmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n} \\ -- \\ y_{21} \\ y_{22} \\ \vdots \\ y_{2n} \end{pmatrix}; x = \begin{pmatrix} A|O \\ -|- \\ O|B \end{pmatrix} = \beta_\sim = \begin{pmatrix} \beta_{1\sim} \\ -- \\ \beta_{2\sim} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \alpha_1{}^* \\ \alpha_2{}^* \\ \beta_1{}^* \\ \beta_2{}^* \\ -- \\ \gamma_1 \\ \gamma_2 \\ \delta_1 \\ \delta_2 \\ \gamma_1{}^* \\ \gamma_2{}^* \\ \delta_1{}^* \\ \delta_2{}^* \end{pmatrix}; \varepsilon_\sim = \begin{pmatrix} \varepsilon_{1\sim} \\ -- \\ \varepsilon_{2\sim} \end{pmatrix} = \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1n} \\ -- \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2n} \end{pmatrix}$$

Dengan :

$$A=\begin{pmatrix} x_{11} & x_{11}^{2} & (x_{11}-k_1)_+^{\ 2} & (x_{11}-k_2)_+^{\ 2} x_{21} & x_{21}^{2} & (x_{21}-k_1^{*})_+^{\ 2} & (x_{21}-k_2^{*})_+^{\ 2} \\ x_{12} & x_{12}^{2} & (x_{12}-k_1)_+^{\ 2} & (x_{12}-k_2)_+^{\ 2} x_{22} & x_{22}^{2} & (x_{22}-k_1^{*})_+^{\ 2} & (x_{22}-k_2^{*})_+^{\ 2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1n} & x_{1n}^{2} & (x_{1n}-k_1)_+^{\ 2} & (x_{1n}-k_2)_+^{\ 2} x_{2n} & x_{2n}^{2} & (x_{2n}-k_1^{*})_+^{\ 2} & (x_{2n}-k_2^{*})_+^{\ 2} \end{pmatrix}$$

$$B=\begin{pmatrix} x_{11} & x_{11}^{2} & (x_{11}-\lambda_1)_+^{\ 2} & (x_{11}-\lambda_2)_+^{\ 2} x_{21} & x_{21}^{2} & (x_{21}-\lambda_1^{*})_+^{\ 2} & (x_{21}-\lambda_2^{*})_+^{\ 2} \\ x_{12} & x_{12}^{2} & (x_{12}-\lambda_1)_+^{\ 2} & (x_{12}-\lambda_2)_+^{\ 2} x_{22} & x_{22}^{2} & (x_{22}-\lambda_1^{*})_+^{\ 2} & (x_{22}-\lambda^{*})_+^{\ 2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1n} & x_{1n}^{2} & (x_{1n}-\lambda)_+^{\ 2} & (x_{1n}-\lambda_2)_+^{\ 2} x_{2n} & x_{2n}^{2} & (x_{2n}-\lambda_1^{*})_+^{\ 2} & (x_{2n}-\lambda_2^{*})_+^{\ 2} \end{pmatrix}$$

Maka untuk memperoleh penaksir $\underset{\sim}{\beta}$ pada persamaan (3.4), dilakukan optimasi WLS yaitu dengan menyelesaikan persamaan sebagai berikut :

$$\underset{\underset{\sim}{\beta}}{min}\{\underset{\sim}{\varepsilon}' W^{-1}\underset{\sim}{\varepsilon}\} = \underset{\underset{\sim}{\beta}}{min}\{(\underset{\sim}{y}-x\underset{\sim}{\beta})' W^{-1}(\underset{\sim}{y}-x\underset{\sim}{\beta})\} \qquad (3.5)$$

Untuk menyelesaikan optimasi pada persamaan (3.5), maka dilakukan derivatif parsial. Dengan memisalkan fungsi $Q(\underset{\sim}{\beta}) = (\underset{\sim}{y}-x\underset{\sim}{\beta})' W^{-1}(\underset{\sim}{y}-x\underset{\sim}{\beta})$, maka

$$\begin{aligned} Q(\underset{\sim}{\beta}) &= (\underset{\sim}{y}-x\underset{\sim}{\beta})' W^{-1}(\underset{\sim}{y}-x\underset{\sim}{\beta}) \\ &= (\underset{\sim}{y}' - \underset{\sim}{\beta}' x')' W^{-1}(\underset{\sim}{y}-x\underset{\sim}{\beta}) \\ &= (\underset{\sim}{y}' - \underset{\sim}{\beta}' x')'(W^{-1}\underset{\sim}{y} - W^{-1}x\underset{\sim}{\beta}) \\ &= \underset{\sim}{y}' W^{-1}\underset{\sim}{y} - \underset{\sim}{\beta}' x' W^{-1}\underset{\sim}{y} - \underset{\sim}{y}' W^{-1}x\underset{\sim}{\beta} \\ &\quad + \underset{\sim}{\beta}' x' W^{-1}x\underset{\sim}{\beta} \\ &= \underset{\sim}{y}' W^{-1}\underset{\sim}{y} - 2\underset{\sim}{\beta}' x' W^{-1}\underset{\sim}{y} \\ &\quad + \underset{\sim}{\beta}' x' W^{-1}x\underset{\sim}{\beta} \qquad (3.6) \end{aligned}$$

Proses selanjutnya adalah menurunkan persamaan (3.6) terhadap $\underset{\sim}{\beta}$ dan dihasilkan :

$$\frac{\partial Q(\beta_{\sim})}{\partial \beta_{\sim}} = \frac{\partial (y'_{\sim} W^{-1} y_{\sim} - 2\beta_{\sim}' x' W^{-1} y_{\sim} \quad + \beta_{\sim}' x' W^{-1} x \beta_{\sim} \quad)}{\partial \beta_{\sim}}$$

$$= -2x' W^{-1} y_{\sim} + 2x' W^{-1} x \beta_{\sim} \qquad (3.7)$$

Setelah diturunkan terhadap $\beta_{\sim}$ hasil (3.7) disamakan dengan nol.

$$-2x' W^{-1} y_{\sim} + 2x' W^{-1} x \beta_{\sim} = 0$$

$$x' W^{-1} x \hat{\beta}_{\sim} = x' W^{-1} y_{\sim} \qquad (3.8)$$

Berdasarkan persamaan (3.8), didapatkan penaksir $\beta_{\sim}$ sebagai berikut :

$$\hat{\beta}_{\sim} = (x' W^{-1} x)^{-1} x' W^{-1} y_{\sim}$$

Sehingga bentuk taksiran model spline dalam regresi nonparametrik birespon spline menjadi sebagai berikut :

$$\hat{y}_{\sim} = (x' W^{-1} x)^{-1} x' W^{-1} y_{\sim}$$

Jika matriks $\mathrm{H}(k_{\sim}) = x(x' W^{-1} x)^{-1} x' W^{-1}$, maka diperoleh$\hat{y}_{\sim} = \mathrm{H}(k_{\sim}) y_{\sim}$. Matrik $\mathrm{H}(k_{\sim})$ merupakan fungsi dari titik knot, sedangkan $k_{\sim} = (k_1, k_2, \ldots, k_K)'$ adalah titik-titik knot.