

I'im Fatimah., 2006 *Penalised Spline Estimator Approach in Nonparametric Regression Model*. This *skripsi* under guidance of Nur Chamidah S.Si, M.Si and Drs. Suliyanto M.Si, Departement of Mathematics, Faculty of Mathematics and Natural Science Airlangga University.

ABSTRACT

Regression analysis is statistic method that is used to determine relationship between two variable or more. Suppose Y is a respon variable and X is predictor variable for n observations. Then relationship between those variables can be specified:

$$y_i = g(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n$$

$g(x_i)$ is unknown regression function, ε_i are observation errors, which assumed to be independence with mean 0 and variance σ^2 .

The purpose of this *skripsi* is to estimate regression function, $\hat{g}(x_i)$ by using *penalised spline estimator approach*. Based on the *penalised spline estimator approach*, we got the estimator of regression function is :

$$\hat{g}(x_i) = X^* \hat{\beta} \quad \text{where} \quad \hat{\beta} = (X^{*T} X^* + n\lambda D)^{-1} X^{*T} Y$$

Application result of nonparametric regression model was applied on *Canadian workers income data* and generated dat which had function: $g(x) = 2x^3 - 21x^2 + 20$. On *Canadian workers income data*, where the number of *Log income Canadian workers* is as a response variable and age of *Canadian workers* is as a predictor variable, using the estimator penalized spline approach was obtained the regression function estimator $\hat{g}(x)$:

$$\begin{aligned} \hat{g}(x_i) = & 0.00001057 + 0.00041607x_i + 0.01227698x_i^2 - 0.00014266x_i^3 + 0.00080039(x_i - 25)_+^3 \\ & + 0.00324229(x_i - 29)_+^3 + 0.00106499(x_i - 33)_+^3 - 0.00205906(x_i - 37)_+^3 \\ & + 0.00045810(x_i - 41)_+^3 + 0.00206536(x_i - 45)_+^3 - 0.00108949(x_i - 49)_+^3 \\ & - 0.00145031(x_i - 53)_+^3 + 0.00255816(x_i - 57)_+^3 - 0.00129475(x_i - 61)_+^3 \end{aligned}$$

While the application result of nonparametric regression model on generated data which have function : $g(x) = 2x^3 - 21x^2 + 20$, with error variance 60 and number of data is 100, using the estimator penalized spline approach was obtained the regression function estimator $\hat{g}(x)$:

$$\hat{g}(x_i) = 63.443 - 33.972x_i - 15.486x_i^2 + 1.745x_i^3 + 0.573(x_i - 4.023)_+^3$$

Key Word : Nonparametric Regression, *Penalised Spline Estimator*.

I'im Fatimah., 2006 *Pendekatan Estimator Penalised Spline Pada Model Regresi Nonparametrik*. Skripsi ini dibawah bimbingan Nur Chamidah S.Si, M.Si dan Drs. Suliyanto M.Si, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Airlangga.

ABSTRAK

Analisis regresi merupakan suatu metode statistika yang digunakan untuk menentukan hubungan antara dua variabel atau lebih. Misalkan Y adalah variabel respon dan X adalah variabel prediktor untuk n pengamatan, maka hubungan antara variabel-variabel tersebut dapat dinyatakan sebagai berikut :

$$y_i = g(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n$$

$g(x_i)$ merupakan fungsi regresi yang tidak diketahui, ε_i error random yang diasumsikan independent dengan mean 0 dan variansi σ^2 .

Tujuan dari skripsi ini adalah untuk mengestimasi fungsi regresi $g(x_i)$ dengan pendekatan estimator *penalised spline*. Berdasarkan pendekatan estimator *penalised spline* didapatkan bentuk estimator untuk fungsi regresi adalah

$$\hat{g}(x_i) = X^* \hat{\beta} \quad \text{dengan} \quad \hat{\beta} = (X^{*T} X^* + n\lambda D)^{-1} X^{*T} Y$$

Hasil penerapan model regresi nonparametrik dilakukan pada data pendapatan pekerja *Canadian* dan data bangkitan yang berfungsi $g(x) = 2x^3 - 21x^2 + 20$. Pada data pendapatan pekerja *Canadian*, dimana besarnya *Log income* pekerja *Canadian* sebagai variabel respon dan usia pekerja *Canadian* sebagai variabel predictor, maka dengan pendekatan estimator *penalised spline* diperoleh bentuk estimator fungsi regresi $\hat{g}(x_i)$ sebagai berikut:

$$\begin{aligned} \hat{g}(x_i) = & 0.00001057 + 0.00041607x_i + 0.01227698x_i^2 - 0.00014266x_i^3 + 0.00080039(x_i - 25)_+^3 \\ & + 0.00324229(x_i - 29)_+^3 + 0.00106499(x_i - 33)_+^3 - 0.0020590(x_i - 37)_+^3 \\ & + 0.00045810(x_i - 41)_+^3 + 0.00206536(x_i - 45)_+^3 - 0.00108949(x_i - 49)_+^3 \\ & - 0.0014503(x_i - 53)_+^3 + 0.00255816(x_i - 57)_+^3 - 0.00129475(x_i - 61)_+^3 \end{aligned}$$

Sedangkan hasil penerapan model regresi nonparametrik pada data bangkitan yang berfungsi $g(x) = 2x^3 - 21x^2 + 20$, dengan varians error 60 dan jumlah data sebanyak 100, maka dengan pendekatan estimator *penalised spline* diperoleh bentuk estimator fungsi regresi $\hat{g}(x_i)$ sebagai berikut:

$$\hat{g}(x_i) = 63.443 - 33.972x_i - 15.486x_i^2 + 1.745x_i^3 + 0.573(x_i - 4.023)_+^3$$

Kata Kunci : Regresi Nonparametrik, Estimator *Penalised Spline*.