INTISARI

Estimator Spline Dalam Regresi Nonparametrik dan Semiparametrik

Oleh
I Nyoman Budiantara

Diberikan data (t_j, y_j) dan hubungan antara t_j dan y_j diasumsikan mengikuti model regresi nonparametrik :

$$y_j = f(t_j) + \varepsilon_j,\ j = 1,2,\ldots,n.$$

Bentuk kurva regresi f tidak diketahui dan diasumsikan licin (*smooth*), dalam arti f merupakan anggota ruang Sobolev $W_2^m[0,1]$. Sesatan random ε_j berdistribusi independen dengan mean nol dan variansi σ^2/w_j, $0 < w_j < \infty$. Estimator f_λ diperoleh dengan meminimumkan *Penalized Least Square* Terbobot :

$$n^{-1}\sum_{j=1}^{n} w_j (y_j - f(t_j))^2 + \lambda \int_0^1 [f^{(m)}(t)]^2\, dt.$$

Parameter penghalus λ merupakan kontrol antara *goodness of fit* dan *penalty*. Pendekatan *Reproducing Kernel Hilbert Space* memberikan estimator f_λ berupa polinomial spline natural terbobot berordo $2m-1$.

Diberikan data (t_j, x_j, y_j) dan hubungan antara t_j, x_j dan y_j diasumsikan mengikuti model regresi semiparametrik :

$$y_j = x_j'\gamma + f(t_j) + \varepsilon_j,\ j = 1,2,\ldots,n.$$

Variabel $x_j' = (x_{j1},\ldots,x_{jp})$ dan t_j, $j = 1,2,\ldots,n$ merupakan variabel-variabel prediktor. Vektor parameter $\gamma = (\gamma_1,\ldots,\gamma_p)' \in R^p$ tidak diketahui dan f merupakan anggota ruang Sobolev $W_2^m[0,1]$. Sesatan random ε_j berdistribusi independen dengan mean nol dan variansi σ^2/w_j, $0 < w_j < \infty$. Estimator γ_λ dan f_λ diperoleh dengan meminimumkan *Penalized Least Square* Terbobot :

$$n^{-1}\sum_{j=1}^{n} w_j (y_j - x_j'\gamma - f(t_j))^2 + \lambda \int_0^1 [f^{(m)}(t)]^2\, dt.$$

Penyelesaian yang meminimumkan optimasi di atas adalah polinomial spline parsial terbobot.

Diberikan metode memilih λ optimal, sifat asimtotik, pandangan Bayes dan interval konfidensi untuk estimator spline terbobot. Diberikan pula sifat estimator spline parsial terbobot untuk $\lambda \rightarrow \infty$ dan diselidiki sifat konsisten serta distribusi asimtotik estimator komponen parametrik dalam spline parsial terbobot. Akhirnya diberikan suatu visualisasi estimator spline terbobot dan spline parsial terbobot.

ABSTRACT

Spline Estimator in Nonparametric and Semiparametric Regression

By
I Nyoman Budiantara

We consider the data (t_j,y_j) and we assume that the t_j and y_j are related by the nonparametric regression model :

$$y_j = f(t_j) + \varepsilon_j\,,\; j = 1,2,\ldots,n.$$

The function f is assumed to be in Sobolev space $W_2^m[0,1]$. The random error ε_j are independently distributed with a zero mean and with a variance σ^2/w_j, $0 < w_j < \infty$. The estimator f_λ is obtained by minimizing Weighted Penalized Least Square (WPLS) :

$$n^{-1}\sum_{j=1}^{n} w_j (y_j - f(t_j))^2 + \lambda \int_0^1 [f^{(m)}(t)]^2\, dt$$

where λ is the smoothing parameter. Under Reproducing Kernel Hilbert Space approach we obtained the solution of minimized WPLS as a weighted natural spline polynomial of order $2m-1$.

We consider the data (t_j,x_j,y_j) and we assume that the t_j, x_j and y_j are related by the semiparametric regression model :

$$y_j = x_j'\gamma + f(t_j) + \varepsilon_j\,,\; j = 1,2,\ldots,n,$$

where $x_j' = (x_{j1},\ldots,x_{jp})$ and t_j, $j = 1,2,\ldots,n$ are predictor variables. The vector parameter $\gamma = (\gamma_1,\ldots,\gamma_p)' \in R^p$ is unknown and the curve f is assumed in Sobolev space $W_2^m[0,1]$. The random errors ε_j are independently distributed with a zero mean and with a variance σ^2/w_j, $0 < w_j < \infty$. The estimator γ_λ and f_λ are obtained by minimizing WPLS :

$$n^{-1}\sum_{j=1}^{n} w_j (y_j - x_j'\gamma - f(t_j))^2 + \lambda \int_0^1 [f^{(m)}(t)]^2\, dt\,.$$

The solution of minimized WPLS above is a weighted partial spline polynomial.

Further, we give a method how to choose the smoothing parameter λ, asymptotic properties, Bayes estimator and the confidence interval for a weighted spline estimator. Also we give properties of weighted partial spline estimator, as $\lambda \rightarrow \infty$. We have investigated the consistent property and asymptotic distribution of the parametric component estimator in a weighted partial spline. Finally, we give a visualization of the weighted spline and the weighted partial spline.