

INTISARI

ESTIMASI M ROBUST UNTUK REGRESI SPLINE TERPINALTI (Studi Kasus: DATA PENGAMATAN LALU LINTAS INTERNET)

Oleh

JUSRIATY M PURBA
11/314053/PA/13755

Regresi nonparametrik merupakan analisis regresi dengan pendugaan model dilakukan berdasarkan pendekatan yang tidak terikat asumsi bentuk kurva regresi tertentu, namun dibentuk sesuai dengan informasi yang ada dalam data. Salah satu jenis fungsi yang dapat digunakan untuk menduga bentuk regresi nonparametrik adalah fungsi Huber pada Regresi Spline Terpinalti. Regresi spline terpinalti ialah salah satu metode yang sering dipakai untuk *smoothing noisy* data.

Dalam analisis data, saat data mempunyai outlier dan outlier yang ada bukan merupakan suatu kesalahan taksiran yang diperoleh dengan metode *Ordinary Least Square* akan bias karena metode OLS tidak robust terhadap outlier. Pada regresi spline, estimasi kurva regresi dapat diselesaikan dengan kuadrat terkecil terpinalti atau *Penalized Least Square*. Namun metode estimasi ini rentan terhadap kehadiran pencilan. Sehingga diperkenalkan metode *estimasi – M robust* terpinalti.

Metode estimasi *Penalized Least Square* pada regresi spline terpinalti diganti dengan metode estimasi – M robust yang mampu menangani kehadiran pencilan data. Pembobot dalam estimasi – M bergantung pada residual dan koefisien, sehingga untuk menyelesaikan masalah tersebut perlu dilakukan prosedur iterasi yang disebut *iteratively reweighted least squares* (IRLS) yang dimana sering disebut dengan *pseudo data*. Pada skripsi ini juga mempelajari bagaimana memilih secara robust parameter penalti ketika kemungkinan terdapat outlier pada data. Diberikan kriteria pemilihan parameter penalti robust berdasarkan *generalized cross –validation*. Contoh data ialah simulasi data dan data riil yang digunakan untuk menggambarkan efektivitas prosedur.

Kata kunci: Estimasi – M robust, Regresi Spline Terpinalti, Estimasi – M terpinalti, parameter pemulus, GCV, IRLS, pseudo data, regresi nonparametrik.

ABSTRACT

ROBUST – M ESTIMATION FOR PENALIZED REGRESSION SPLINE

(Case Study: Internet Traffic Data Observation)

Oleh

JUSRIATY M PURBA

11/314053/PA/13755

Nonparametric regression is a regression analysis to estimate the model is based on the approach that is not tied assuming the form of a specific regression curve, but formed in accordance with the information contained in the data. One type of function that can be used to predict the form of the nonparametric regression function is in Penalized Regression Spline with Huber. Penalized Regression Spline is one method that is often used for smoothing noisy data.

In the analysis of the data, when the data have outliers and outliers that there is not a fault estimate obtained by the method of Ordinary Least Square OLS will be biased because the method is not robust to outliers. In spline regression, regression curve estimation can be solved by the least squares penalized or Penalized Least Square. However, this estimation method is vulnerable to the presence of outliers. So the estimation method introduced - M robust penalized.

The estimation method Penalized Least Squares regression spline terpinalti replaced with the estimation method - M robust presence of outliers that can handle data. The weighting in the estimation - M depend on the residual and coefficients, so as to solve this problem is iterative procedure called iteratively reweighted least squares (IRLS) that which is often called the pseudo data. In this thesis also learn how to choose the robust parameter penalty when there might be an outlier in the data. Awarded a penalty parameter robust selection criteria based on generalized cross -validation. Sample data was simulated data and real data are used to illustrate the effectiveness of the procedure.

Keywords: Estimation - M robust, Penalized Regression Spline, estimation – M Penalized, IRLS, pseudo data, nonparametric regression.