



A Non-parametric Regression Model for Consumption of Urban Residents and Foreign Exchange Reserve

Yi Ning & Xiangdong Song

Department of Probability Theory and Mathematical Statistics, Yanshan University

Qinhuangdao 066004, China

E-mail: iamddt123@126.com

Abstract

The consumption of urban residents and the foreign exchange reserves are important indicators of economic development and between which there existed a stable relationship. Local polynomial regression estimation method of the non-parametric regression model is discussed and considering the historical data of China's consumption of urban residents and foreign exchange reserves from 1981 to 2003, which was researched using the local polynomial regression estimation method of the non-parametric regression model, we built its non-parametric regression model and compare it with the polynomial fitting model. The results showed that non-parametric regression model is superior to polynomial fitting model.

Keywords: Non-parametric regression, Local polynomial regression estimation, Polynomial fitting

The consumption of urban residents is a significant economic indicator, which reflects the economic situation of one country. With the development of China's economic reform, and the rising level of people's income and consumption at the same time, foreign trade status is having a profound impact on China's economic growth and people's lives (Xiaodong, Shen and Weiya, Zhao, 2005, 1, 18-22). Currently, as foreign trade continuously extends and domestic demand grows, the research of urban residents' consumption demand and patterns can be helpful to explain the trend of changes in consumer demand and the law of the consumer goods market, which can be useful to predict future demand for various types of consumer goods and help the relevant government departments to adjustment product structure and make reasonable arrangements for foreign trade and circulation which is of great significance for the healthy development of economy (Huiming, Zhu and Yuqi, Han, 2002, 18(5): 7-10) (YE, A-zhong, 2002, 21(1): 47). The foreign exchange reserve is also a significant economic indicator which has a stable relationship with the consumption level of urban residents under the condition of macroeconomic in China. This paper investigates the local polynomial regression estimation method of the non-parametric regression model which is used to empirically research the relationship between China's consumption of urban residents and foreign exchange reserve.

1. Nonparametric regression model and local polynomial regression estimation

1.1 nonparametric regression model

Definition 1.1: With given n pairs of observed value $(x_1, Y_1), \dots, (x_n, Y_n)$, the relationship between response variable Y and covariate variables x is defined by following equation:

$$Y_i = r(x_i) + \varepsilon_i, E(\varepsilon_i) = 0, i = 1, 2, \dots, n, \quad (1.1)$$

where r is regression function. Variable x is also known as characteristic. The estimation of $r(x)$ is represented by $\hat{r}_n(x)$.

1.2 Local polynomial estimation

Under the Nadaraya-Watson kernel estimation condition, consider an estimator $a \equiv \hat{r}_n(x)$ to minimize $\sum_{i=1}^n (Y_i - a)^2$.

Define weight function $w_i(x) = K((x_i - x)/h)$, and chose $a \equiv \hat{r}_n(x)$ to minimize the following weighted quadratic sum

$$\sum_{i=1}^n w_i(x)(Y_i - a)^2 \quad (1.2)$$

Solution obtained is

$$\hat{r}_n(x) \equiv \frac{\sum_{i=1}^n w_i(x) Y_i}{\sum_{i=1}^n w_i(x)}, \quad (1.3)$$

which is kernel regression estimation.

To improve the estimation, a p -order local polynomial is used. x is a fixed value which is used to estimate $r(x)$. For value u that belongs to the neighborhood of x , define polynomial

$$P_x(u; a) = a_0 + a_1(u-x) + \frac{a_2}{2!}(u-x)^2 + \cdots + \frac{a_p}{p!}(u-x)^p, \quad (1.4)$$

which approximates a smooth regression function $r(u)$ in the neighborhood of target value x with the following polynomial:

$$r(u) \approx P_x(u; a). \quad (1.5)$$

To estimate $\mathbf{a} = (a_0, \dots, a_p)^T$, chose $\hat{\mathbf{a}} = (\hat{a}_0, \dots, \hat{a}_p)^T$ which minimizes the following weighted quadratic sum:

$$\sum_{i=1}^n w_i(x) [Y_i - P_x(X_i; a)]^2. \quad (1.6)$$

Estimator $\hat{\mathbf{a}}$ depends on target value x . While $\hat{\mathbf{a}}(x) = (\hat{a}_0(x), \dots, \hat{a}_p(x))^T$, the local estimation of r is $\hat{r}_n(u) = P_x(u; \hat{\mathbf{a}})$.

While

$$\mathbf{X}_x = \begin{bmatrix} 1 & x_1 - x & \cdots & \frac{(x_1 - x)^p}{p!} \\ 1 & x_2 - x & \cdots & \frac{(x_2 - x)^p}{p!} \\ \vdots & \vdots & & \vdots \\ 1 & x_n - x & \cdots & \frac{(x_n - x)^p}{p!} \end{bmatrix}, \quad (1.7)$$

And $\mathbf{W}_x$ is a $n \times n$ diagonal matrix, where $w_{ij} = w_i(x)$. (1.6) could be written as

$$(\mathbf{Y} - \mathbf{X}_x \mathbf{a})^T \mathbf{W}_x (\mathbf{Y} - \mathbf{X}_x \mathbf{a}). \quad (1.8)$$

By minimizing (1.8), a weighted least square estimation is obtained

$$\hat{\mathbf{a}}(x) = (\mathbf{X}_x^T \mathbf{W}_x \mathbf{X}_x)^{-1} \mathbf{X}_x^T \mathbf{W}_x \mathbf{Y}. \quad (1.9)$$

As $\hat{r}_n(x) = \hat{a}_0(x)$ is the inner product of the first line from $(\mathbf{X}_x^T \mathbf{W}_x \mathbf{X}_x)^{-1} \mathbf{X}_x^T \mathbf{W}_x$ and $\mathbf{Y}$, then there is:

Theorem 1.1: estimation for the local polynomial regression is:

$$\hat{r}_n(x) = \sum_{i=1}^n l_i(x) Y_i, \quad (1.10)$$

where

$$\mathbf{L}(x)^T = (l_1(x), \dots, l_n(x)), \quad (1.11)$$

$$\mathbf{L}(x) = \mathbf{e}_1^T (\mathbf{X}_x^T \mathbf{W}_x \mathbf{X}_x)^{-1} \mathbf{X}_x^T \mathbf{W}_x, \quad (1.12)$$

$\mathbf{e}_1 = (1, 0, \dots, 0)^T$, and $\mathbf{X}_x$ and $\mathbf{W}_x$ were defined in (1.7).

The mean value and covariance value of this estimation are:

$$E(\hat{r}_n(x)) = \sum_{i=1}^n l_i(x) r(x_i), \quad (1.13)$$

$$\text{Var}(\hat{r}_n(x)) = \sigma^2 \sum_{i=1}^n l_i(x)^2 = \sigma^2 \|\mathbf{L}(x)\|^2. \quad (1.14)$$

Theorem 1.2: If $\hat{r}_n$ is a linear smoothing value, then the leave-one-out cross-validation score $\hat{R}(h)$ can be written as

$$\hat{R}(h) = \frac{1}{n} \sum_{i=1}^n \left[\frac{Y_i - \hat{r}_n(x_i)}{1 - L_{ii}} \right], \quad (1.15)$$

in which, $L_{ii} = l_i(x_i)$ is the i th diagonal elements of the smooth matrix.

The bandwidth h can be obtained by minimizing cross-validation formulas from theorem 1.2 (Larry Wasserman, 2008, p.57-62).

2. A non-parametric regression model of the consumption of urban residents and the foreign exchange reserve

Define $\mathbf{Y}$ as foreign exchange reserves, $\mathbf{X}$ as consumption of urban residents. The research sample is the historical data of China's foreign exchange reserves and consumption of urban residents date from 1981 to 2003. Data sources is from *Statistical Yearbook of China 2005*

2.1 Polynomial fitting model

Build 3-order polynomial fitting model based on least squares method:

$$\hat{y}_i = 0.0000000140 \quad 4x_i^3 - 0.0000204906 \quad 3x_i^2 + 0.1323783246 \quad 2x_i - 37.5053646378, \text{ where } i = 1, 2, \dots, n.$$

2.2 Local polynomial regression estimation model

Build non-parameter local polynomial regression estimation model of consumption of urban residents and foreign exchange reserves according to theorem 1.1 and theorem 1.2, where $p=3$, $h=250$. The results are shown in Figure 1, Figure 2, Table 1, and Table 2.

The fitting results of the foreign exchange reserves by polynomial fitting model and non-parameter local polynomial regression estimation model are shown in Figure 1, in which the dot-line represents the fitting curve of polynomial fitting model and the solid line represents the fitting curve of non-parameter local polynomial regression estimation model. Figure 2 shows the estimate value of the two models at x_i , in which $\square$ represents the fitting value y_i at point x_i based on polynomial fitting model, while $\circ$ represents the estimate value at point x_i based on non-parameter local polynomial regression estimation model. Symbol $*$ represents the actual value in both Figure 1 and Figure 2.

As shown in Figure 1, the fitting curve in non-parameter local polynomial regression estimation model tallies with the actual situation better than that of polynomial fitting model. In Figure 2, fitting values in both models tallies with the actual values, but the former is superior to the latter. This can be drawn from table 1 and table 2.

As can be seen from the above chart, at the beginning of China's reform and opening up, the relationship between the consumption of urban residents and foreign exchange reserves was not stable and the volatility couldn't be neglected. However, with the continuous deepening of China's reform and opening up, expanding trade with other countries, improvement of the market economic system and perfection of national policy, the relationship between the consumption of urban residents and foreign exchange reserves in more and more stable, which creates a favorable condition for our studies

Using polynomial fitting model and non-parameter local polynomial regression estimation model respectively, to predict the value of China's foreign exchange reserve in 2004, the difference between the former result and the actual result is large (error is larger than 1000) while the later is close to (error is less than 50).

3. Conclusion

In actual economic life, due to the unpredicted reasons, it is difficult to make specific assumption. With the improvement of the market economic system and increasingly close connection between China and world economy, the correlation between consumption of urban residents and foreign exchange reserve is more and more stable. Non-parametric local polynomial regression estimation model has the regression function form of arbitrariness, lower requirement than polynomial fitting model, and no need to consider the distribution of the sample, making non-parametric local polynomial regression estimation the best reflection of correlation between variables, and the errors are smaller than the polynomial fitting model.

References

- Huiming, Zhu and Yuqi, Han. (2002). An empirical analysis of the equilibrium relationship between the per capital annual income and living expenditure in urban households. *Journal of Lanzhou Commercial College*. 18 (5):7-10. (in Chinese)
- Larry, Wasserman, Translate by Xizhi, Wu. (2008). *All of Nonparametric Statistics[M]*. Beijing: Science

press.pp.57-62.(in Chinese)

Xiaodong, Shen and Weiya, Zhao. (2005). The dynamic relationship between the per capital annual income and living expenditure in urban households - an empirical analysis based on nonparametric regression model [J]. *Economic Science*.1, 18-22 (in Chinese)

Xiaoqun, He. (1997). *Regression Analysis and Economic Data Modeling [M]*. Beijing: China Renmin University Press. (Chapter3-4) (in Chinese)

Xiru, Chen and Songgui, Wang. (1987). *Modern Regression Analysis- principle method and application[M]*. Hefei: Anhui Education Press.pp.155-161 (in Chinese)

Xizhi, Wu and Zhaojun, Wang. (1996). *Nonparametric Statistical Method [M]*.Beijing: Higher Education Press,(Chapter4) (in Chinese)

YE, A-zhong. (2002). Nonparametric Regression Model of Chinese Inflation[J]. *Aplication of Statistics and Management*.21 (1):47.(in Chinese)

YE, A-zhong. (2003). *Nonparameter Econometrics[M]*. Tianjing: Nankai University Press,(Chapter6).(in Chinese)

Zinai, Li. (2000). *Econometrics [M]*. Beijing. Higher Education Press, (Chapter5).(in Chinese)

Table 1. Comparison of MSE in two models

Polynomial fitting model	103.8655
Non-parameter local polynomial regression estimation model	17.0623

Table 2. Comparison of fitting values in two models

Year	x_i	y_i	$\hat{y}_i$	$\hat{r}_n(x_i)$
1981	456.80	27.08	20.0274	54.5065
1982	471.00	69.86	21.7659	59.2710
1983	505.92	89.01	26.0406	65.9079
1984	559.44	82.20	32.5972	65.8314
1985	673.20	26.44	46.6082	47.1090
1986	798.96	20.72	62.3389	27.2955
1987	884.40	29.23	73.2534	21.9781
1988	1103.98	33.72	102.552	46.2960
1989	1210.95	55.50	117.678	80.6515
1990	1278.89	110.93	127.641	109.192
1991	1453.81	217.12	154.773	176.117
1992	1671.73	194.43	192.114	208.330
1993	2110.81	211.99	282.644	211.757
1994	2851.34	516.20	498.775	516.180
1995	3537.57	735.97	795.817	735.526
1996	3919.47	1050.49	1011.80	1059.72
1997	4185.64	1398.90	1186.98	1360.56
1998	4331.61	1449.59	1292.33	1459.37
1999	4615.91	1546.75	1517.54	1537.87
2000	4998.00	1655.74	1864.86	1666.11
2001	5309.01	2121.65	2188.30	2120.80
2002	6029.88	2864.07	3093.34	2864.07
2003	6510.94	4032.51	3830.33	4032.51

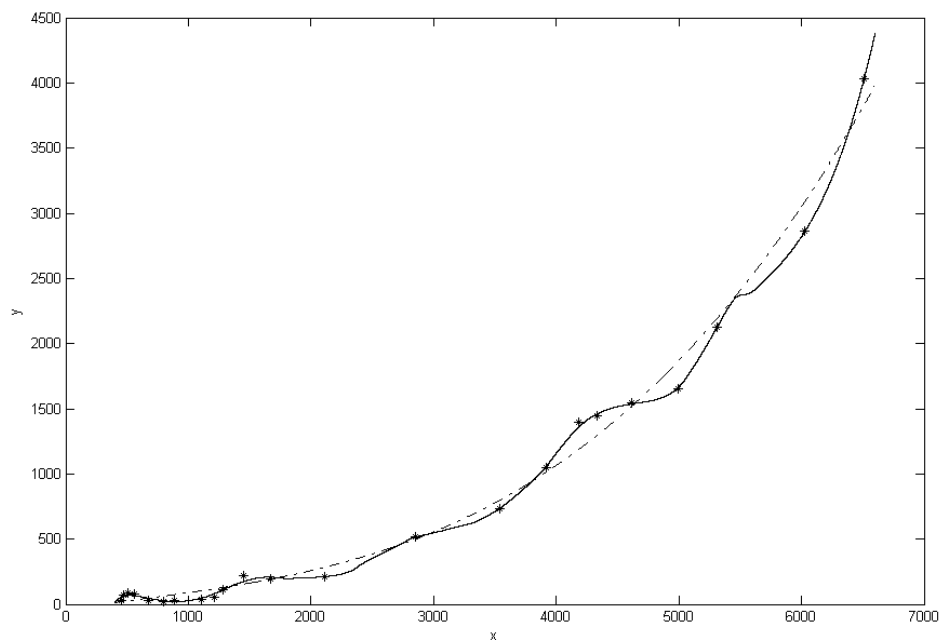


Figure 1. Fitting curve in two models

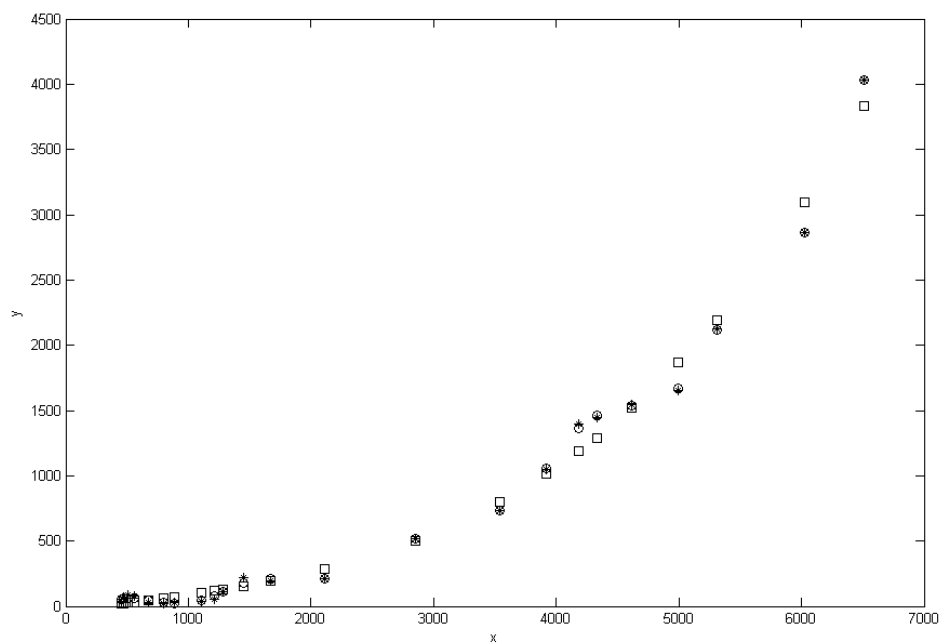


Figure 2. Fitting value in two models