行政院國家科學委員會專題研究計畫 期中進度報告

多維及相關數據的非參數估計問題(1/2)

計畫類別： 個別型計畫
計畫編號： NSC91-2118-M-002-001-
執行期間： 91 年 08 月 01 日至 92 年 07 月 31 日
執行單位： 國立臺灣大學數學系暨研究所

計畫主持人： 鄭明燕

報告類型： 精簡報告
報告附件： 國外研究心得報告
處理方式： 本計畫可公開查詢

中 華 民 國 92 年 5 月 15 日

行政院國家科學委員會補助專題研究計畫 □成果報告 ☑期中進度報告

多維及相關數據的非參數估計問題(1/2)

計畫類別：☑ 個別型計畫　　□ 整合型計畫
計畫編號：NSC　91－2118－M－002－001－
執行期間：　91 年　8 月　1 日至　92 年　7 月 31 日

計畫主持人：鄭明燕
共同主持人：
計畫參與人員：林思成、黃怡碧、洪弘

成果報告類型(依經費核定清單規定繳交)：☑精簡報告　□完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，□一年□二年後可公開

查詢

執行單位：臺灣大學數學系

中　華　民　國　　92 年　5 月　14 日

中文摘要：非參數估計的核方法已被廣泛應用於實務。區域線性迴歸有許多優點。我們研究區域線性迴歸的二次內插法，發現如此可大幅降低變異。這在估計多維迴歸曲面時特別有用。非參數濾波問題常見於工程及財務經濟。非參數濾波法通常包含一些濾波參數，這些參數可以針對每一時間點區域選擇或針對一段時間區間選擇使得表現最佳。我們提出以預測誤差選擇濾波參數，並且證明在很弱的條件下這種適濾波法與理想濾波法表現一樣好。這些方法可應用於財務經濟的輕變異估計。

關鍵詞：適濾波、自迴歸、指數平滑法、GARCH、內插法、區域線性迴歸、降變異法、輕變異

英文摘要：Kernel methods for nonparametric estimation have been widely used in practice. Local linear regression has many advantages. We study quadratic interpolation of local linear smoothers and show that it substantially enhances the stability. This is of particular value when estimating multivariate surfaces. Problems of nonparametric filtering arise frequently in engineering and financial economics. Nonparametric filters often involve some filtering parameters. These parameters can be chosen to optimize performance locally at each time point or globally over a time interval. We propose to choose the filtering parameters via minimizing the prediction error and show that, under mild conditions, the adaptive filter performs nearly as well as the ideal filter. The techniques can also be applied to volatility estimation in financial economics.

英文關鍵詞：adaptive filtering, autoregression, exponential smoothing, GARCH, interpolation, local linear regression, variance reduction, volatility

前言：Kernel methods for nonparametric estimation of curves and surfaces have been widely used in applications. Local linear regression enjoys many advantages. It is minimax optimal among all linear estimators, automatically achieves boundary corrections, adapts to both random and fixed designs, et al. See Fan (1993) and Fan and Gijbels (1996).

Problems of nonparametric filtering arise frequently in engineering, financial economics, and many other scientific disciplines. Given a time series, the nonparametric filtering problem is to dynamically predict the series based on past observations. This is a specific problem of time domain smoothing. Traditional nonparametric filters include the moving average filter, exponential smoothing, kernel smoothing, autoregressive filtering, and ARCH and GARCH filtering. The concept of volatility is associated with the notion of risks. It is very important for portfolio optimization, option pricing and management of financial risks. The problem of dynamic prediction of volatility is closely related with a filtering problem.

研究目的：Local linear techniques naturally extends to estimation of regression surfaces defined by multivariate covariates. The major difficulty is the design sparseness problem. Variance reduction methods greatly enhance the feasibility of local linear methods.

All of the nonparametric filters depend on certain filtering parameters. An interesting and challenging issue is how to choose these parameters adaptively. There are two versions of filtering parameters, local and global versions. The local version chooses the filtering parameter to optimize the performance at each time point. The global version is to set an in-sample period, then to choose the filtering parameters to optimize the performance in the in-sample period, and finally to predict the data in the out-sample period using the chosen filtering parameters. A natural criterion for choosing filtering parameters is the prediction error.

文獻探討：Jones (1989) and Deng and Chu (2000) studied linear interpolation of local linear estimates. It is found that, linear interpolation introduces extra first order bias terms and reduces variance. Overall, the asymptotic mean squared error is inflated for almost all cases. The problem of choice of filtering parameters is analogous to bandwidth selection in nonparametric smoothing. Brockmann, et al.

(1993) and Fan and Gijbels (1995) studied global and local bandwidth selection rules.

研究方法：Suppose $\{(X_1, Y_1),\ldots,(X_n, Y_n)\}$ is a sample generated from a regression model $Y = m(X) + \pounds$. The quadratic interpolation of the local linear estimator $\hat{m}(x)$ over an interval $[\sim_{2k}, \sim_{2k+2}]$ is given by $\tilde{m}(x) = \frac{r(r-1)}{2}\hat{m}(\sim_{2k}) + (1-r^2)\hat{m}(\sim_{2k+1}) + \frac{r(r+1)}{2}\hat{m}(\sim_{2k+2})$, where $x = \sim_{2k+1} + r\mu_n$ with $\mu_n = ¢h = \sim_{2k+2} - \sim_{2k+1} = \sim_{2k+1} - \sim_{2k}$. Another approach is the moving quadratic interpolation of the local linear estimator. Define the left- and right-shifted versions as

$$\tilde{m}_1(x) = \frac{1-\sqrt{2}}{4}\hat{m}\left(x - \left(\sqrt{1/2}+1\right)\mu_n\right) + \frac{1}{2}\hat{m}\left(x - \sqrt{1/2}\mu_n\right) + \frac{1+\sqrt{2}}{4}\hat{m}\left(x - \left(\sqrt{1/2}-1\right)\mu_n\right),$$
$$\tilde{m}_2(x) = \frac{1+\sqrt{2}}{4}\hat{m}\left(x + \left(\sqrt{1/2}-1\right)\mu_n\right) + \frac{1}{2}\hat{m}\left(x + \sqrt{1/2}\mu_n\right) + \frac{1-\sqrt{2}}{4}\hat{m}\left(x + \left(\sqrt{1/2}+1\right)\mu_n\right).$$

The moving averaged quadratic interpolation estimator is $\tilde{m}_a(x) = 0.5\tilde{m}_1(x) + 0.5\tilde{m}_2(x)$.

Consider a time series $Y_1,\ldots,Y_T$ which is progressively measurable with respect to a filteration $F = (F_t)$ and allows a semi-martinagle representation $Y_t = f_t + v_t\pounds_t$, where f_t and v_t are predictable and the innovation $\pounds_t$ form a standardized martingale difference. The function f_t is a trend or a drift series and v_t is the conditional standard deviation. The following are a few examples of nonparametric filters.

Example 1 *(Moving average filtering)* For every integer m, the moving average filter is defined as

$$\hat{f}_{t,m} = \frac{1}{m}\sum_{s<t} Y_s .$$

Example 2 *(Exponential smoothing)* For a positive parameter a, the exponential smoothing filter is defined by

$$\hat{f}_{t,a} = \frac{1}{1-e^{-a}}\sum_{s<t} e^{-(t-s)a} Y_s .$$

Example 3 *(Autoregression)* Suppose that the process $\{Y_t\}$ is to be approximated by an autoregressive (AR) equation $Y_t = \sim_1 Y_{t-1} + \ldots + \sim_p Y_{t-p} + v_t\pounds_t$. Denote by

$X_t = \left(Y_{t-1},\dots,Y_{t-p}\right)^T$ and $\sim = \left(\sim_1,\dots,\sim_p\right)^T$. Then the least squares estimate of the parameter $\sim$ from the past observations is

$$\hat{\sim}_{t,p} = \left(\sum_{s=t_0}^{t-1} X_{s,p} X_{s,p}^T\right)^{-1} \sum_{s=t_0}^{t-1} X_{s,p} Y_s \,,$$

and the corresponding filter is deifned as $\hat{f}_{t,p} = X_{t,p}^T \hat{\sim}_{t,p_s}$.

The performance of a filter can be measured by the sum of squared filtering errors

$$R(a) = \sum_{t=t_0}^{T} \left(f_t - \hat{f}_{t,a}\right)^2 .$$

An empirical analog of the filtering error is the prediction error

$$\circ(a) = \sum_{t=t_0}^{T} \left(Y_t - \hat{f}_{t,a}\right)^2 .$$

This criterion leads to a global data-driven selection rule $\hat{a} = \arg\inf_{a\in\Lambda} \circ(a)$. To enhance the flexibility of the family of filters to adapt to possible structure changes over time, the filtering parameter should be allowed to vary over time. For each time t, choose $\hat{a}_t = \arg\inf_{a\in\Lambda} \sum_{s=t-M+1}^{t} \left(Y_s - \hat{f}_{s,a}\right)^2$, where M is the size of the neighborhood over which we wish to optimize the performance. These methods can be applied to estimation of volatility by letting $Y_t = |R_t|^{¥}$ where R_t is the log-return of an asset or bond process.

結果與討論：We show that the quadratic interpolation estimators have the same asymptotic bias as the original local linear estimator and they have smaller asymptotic variances than the local linear estimator. The variance reduction can be as much as 60%. A simulation study was conducted to examine finite sample performances. We find that if the regression curve is relatively smooth, the effect of the variance reduction is apparent for sample sizes as small as 100. If the regression curve has sharp features, quadratic interpolation may introduce second order biases and relatively small values of $¢$ are recommended.

We show that the data-driven choices of filtering parameters give filters that perform nearly as well as the ideal filters which use the ideal filtering parameters. In a simulation study, we find that the ES global and GARCH(1,1) methods are quite robust. When the true model is GARCH(1,1), GARCH(1,1) performs the best, followed by ES global and then AR global. When the true model is GARCH(1,3) the GARCH(1,1) method and the ES global method perform nearly the same. We applied the methods to the three-month treasury bills data and the S&P 500 data. The adaptive

ES global and AR global perform better than the GARCH(1,1) method most of the time, and the relative efficiency can be as large as 1/0.696.

計劃成果自評：As a result of this project two papers, Cheng, Wu and Yen (2003) and Cheng, Fan and Spokoiny (2003), have been produced. I plan to submit the former paper to either *Statistics and Computing* or *Journal of Computational and Graphical Statistics*. The latter paper will appear in *Journal of Nonparametric Statistics*.

參考文獻：

1. Brockmann, M., Gasser, T. and Herrmann, E. (1993) Locally adaptive bandwidth choice for kernel regression estimators. *Journal of the American Statistical Association* **88**, 1302—1309.
2. Cheng, M.-Y., Wu, J.S., and Yen, J.H. (2003) Quadratic interpolation of local linear smoothing. In preparation.
3. Cheng, M.-Y., Fan, J., and Spokoiny, V. (2003) Dynamic nonparametric filtering with application to volatility estimation. *J. Nonparametric Statistics,* to appear.
4. Deng, W.S. and Chu, C.K. (2000) On study of kernel regression function polygons. *Journal of Nonparametric Statistics* **12**, 597—609.
5. Fan, J. (1993) Local linear regression smoothers and their minimax efficiencies. *Annals of Statistics* **21**, 196—216.
6. Fan J. and Gijbels, I. (1995) Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. *Journal of the Royal Statistical Society* B **57**, 371—394.
7. Fan, J. and Gijbels, I. (1996) *Local polynomial modeling and its applications*. Chapman and Hall.
8. Jones, M.C. (1989) Discretized and interpolated kernel density estimates. *Journal of the American Statistical Association* **84**, 733-741.