ممیزی سری‌های زمانی به روش هسته

راضیه دهقانیان [۱]، رحیم چینی‌پرداز [۲]، بهزاد منصوری [۳]

چکیده:

به دلیل عدم کارایی روش‌های کلاسیک ممیزی مانند خطی و درجه دوم در بسیاری از مدل‌های سری زمانی، لازم است مشاهدات سری زمانی به‌صورت دیگری رده‌بندی شوند. روش ناپارامتری، ممیزی هسته مبتنی بر استفاده از برآورد تابع چگالی هسته به جای استفاده از مقادیر واقعی آن‌ها است. مهم‌ترین مسئله در برآورد تابع چگالی هسته انتخاب مقدار مناسب پارامتر هموارکننده است. در این مقاله روش‌های مختلف برآورد تابع چگالی احتمال هسته و روش مناسب انتخاب پارامتر هموارکننده که منتهی به مقدار بهینه آن در ممیزی سری‌های زمانی می‌گردد، مورد بررسی قرار گرفته است.

واژه‌های کلیدی: فرآیند میانگین متحرک اتورگسیو، ممیزی، روش‌های ممیزی هسته، پارامتر هموارکننده.

۱ مقدمه

تحلیل ممیزی[۴] از مباحث پایه‌ای آمار چند متغیره و ممیزی سری‌های زمانی از اساسی‌ترین کاربردهای داده‌های سری زمانی است. به عنوان مثال می‌توان تحلیل ممیزی سری‌های زمانی نوار مغزی (EEG)[۵] و ممیزی بین موج‌های زلزله و انفجارهای زیر زمینی را به عنوان دو کاربرد مهم نام برد. شاموی مثال‌های گسترده‌ای در مبحث ممیزی داده‌های سری زمانی ارائه داده است. روش‌های ممیزی ناپارامتری در داده‌های سری زمانی می‌توانند با انگیزه‌های زیر مورد استفاده قرار گیرند:

۱) در اغلب روش‌های کلاسیک که برای ممیزی داده‌های سری زمانی مورد استفاده قرار گرفته‌اند، اغتشاش‌ها نرمال فرض می‌شوند، در حالی‌که در روش‌های ناپارامتری فرض غیرنرمال بودن اغتشاش‌ها مجاز است.

۲) هرگاه اندازه نمونه راهنما نسبت به بعد سری زمانی کوچک باشد، به این دلیل که اطلاعات موجود در مشاهدات راهنما[۶] برای برآورد تعداد زیادی از پارامترها کافی نیست، احتمال رده‌بندی نادرست[۷] به طور معنی‌داری افزایش می‌یابد (چینی‌پرداز و کاکس، ۲۰۰۴) .

با توجه به این‌که برای رده‌بندی معمولاً از نسبت تابع درستنمایی جامعه‌ها استفاده می‌شود، فیکس و هادچ [9] استفاده از برآورد تابع درستنمایی‌ها را به جای تابع درستنمایی‌ها پیشنهاد کردند. مهم‌ترین روش برآورد توابع درستنمایی روش هسته[۸] است اما کارایی[۹] این برآورد به شدت تحت تاثیر پارامتر هموارکننده[۱۰] است (سیلورمن). به همین دلیل روش‌های مختلفی برای برآورد تابع چگالی هسته وجود دارد که پارامترهای هموارکننده متفاوتی را نیز تولید می‌کنند. شایان ذکر است که این روش‌ها لزوماً بهترین مقدار را برای کاربردهای تابع چگالی هسته ارائه نمی‌دهند. با توجه به اهمیت پارامتر هموارکننده در ممیزی ناپارامتری، در این مقاله

[۱] کارشناس ارشد آمار-دانشگاه شهید چمران اهواز

[۲] استاد گروه آمار-دانشگاه شهید چمران اهواز

[۳] استادیار گروه آمار-دانشگاه شهید چمران اهواز

[۴] Discrimination

[۵] Electroencephalography

[۶] Training observation

[۷] Misclassification

[۸] Kernel method

[۹] Efficiency

[۱۰] Bandwidth

روش‌های مهم برآورد مورد بررسی قرار می‌گیرند، تا روش بهینه برای آنالیز ممیزی به دست آید. مقاله در هشت بخش به‌صورت زیر تنظیم شده است: در بخش دوم ممیزی ناپارامتری برای سری‌های زمانی توضیح داده شده است. در بخش سوم معیارهای تغییرپذیری برای برآورد تابع چگالی مطرح شده و روش رسیدن به پارامتر هموارکننده ارائه شده است. بخش بعدی مقاله به انتخاب پارامتر هموارکننده اختصاص دارد که در این بخش روش‌های مهم پیشنهاد شده برای پارامتر هموارکننده آورده می‌شوند. در بخش پنجم روش برآورد هسته توافقی به دلیل اهمیت آن مطرح شده است و در بخش ششم روش پیشنهادی برای آنالیز ممیزی مورد بحث قرار گرفته است. در بخش هفتم به کمک شبیه سازی روش‌های مختلف انتخاب پارامتر هموارکننده برای ممیزی مورد مقایسه قرار گرفته است. بخش نهایی مقاله نیز به بحث و نتیجه‌گیری ممیزی بهینه اختصاص یافته است.

۲ ممیزی ناپارامتری داده‌های سری زمانی

فرض کنید یک فرآیند اتورگرسیو تحت فرض $H_i (i = 1,2)$ به صورت زیر تعریف شود:

$$\begin{aligned} H_1 : x_t &= \alpha_1 x_{t-1} + \cdots + \alpha_p x_{t-p} + \varepsilon_t, \\ H_2 : x_t &= \beta_1 x_{t-1} + \cdots + \beta_p x_{t-p} + \varepsilon_t, \end{aligned}$$

که ε_t، اغتشاش خالص نرمال است. همچنین فرض کنید واریانس ε_t برای هر دو جامعه H_1 و H_2 یکسان و برابر با σ^2 بوده و طول سری برابر با n است و α_jها و β_jها در شرط ایستایی صدق می‌کنند. آن‌گاه توزیع $\mathbf{x} = (x_t, x_{t-1}, \ldots, x_{t-p})$ نرمال چندمتغیره با ماتریس کواریانس Σ به صورت زیر خواهد بود:

$$\Sigma_i = \{\sigma_i(s-r), r, s = 0, 1, \ldots, T-1\}, \quad i = 1, 2 \ ,$$

که در آن σ_i بر اساس معادلات یول - والکر و تابعی از α_iها برای H_1 و یا تابعی از β_iها برای H_2 خواهد بود (شاموی و استوفر، ۲۰۱۱). در روش کلاسیک تابع ممیزی مبتنی بر نسبت توابع درستنمایی به دو صورت خطی (اگر ماتریس واریانس کوواریانس Σ_1 و Σ_2 یکسان باشند) و یا درجه دوم (اگر ماتریس واریانس کوواریانس‌ها یکسان نباشند) از $\mathbf{x}$ به دست می‌آید (ولش ۱۹۳۹، والد ۱۹۴٤ و پنرس ۱۹٤۷). اگر مشاهدات نرمال نباشند نمی‌توان در مورد شکل تابع ممیزی اظهارنظر کرد و بنابراین خطای آماری را نمی‌توان به‌صورت تحلیلی به دست آورد. قاعده ممیزی ناپارامتری مبتنی بر استفاده از برآورد توابع درستنمایی به جای توابع درستنمایی واقعی است (فیکس و هادچ، ۱۹۵۱). بنابراین مسئله منجر به استفاده از روش‌هایی، برای برآورد توابع درستنمایی خواهد شد. از میان روش‌های برآورد تابع چگالی روش هسته از دیگر روش‌ها مهم‌تر است.

استفاده از برآورد توابع درستنمایی در ممیزی برای مشاهدات مستقل مورد استفاده آماردانان قرار گرفته است (به عنوان مثال رائو و سیلورمن). یک سؤال اساسی این است که آیا می‌توان این روش را برای داده‌های سری زمانی که کاملاً به‌هم وابسته هستند به‌کار برد. ضرایب اتوکوواریانس[۱۱]، خودهمبستگی[۱۲] و خودهمبستگی جزئی[۱۳] نمونه‌ای اطلاعات خوبی را برای ممیزی بین دو سری زمانی ارائه می‌دهند. چان ۱۹۹۱ نشان داد که لگاریتم نسبت درستنمایی برای دو فرآیند اتورگرسیو نرمال بر اساس $(p+1)$ ضریب اتوکوواریانس نمونه‌ای، یعنی $\hat{\gamma}_0, \ldots, \hat{\gamma}_p$ به‌صورت زیر محاسبه می‌شود (چینی پرداز و کاکس، ۲۰۰٤)

$$lr = -\frac{n}{2\sigma^2}\Big[\sum_{i=0}^{p}\sum_{j=0}^{p}(\beta_i\beta_j - \alpha_i\alpha_j)\hat{\gamma}_{|i-j|}\Big]. \qquad (۱)$$

ممیزی ناپارامتری سری‌های زمانی با استفاده از روش هسته بر مبنای $(p+1)$ خودهمبستگی نمونه‌ای اولین بار توسط چینی‌پرداز و کاکس مطرح شد. آن‌ها نشان دادند بهترین مرتبه برای یک فرآیند $AR(p)$، $(p+1)$ خودهمبستگی است. همچنین نشان دادند که اگر توزیع مشاهدات نرمال نباشد و همچنان ممیزی با استفادها از برآورد هسته به تابعی از ضرایب خودهمبستگی حساس است. براورد هسته تابع چگالی به تابع هسته و به مقدار زیادی به پهنای باند [۱٤] وابسته است (سیلورمن، ۱۹۸٦). چینی پرداز و کاکس ۱۹۹٦

[۱۱] Autocovariance

[۱۲] Autocorrelation

[۱۳] Partial autocorrelation

[۱٤] Bandwidth

تنها به یک روش برآورد تابع درستنمایی برای ممیزی پرداخته‌اند. در این مقاله هدف انتخاب روش یافتن پهنای باند در برآورد تابع چگالی هسته بهینه است.

۳ معیارهای تغییرپذیری برای برآورد تابع چگالی

فرض کنید x بردار p بعدی با تابع چگالی احتمال $f(x)$ باشد. تابع براورد هسته $f(x)$ با $\hat{f}(x)$ نمایش داده شده و به صورت زیر تعریف می‌شود:

$$\hat{f}(x) = \frac{1}{nh^p}\sum_{i=1}^{n} K_p\Big(\frac{x-x_i}{h}\Big).$$

در اینجا تابع K تابع هسته و h پهنای باند برآورد هسته است که پارامتر هموار کننده نیز نامیده می‌شود. مقادیر کوچک پهنای باند بیانگر نزدیکی برآورد تابع هسته به مشاهدات است اما از طرف دیگر باعث ناهمواری برآورد خواهد بود. رزنبلات ۱۹۵۶و هند ۱۹۸۱ نشان دادند که در حالت ناپارامتری برخلاف حالت پارامتری هیچ برآورد ناااریبی برای تابع چگالی وجود ندارد و باید در میان دنباله‌ای از برآوردهای به طور مجانبی نااریب بهترین برآورد را پیدا کرد.

بنابراین به دلیل اریب بودن $\hat{f}(x)$، برای بررسی میزان اختلاف بین $\hat{f}(x)$ و $f(x)$ از معیارهای میانگین مربع خطا[۱۵]، (MSE) و میانگین انتگرال مربع خطا[۱۶]، $(MISE)$ استفاده می‌شود.

اگر $\hat{f}(x)$ برآورد $f(x)$ در نقطه $x \in R$ باشد آن‌گاه:

$$MSE\Big(\hat{f}(x)\Big) = Var\Big(\hat{f}(x)\Big) + \Big(E(\hat{f}(x)) - f(x)\Big)^2. \qquad (۲)$$

حال مقادیر اریبی و واریانس محاسبه شده و در رابطه (۲) جایگذاری می‌شود. فرض می‌شود f دارای مشتق پیوسته از مرتبه دوم باشد. برای اریبی $\hat{f}(x)$ خواهیم داشت:

$$\begin{aligned} Bias\Big(\hat{f}(x)\Big) &= E\Big(\hat{f}(x)\Big) - f(x) \\ &= \int \frac{1}{h}K\Big(\frac{x-y}{h}\Big)f(y)dy - f(x) \\ &= \int K(t)f(x-ht)dt - f(x). \end{aligned}$$

با استفاده از تغییر متغیر $t = \frac{x-y}{h}$ و بسط تیلور $f(x-ht)$ حول نقطه x خواهیم داشت:

$$\begin{aligned} Bias\Big(\hat{f}(x)\Big) &= -hf'(x)\int tK(t)dt \\ &+ \frac{h^2 f''(x)}{2}\int t^2K(t)dt + o(h^2) \\ &= \frac{h^2 f''(x)}{2}\int t^2K(t)dt + o(h^2). \qquad (۳) \end{aligned}$$

به همین ترتیب برای $Var\Big(\hat{f}(x)\Big)$ خواهیم داشت:

$$Var\Big(\hat{f}(x)\Big) = \frac{f(x)}{nh}\int K^2(t)dt + o\Big((nh)^{-1}\Big). \qquad (۴)$$

اگر $\mu_2(K) = \int t^2K(t)dt$ و $R(g) = \int g^2(t)dt$ تعریف شود آن‌گاه با جایگذاری روابط (۳) و (۴) در رابطه (۲)، MSE به صورت زیر حاصل می‌شود:

$$\begin{aligned} MSE\Big(\hat{f}(x)\Big) &= \frac{R(K)}{nh}f(x) + \frac{h^4\mu_2^2(K)}{4}f''(x)^2 \\ &+ o\Big((nh)^{-1}+h^4\Big). \qquad (۵) \end{aligned}$$

با انتگرال‌گیری از رابطه (۵) خواهیم داشت:

$$MISE\Big(\hat{f}(.,h)\Big) = AMISE\Big(\hat{f}(.,h)\Big) + o\Big((nh)^{-1}+h^4\Big),$$

که در آن

$$AMISE\Big(\hat{f}(.,h)\Big) = \frac{R(K)}{nh} + \frac{h^4\mu_2^2(K)}{4}R(f''). \qquad (۶)$$

لازم به ذکر است که MSE برآورد تابع چگالی را در یک نقطه و $MISE$ برآورد آن را روی خط اعداد حقیقی در نظر می‌گیرد که برآورد اخیر با نماد $\hat{f}(.,h)$ نشان داده می‌شود.

همان‌طور که مشاهده می‌شود مربع اریبی به طور مجانبی متناسب با h^4 است، بنابراین با افزایش h، مقدار اریبی افزایش می‌یابد از طرفی واریانس به طور مجانبی معادل با $(nh)^{-1}$ است یعنی با افزایش h واریانس کاهش می‌یابد. بنابراین چنان‌چه n افزایش یابد، باید h به گونه‌ای تغییر پیدا کند که هر کدام از مولفه‌های $AMISE$ کاهش یابد. برای ایجاد تعادل، با مشتق گیری از رابطه (۶) نسبت به h و مساوی صفر قرار دادن آن خواهیم داشت:

$$h = h_{opt} = \left[\frac{R(K)}{\mu_2^2(K)R(f'')n}\right]^{\frac{1}{5}}. \qquad (۷)$$

[۱۵] Mean squared error

[۱۶] Mean integrated squared error

رابطه (۷) نشان می‌دهد که h_{opt} علاوه بر این‌که به K و n وابسته است، به طور معکوس با $R(f'')^{\frac{1}{5}}$ متناسب است. تابع $R(f'')$ مجموع انحنای f را اندازه می‌گیرد. بنابراین وقتی انحنای چگالی کم باشد $R(f'')$ کوچک و در نتیجه پارامتر هموارکننده بزرگتری لازم است و بالعکس.

با جایگذاری رابطه(۷) در رابطه (٦) خواهیم داشت:

$$\inf_{h>0} AMISE\Big(\hat{f}(.,h)\Big) = \frac{5}{4}C(K)R^{\frac{1}{5}}(f'')n^{\frac{-4}{5}}, \qquad (۸)$$

که در آن $C(K) = \mu_2^{\frac{2}{5}}(K)R^{\frac{4}{5}}(K)$. چنان‌چه $n \to \infty$، پارامتر هموارکننده حاصل از مینیمم کردن $AMISE$ و $\inf AMISE$ به صفر میل می‌کند. تابع هسته K باید به گونه‌ای اختیار شود که $C(K)$، مربوط به آن مینیمم مقدار خود را داشته باشد. اپانچنیکوف [8] نشان داد که $C(K)$ به ازای تابع $K_e(t)$ مینیمم می‌شود، که

$$K_e(t) = \begin{cases} \frac{3}{4\sqrt{5}}(1-\frac{t^2}{5}) & ,-\sqrt{5} \le t \le \sqrt{5} \\ 0 & ,\ \text{در غیر این‌صورت}\ . \end{cases} \qquad (۹)$$

حال می‌توان کارایی هر هسته متقارن را با هسته اپانچنیکوف مقایسه کرد. میزان کارایی تابع هسته K، برابر است با:

$$eff(K) = (C(K_e)/C(K))^{\frac{5}{4}} = \frac{3}{5\sqrt{5}}\ \ \mu_2^{\frac{-1}{2}}(K)R^{-1}(K).$$

٤ انتخاب پارامتر هموارکننده

پارامتر هموارکننده بهینه مینیمم کننده $AMISE$ شامل تابع نامعلوم f'' است. در این بخش روش به دست آوردن پارامتر هموارکننده، وقتی f نامعلوم است بیان می‌شود.

روش‌های انتخاب پارامتر هموارکننده به دو گروه تقسیم می‌شوند: گروه اول شامل روش‌هایی می‌شود که بر پایه اطلاعات قبلی از داده‌هاست، که به این روش‌ها اصطلاحاً برآوردکننده‌های سریع و ساده گفته می‌شود. گروه دوم روش‌هایی مبتنی بر استفاده از داده‌ها در برآورد پارامتر هموارکننده هستند و اصطلاحاً به آن‌ها روش انتخاب کاملاً خودکار و سازگار پارامتر هموارکننده[۱۷] گفته می‌شود. در اینجا به معرفی برخی از روش‌های گروه دوم که در این مقاله استفاده شده‌اند، می‌پردازیم.

۱.٤ هم‌روایی حداقل مربعات

در روش هم‌روایی حداقل مربعات[۱۸] که توسط رادمو ۱۹۸۲ و باومن ۱۹۸٤ مطرح شد، از مینیمم کردن انتگرال مربع خطا[۱۹]، (ISE) برای برآورد پارامتر هموارکننده استفاده می‌شود،

$$\begin{aligned} ISE(\hat{f}) &= \int (\hat{f}(x) - f(x))^2 dx \\ &= \int \hat{f}^2(x)dx + \int f^2(x)dx - 2\int \hat{f}(x)f(x)dx. \end{aligned}$$

چون عبارت دوم به پارامتر هموارکننده h وابسته نیست از آن صرفنظر می‌شود، بنابراین مینیمم کردن (ISE) منجر به رابطه زیر می‌شود که با $R(\hat{f})$ نشان داده می‌شود

$$R(\hat{f}) = \int \hat{f}^2(x)dx - 2\int \hat{f}(x)f(x)dx \qquad (۱۰)$$

رابطه (۸) شامل تابع مجهول f است و باید با استفاده از داده‌ها برآورد شود و سپس برآورد آن نسبت به h مینیمم شود. می‌توان نشان داد که تابع $M_0(h)$ به صورت زیر، یک برآوردکننده نااریب برای $E\Big(R(\hat{f})\Big)$ است

$$M_0(h) = \int \hat{f}^2(x)dx - \frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-i}(X_i), \qquad (۱۱)$$

که در آن $\hat{f}_{-i}(X_i)$ برآورد تابع چگالی با استفاده از تمامی مشاهدات به جز مشاهده i–ام است و به صورت زیر تعریف می‌شود:

$$\hat{f}_{-i}(X_i) = \frac{1}{(n-1)h}\sum_{\substack{j=1\\ j\neq i}}^{n} K\Big(\frac{X_i - X_j}{h}\Big).$$

چنین برآوردی را برآورد با یک مشاهده کنار گذاشته شده[۲۰] می‌نامند. از آن‌جایی که در این روش از یک قسمت از مشاهدات، برای به دست آوردن اطلاعات در مورد قسمت دیگر مشاهدات استفاده می‌شود، این روش را هم روایی نامیده‌اند. برای به دست آوردن یک رابطه روشن‌تر از $M_0(h)$، تابع $K^{(2)}$ به عنوان پیچش

[۱۷] Fully automatic and consistent bandwidth selectors

[۱۸] Least square cross-validation

[۱۹] Integrated square error

[۲۰] Leave-one-out

تابع هسته K با خودش تعریف می‌شود. به عبارت دیگر

$$\int K^{(2)}(x)dx = \int K(x-y)K(y)dy.$$

فرض می‌شود K یک تابع متقارن باشد، با تغییر متغیر $u = h^{-1}x$ خواهیم داشت:

$$\begin{aligned}\int \hat{f}^2(x)dx &= \int \frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{x-X_i}{h}\Big) \\ &\times \frac{1}{nh}\sum_{j=1}^{n} K\Big(\frac{x-X_j}{h}\Big)dx \\ &= \frac{1}{n^2h}\sum_{i=1}^{n}\sum_{j=1}^{n} K^{(2)}\Big(\frac{X_i-X_j}{h}\Big). \qquad (۱۲)\end{aligned}$$

ازطرف دیگر

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\hat{f}_{-i}(X_i) &= \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\frac{1}{h}K\Big(\frac{X_i-X_j}{h}\Big) \\ &= \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{h}K\Big(\frac{X_i-X_j}{h}\Big) \\ &- \frac{K(0)}{(n-1)h}. \qquad (۱۳)\end{aligned}$$

با جایگذاری روابط (۱۰) و (۱۱) دررابطه (۹) خواهیم داشت:

$$\begin{aligned}M_0(h) &= \frac{2}{n^2h}\Big[\frac{n}{2}K^2(0)\Big] \\ &+ \frac{2}{n^2h}\Big[\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} K^{(2)}\Big(\frac{X_j-X_i}{h}\Big) \\ &- \frac{2n}{n-1}K\Big(\frac{X_j-X_i}{h}\Big)\Big].\end{aligned}$$

پارامتر هموارکننده حاصل از مینیمم کردن $M_0(h)$، پارامتر هموارکننده بهینه است و با $\hat{h}_{LSCV}$ نشان داده می‌شود.

۲.۴ هم‌روایی ماکزیمم درستنمایی

در روش هم‌روایی ماکزیمم درستنمایی [۲۱]، اساس کار بر پایه انجام آزمون $H_0 : \hat{f}(x) = f(x)$ در مقابل $H_1 : \hat{f}(x) \neq f(x)$ با آماره آزمون نسبت درستنمایی $\frac{f(x)}{\hat{f}(x)}$ انجام می‌شود. پارامتر هموارکننده‌ای مناسب است که به ازای آن، آماره آزمون نزدیک به یک و یا $E\Big[\log\Big(\frac{f}{\hat{f}}\Big)(x)\Big]$ نزدیک صفر شود. بنابراین پارامتر هموارکننده مناسب از مینیمم کردن فاصله کولبک لایبلر [۲۲]، $d_{KL} = \int \log\Big(\frac{f}{\hat{f}}\Big)(x)f(x)dx$ به دست می‌آید. از آن‌جا که در عمل مشاهدات آزمون [۲۳] در دسترس نیستند، در هر مرحله یکی از مشاهدات موجود حذف و به عنوان مشاهده مستقل از بقیه در نظر گرفته می‌شود. این کار n مرتبه تکرار می‌شود، تابع چگالی برای مشاهده حذف شده به صورت زیر حاصل می‌شود:

$$\hat{f}_{-i}(X_i) = \frac{1}{(n-1)h}\sum_{\substack{j=1\\ j\neq i}}^{n} K\Big(\frac{X_i-X_j}{h}\Big).$$

در نتیجه تابع درستنمایی برابر مقدار

$$\prod_{i=1}^{n}\hat{f}_{-i}(X_i) = \prod_{i=1}^{n}\frac{1}{(n-1)h}\sum_{\substack{j=1\\ j\neq i}}^{n} K\Big(\frac{X_i-X_j}{h}\Big),$$

خواهد شد.

برای به دست آوردن h مناسب، لازم است لگاریتم تابع درستنمایی، ماکزیمم و یا به دلیل نسبت عکس آن با $d_{KL}(f,\hat{f})$، فاصله کولبک لایبلر مینیمم شود،

$$E\Big(CV_{KL}(h)\Big) \approx -E\Big[d_{KL}(f,\hat{f})\Big] + \int \log\Big(f(x)\Big)f(x)dx,$$

که در آن $CV_{KL}(h) = \frac{1}{n}\log\Big[\prod_{i=1}^{n}\hat{f}_{-i}(X_i)\Big]$. پارامتر هموارکننده حاصل از این روش با $\hat{h}_{CV.KL}$ نشان داده می‌شود.

۳.۴ هم‌روایی اریب

روش هم‌روایی اریب [۲۴] توسط اسکات و ترل [19] مطرح شد و بر مبنای استفاده از $MISE$ است و به اختصار با BCV نشان داده می‌شود،

$$BCV(h) = \frac{1}{nh}R(K) + \frac{h^4}{4}\mu_2^2(K)\hat{R}(f'').$$

همان‌گونه که قبلاً نشان داده شد پارامتر هموارکننده حاصل از مینیمم کردن $AMISE$ متناسب با $n^{-1/5}$ است بنابراین اگر h برای بهینه‌سازی $BCV(h)$ طوری انتخاب شود که متناسب با $n^{-1/5}$ باشد، واریانس $\hat{f}''(x)$ به ازای انتخاب $h \propto n^{-1/5}$ به صفر همگرا

[۲۱] Maximum likelihood cross-validation

[۲۲] Kullback-Leibler distance

[۲۳] Testing observation

[۲۴] Baised cross-validation

نخواهد بود. در نتیجه تابع $BCV(h)$ نمی‌تواند تقریب مناسب $MISE(\hat{f})$ باشد. به این دلیل اسکات و ترل ۱۹۸۷ از پارامتر هموارکننده ثابت h برای برآورد $\int f''^2(x)dx$ استفاده کردند. آن‌ها رابطه زیر را برای $E\Big(\int \hat{f}''^2(x)dx\Big)$ ارائه دادند:

$$E\int {\hat{f}''}^2(x)dx = \int f''^2(x)dx + \frac{1}{nh^5}\int K''^2(x)dx + O(h^2). \qquad (۱۴)$$

اریبی در رابطه (۱۲) را می‌توان به وسیله رابطه زیر تصحیح کرد:

$$\int {\hat{f}''}^2(x)dx = \int {\hat{f}''}^2(x)dx - \frac{1}{nh^5}\int K''^2(x)dx.$$

تابع هم‌روایی اریب به صورت زیر حاصل می‌شود:

$$BCV(h) = \frac{R(K)}{nh} + \frac{h^4\mu_2^2(K)}{4}\Big[R(\hat{f}'') - \frac{R(K'')}{nh^5}\Big].$$

پارامتر هموارکننده حاصل از مینیمم کردن رابطه بالا با $\hat{h}_{BCV}$ نشان داده می‌شود. اسکات و ترل ۱۹۸۷ نشان دادند که $\hat{h}_{BCV}$ به طور مجانبی ناااریب است.

با تغییر متغیر $u = \frac{x}{h}$ خواهیم داشت:

$$\int {\hat{f}''}^2(x)dx = \frac{1}{n^2h^5}\sum_{i=1}^{n}\sum_{j=1}^{n} K''^{(2)}\Big(\frac{X_j - X_i}{h}\Big).$$

از طرفی

$$K''^{(2)}(0) = \int K''(0-y)K''(y)dy = \int K''^2(y)dy.$$

در نتیجه (شیدر ۲۰۰۴):

$$BCV(h) = \frac{R(K)}{nh} + \frac{\mu_2^2(K)}{2n^2h}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} K''^{(2)}\Big(\frac{X_j - X_i}{h}\Big). \qquad (۱۵)$$

۴.۴ روش جایگزینی

همان طور که در رابطه (۷) ملاحظه شد، پارامتر هموارکننده بهینه مجانبی از طریق $\int f''^2(x)dx$ به تابع چگالی f وابستگی پیدا می‌کند بنابراین باید ابتدا این مقدار برآورد شود.

قبل از پرداختن به روش جایگزینی[۲۵] ابتدا در حالت کلی برآورد انتگرال مربع مشتق s–ام تابع چگالی مجهول f را خواهیم آورد.

فرض کنید $f^{(s)}$ نشان‌دهنده مشتق sام تابع چگالی f باشد. تعریف می‌کنیم:

$$R\Big(f^{(s)}\Big) = (-1)^s \int f^{(2s)}(x)f(x)dx.$$

توابع $\psi_r = E\Big(f^{(r)}(x)\Big)$ را به صورت :

$$\psi_r = \int f^{(r)}(x)f(x)dx, \quad r = 2n,$$

که دارای گشتاورهای فرد صفر است در نظر بگیرید. هال و مارون ۱۹۸۷ و جونز و شیدر ۱۹۹۱ نشان دادند که یک برآوردگر هسته برای ψ_r به صورت زیر است:

$$\hat{\psi}_r(g) = \frac{1}{n}\sum_{i=1}^{n}\hat{f}^{(r)}(X_i; g) = \frac{1}{n^2g}\sum_{i=1}^{n}\sum_{j=1}^{n} L^{(r)}\Big(\frac{X_i - X_j}{g}\Big), \qquad (۱۶)$$

که در آن g و L به ترتیب پارامتر هموارکننده و تابع هسته هستند که از پارامتر هموارکننده h و تابع هسته K به کار گرفته شده در برآورد f متفاوت هستند. در این‌جا نیز برای برآورد ψ_r نیاز به انتخاب پارامتر هموارکننده g است. برای این منظور از میانگین مربع خطا (MSE) استفاده می‌شود. در این‌جا فرض‌های زیر در نظر گرفته می‌شود:

۱) هسته L، یک هسته متقارن از مرتبه $k = 2, 4, \ldots$ و مشتق‌پذیر از مرتبه r–ام است به طوری‌که:

$$(-1)^{[\frac{(r+k)}{2}+1]}L^{(r)}(0)\mu_k(L) > 0,$$

که در آن $\mu_k(L) = \int x^k L(x)dx$ می‌باشد.

۲) تابع چگالی f دارای مشتق p–ام $(p > k)$ پیوسته و کران‌دار است.

۳) $g = g_n$ وها دنباله‌ای مثبت از پارامترهای هموارکننده هستند که در شرایط زیر صدق می‌کنند:

$$\lim_{n\to\infty} g = 0 \qquad \lim_{n\to\infty} ng^{2r+1} = \infty.$$

واند و جونز [26] نشان دادند که:

$$MSE\Big(\hat{\psi}_r(g)\Big) = \Big(\frac{L^{(r)}(0)}{ng^{r+1}} + \frac{g^k\mu_k(L)\psi_{r+k}}{k!}\Big)^2 + \frac{2}{n^2g^{2r+1}}R\Big(L^{(r)}\Big)\psi_0 + \frac{4}{n}\Big(\int f^{(r)}(x)^2f(x)dx - \psi_r^2\Big) + O\Big(g^{2k+2}\Big) + o\Big(n^{-2}g^{-2r-1} + n^{-1}\Big).$$

[۲۵] Plug-in

با مینیمم کردن این معیار نسبت به g خواهیم داشت:

$$g_{AMSE} = \left[\frac{k!L^{(r)}(0)}{-\mu_k(L)\psi_{r+k}n}\right]^{1/(r+k+1)}. \quad (۱۷)$$

آن‌گاه g_{AMSE} پارامتر هموارکننده بهینه مجانبی است. اما همان‌طور که دیده می‌شود g_{AMSE} از طریق ψ_{r+k} به تابع چگالی مجهول f وابسته است.

جایگزینی :

روش جایگزینی بر مبنای ایده ساده جایگزینی برآورد کمیت نامعلوم $\int f''^2(x)dx$، در رابطه (۷) می‌باشد. که در آن $R(f'') = \psi_4$ زیرا:

$$R(f'') = R\left(f^{(2)}\right) = (-1)^2 \int f^{(4)}(x)f(x)dx = \psi_4.$$

در نتیجه خواهیم داشت:

$$h_{opt} = \left[\frac{R(K)}{\mu_2^2(K)\psi_4 n}\right]^{\frac{1}{5}}. \quad (۱۸)$$

با توجه به رابطه (۱۶)، h_{opt} از طریق ψ_4 به تابع چگالی مجهول f وابسته است. بنابراین باید برآورد ψ_4، را که با $\hat{\psi}_4(g)$ نشان داده می‌شود، در رابطه (۱۶) جایگزین کرد. اگر پارامتر هموارکننده حاصل با h_{PI} نشان داده شود خواهیم داشت:

$$h_{PI} = \left[\frac{R(K)}{\mu_2^2(K)\hat{\psi}_4(g)n}\right]^{\frac{1}{5}}. \quad (۱۹)$$

این پارامتر هموارکننده به پارامتر هموارکننده اولیه g وابسته است.

۵ روش هسته توافقی

ایده اساسی روش هسته توافقی[۲۶] ساختن یک برآوردگر هسته است که سازگار با مجموع برآمدگی‌ها با هسته‌هایی است که روی داده‌ها قرار می‌گیرد، اما پارامتر هموارکننده از یک نقطه به نقطه دیگر تغییر می‌کند.

فرض کنید مشاهدات در یک فضای d بعدی قرار گرفته‌اند. در روش هسته توافقی مراحل زیر باید انجام شود:

۱) پیدا کردن برآورد مقدماتی[۲۷] $\tilde{f}(\mathbf{x})$ به طوری‌که به ازای هر i، $\tilde{f}(X_i) > 0$ باشد.

۲) در نظر گرفتن عوامل هموارکننده مکانی[۲۸] λ_i، که به‌صورت زیر تعریف می‌شوند:

$$\lambda_i = \{\tilde{f}(X_i)/g\}^{-\alpha}, \quad (۲۰)$$

که g میانگین هندسی $\tilde{f}(X_i)$ با پارامتر حساسیت $0 \leq \alpha \leq 1$ $\log g = \frac{1}{n}\sum_{i=1}^{n} \log \tilde{f}(X_i)$ است.

۳) برآورد هسته توافقی $\hat{f}(\mathbf{x})$ به صورت زیر خواهد بود:

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h^d \lambda_i^d} K(\frac{x - X_i}{h\lambda_i}), \quad (۲۱)$$

که K تابع هسته و h پارامتر هموارکننده است. همانند روش هسته با پارامتر هموارکننده ثابت، K یک تابع متقارن با انتگرال یک است (ون کرم ۲۰۰۳).

در اولین مرحله، ساختن برآوردگر مقدماتی نیازمند استفاده از روش‌های دیگر برآورد تابع چگالی مانند روش هسته با پارامتر هموارکننده ثابت یا روش نزدیکترین همسایگی است. اما بریمن و همکاران ۱۹۷۷و آبرامسون ۱۹۸۲ نشان دادند روش هسته توافقی نسبت به برآورد مقدماتی غیرحساس است. یک برآورد مقدماتی متداول، برآوردگر هسته با پارامتر هموارکننده‌ای است که با ارجاع به توزیع استاندارد به دست آمده باشد.

همان‌طور که مشاهده شد عوامل هموارکننده مکانی به α وابسته است. اگر α بزرگ باشد، حساسیت شیوه نسبت به تغییر برآورد مقدماتی زیاد می‌شود و اختلاف بیشتری بین پارامترهموارکننده مورد استفاده قرار گرفته در قسمت‌های مختلف نمونه به وجود می‌آید. اگر $\alpha = 0$ آن‌گاه روش هسته توافقی به روش هسته با پارامتر هموارکننده ثابت تبدیل می‌شود. سیلورمن ۱۹۸۶ بیان می‌کند که به ازای $\alpha = \frac{1}{2}$، روش هسته توافقی عملکرد خوبی دارد.

در مرحله پایانی، عرض هسته قرار گرفته در X_i مساوی با $\lambda_i h$ است. اگر λ_i با استفاده از رابطه (۱۸) ساخته شود آن‌گاه h، میزان همواری داده‌ها را کنترل می‌کند.

برآوردگر حاصل در رابطه (۱۹) فاقد دنباله‌های سنگین است و تمام ویژگی‌های مشتق‌پذیری خود را از تابع هسته می‌گیرد.

[۲۶] Adaptive kernel

[۲۷] Pilot estimate

[۲۸] Local bandwidth factors

۶ ممیزی داده‌های سری زمانی با استفاده از روش هسته

فرض کنید n_i $(i{=}1,2)$ سری زمانی از جامعه i ام، π_i، در دسترس است. برای هر سری زمانی $(p+1)$ خودهمبستگی نمونه‌ای برای فرآیند اتورگسیو محاسبه می‌شود. اگر مجموعه راهنما برای جامعه i، $S_i = \left\{X_1(i), X_2(i), \ldots, X_{n_i}(i)\right\}$ باشد، بردار $(p+1)$ بعدی ضرایب خودهمبستگی عبارتند از:

$$X_j{}^T(i) = \left\{\hat{\rho}_{j,1}(i), \ldots, \hat{\rho}_{j,p+1}(i)\right\}, \quad j = 1, 2, \ldots, n_i,$$

فرض کنید بردار $\hat{\rho} = (\hat{\rho}_1, \ldots, \hat{\rho}_{p+1})$ برای یک سری زمانی با جامعه نامعلوم مشاهده شده است و باید به یکی از جوامع اختصاص داده شود. از آن‌جا که در عمل سری زمانی جدید دردسترس نیست می‌توان از روش جک‌نایف[۲۹] برای ممیزی داده‌های سری زمانی برمبنای روش هسته استفاده کرد:

۱) روش با اعضا مجموعه راهنما متعلق به جامعه اول شروع می‌شود. اولین مشاهده این جامعه را در نظر گرفته و برای آن برآورد چگالی هسته به دست آورده می‌شود. برآورد چگالی هسته متناظر با این مشاهده با $\hat{h}$ نشان داده می‌شود.

۲) با استفاده از (n_1-1) مشاهده باقی مانده از جامعه اول و n_2 مشاهده از جامعه دوم برآورد چگالی هسته به دست آورده می‌شود. برآورد چگالی هسته متناظر با j امین مشاهده راهنما برای (n_1-1) و n_2 مشاهده از جامعه اول و دوم به ترتیب با $\hat{f}_{1,j}$ و $\hat{f}_{2,j}$ نشان داده می‌شود.

۳) فاصله بین برآوردگر هسته برای جامعه اول و دوم با مشاهده حذف شده به صورت زیر است:

$$\begin{aligned}
\delta_{\hat{h}_1} &= \min_j \sum_k \sum_l \left(\hat{f}_{1,j}(x_k, x_l) - \hat{h}(x_k, x_l)\right)^2, \\
\delta_{\hat{h}_2} &= \min_j \sum_k \sum_l \left(\hat{f}_{2,j}(x_k, x_l) - \hat{h}(x_k, x_l)\right)^2.
\end{aligned}$$

آن‌گاه قاعده ممیزی براساس نزدیکترین همسایگی به‌صورت زیر حاصل می‌شود:

$$\begin{aligned}
&\hat{\rho} \in \pi_1 \quad if \quad \delta_{\hat{h}_1} < \delta_{\hat{h}_2} \\
&\hat{\rho} \in \pi_2 \quad if \quad \delta_{\hat{h}_1} > \delta_{\hat{h}_2}
\end{aligned}$$

۴) مراحل ۱تا ۳ برای تمام مشاهدات از جامعه اول و جامعه دوم تکرار می‌شود، سپس برآورد احتمال خطای رده‌بندی به‌صورت $\frac{n_1^{H_1}+n_2^{H_2}}{n_1+n_2}$ محاسبه می‌شود. در اینجا $n_i^{H_i}$ تعداد مشاهداتی از جامعه i-ام اشتباه رده بندی شده است. برای فرآیند میانگین متحرک و میانگین متحرک اتورگسیو نیز به همین ترتیب عمل می‌شود.

۱.۶ انتخاب پارامتر هموارکننده برای ممیزی

همان‌طور که بیان شد از مسائل اساسی در برآورد تابع چگالی هسته انتخاب پارامتر هموارکننده است به طوری که خطای ممیزی مینیمم شود. یکی از روش‌ها برای ممیزی که در این مقاله نیز استفاده شده، در نظر گرفتن پارامتر هموارکننده یکسان برای دو جامعه است.
در این مقاله ابتدا برای هر کدام از متغیرهای جامعه اول و دوم پارامتر هموارکننده به روش‌های هم‌روایی حداقل مربعات، هم‌روایی درستنمایی ماکزیمم، هم‌روایی اریب و روش جایگزینی که به تفصیل بیان شد، انتخاب می‌شود سپس برای انتخاب یک پارامتر هموارکننده از بین پارامترهای به دست آمده به روش زیر عمل می‌شود:
هرکدام از این پارامترها را در قاعده ممیزی بیان شده قرار داده و برای هرکدام خطای ممیزی محاسبه می‌شود آن‌گاه پارامتر هموارکننده‌ای اختیار می‌شود که کمترین خطای ممیزی را دارد.
از آن‌جا که در روش هسته توافقی پارامتر هموارکننده از یک نقطه به نقطه دیگر تغییر می‌کند انتظار داریم استفاده از این روش در ممیزی سری‌های زمانی عملکرد بهتری داشته باشد. به ویژه به دلیل آن‌که در سری‌های زمانی معمولاً با داده‌های پرت مواجه هستیم. بنابراین در این مقاله برای ممیزی داده‌های سری زمانی از روش هسته توافقی نیز استفاده شده است. برآورد چگالی هسته نرمال استاندارد، که استفاده شده به صورت زیر است:

$$\hat{f}_i(\rho) = \frac{1}{n_i h_i^p} \sum_{j=1}^{n_i} K_{p+1}(\rho)$$

که در آن $K_{p+1}(x_i) = (2\pi)^{\frac{-(p+1)}{2}} \exp\left(\frac{-1}{2} x^T x\right)$ است.

[۲۹] Jaknife

۷ شبیه‌سازی

در این بخش با استفاده از شبیه‌سازی به بررسی تاثیر پارامتر هموارکننده در ممیزی ناپارامتری داده‌های سری‌های زمانی بر مبنای روش هسته پرداخته می‌شود.

ابتدا ۱۰۰ سری به طول ۲۰۰ از فرآیند اتورگسیو مرتبه اول، میانگین متحرک مرتبه اول و میانگین متحرک اتورگسیو مرتبه اول، تولید شده است. برای هر سری $(p+1)$ خودهمبستگی نمونه‌ای برای فرآیند اتورگسیو، $(p+1)$ خودهمبستگی جزئی نمونه‌ای برای فرآیند میانگین متحرک و $(p+1)$ خودهمبستگی نمونه‌ای و $(p+1)$ خودهمبستگی جزئی نمونه‌ای برای فرآیند میانگین متحرک اتورگسیو محاسبه شده است. سپس پارامتر هموارکننده با استفاده از روش‌های پیش گفته انتخاب شده و تعداد مشاهداتی که اشتباه رده بندی شده، محاسبه شده و در جدول‌ها ارائه شده است. همین روند برای برای روش توافقی نیز انجام شده است. روش هم‌روایی حداقل مربعات، هم‌روایی درستنمایی ماکزیمم، هم‌روایی اریب و روش جایگزینی به ترتیب با $LSCV$، $CV.KL$، BCV و PI نشان داده شده است. نتایج به ترتیب در جداول ۱تا ٤ ارائه است.

۸ بحث و نتیجه‌گیری

همان‌گونه که در جداول مشخص شده انتخاب پارامتر هموارکننده به روش‌های مختلف در کاهش خطای ممیزی به روش هسته موثر نیست. در حالت کلی این‌گونه تصور می‌شود که هرچه پارامتر هموارکننده کوچکتر باشد دقت بالاتر رفته و خطای ممیزی کاهش می‌یابد، اما همان‌گونه که مشاهده می‌شود در بعضی از جوامع با پارامتر هموارکننده کوچکتر خطای ممیزی بیشتری وجود دارد. روش هم‌روایی اریب عملکرد خوبی ندارد و پارامتر هموارکننده را معمولاً بزرگ اختیار می‌کند. در روش توافقی، چون پارامتر هموارکننده با نحوه پراکندگی داده‌ها متناسب است، انتظار داریم عملکرد بهتری وجود داشته باشد اما با توجه به نتایج به دست آمده این مسئله در حالت کلی صدق نمی‌کند و در اکثر جوامع خطای ممیزی بالاتری وجود دارد. با توجه به نرخ پایین رده‌بندی نادرست به نظر می‌رسد ممیزی داده‌های سری زمانی با روش ناپارامتری هسته، عملکرد مناسبی دارد. با این وجود نتایج شبیه سازی نشان می‌دهد نمی‌توان پارامتر هموارکننده بهینه را نسبت به دیگر پارامترهای برآوردکننده به دست آورد. لازم به ذکر است نتایج مطرح شده در این مقاله تنها به فرآیندهای اتورگرسیو میانگین متحرک مربوط می‌شود و نتیجه‌گیری برای دیگر فرآیندها تحقیق بیشتری را طلب می‌کند.

جدول ۱ : نرخ رده‌بندی نادرست برای یک فرآیند $AR(1)$ با استفاده از $(p+1)$ خودهمبستگی نمونه‌ای با پارامترهای مختلف

مدل‌ها	$LSCV$	$CV.KL$	BCV	PI	روش توافقی
$(-0/2, -0/3)$	48 (0/07)	48 (0/07)	48 (0/07)	48 (0/02572788) 48 (0/03616936)	45
$(-0/2, 0/2)$	1 (0/07)	1 (0/07)	1 (2)	5 (0/0275369) 5 (0/02744027) 5 (0/02699787)	5
$(-0/2, -0/5)$	1 (0/1)	4 (0/07)	3 (2)	5 (0/0275369) 5 (0/02744072) 5 (0/0307473)	6
$(0/2, 0/4)$	12 (0/07)	12 (0/07)	12 (2)	14 (0/02699787) 14 0/02476449	15
$(0/2, 0/6)$	3 (0/1)	4 (0/07)	3 (2)	2 (0/02699787) 2 (0/02548263)	3

پارامتر هموارکننده متناظر با تعداد خطای رده‌بندی داخل پرانتز آورده شده است.

جدول ۲ : نرخ رده‌بندی نادرست برای یک فرآیند $MA(1)$ با استفاده از $(p+1)$ خودهمبستگی جزئی نمونه‌ای با پارامترهای مختلف

مدل‌ها	$LSCV$	$CV.KL$	BCV	PI	روش توافقی
$(-0/2, -0/3)$	45 (0/07)	45 (0/07)	42 (2)	49 (0/03045534)	44
$(-0/2, 0/2)$	0 (0/07)	0 (0/07)	0 (0/07)	0 (0/03501655)	6
$(-0/2, -0/5)$	6 (0/07)	6 (0/07)	6 (0/07)	9 (0/03045534)	9
$(0/2, 0/4)$	11 (0/07)	11 (0/07)	11 (2)	14 (0/03501655)	14
$(0/2, 0/6)$	3 (0/07)	3 (0/07)	3 (0/07) 3 (2)	3 (0/03501655)	6

جدول ۳: نرخ رده‌بندی نادرست برای یک فرآیند $ARMA(1,1)$ با استفاده از $(p+1)$ خودهمبستگی نمونه‌ای با پارامترهای مختلف

مدل‌ها	$LSCV$	$CV.KL$	BCV	PI	روش توافقی
$(-0/2, 0/2), (0/2, 0/4)$	29 (0/07)	29 (0/07)	29 (0/07)	31 (0/04689683)	38
$(-0/2, -0/3), (-0/2, 0/2)$	35 (0/07)	35 (0/07)	35 (0/07)	36 (0/04814323)	36
$(0/2, 0/4), (0/2, 0/6)$	76 (0/07)	76 (0/07)	76 (0/07)	82 (0/04153842) 82 (0/04689683) 82 (0/04222779)	78
$(-0/2, -0/5), (-0/2, 0/2)$	25 (0/07)	25 (0/07)	25 (2)	23 (0/05317702) 23 0/0501303	26
$(0/2, 0/6), (0/2, -0/6)$	8 (0/07)	8 (0/07)	7 (0/2)	8 (0/04487126) 8 (0/04222779) 8 (0/0414056) 8 (0/05021266)	10

جدول ٤: نرخ رده‌بندی نادرست برای یک فرآیند $ARMA(1,1)$ با استفاده از $(p+1)$ خودهمبستگی جزئی نمونه‌ای با پارامترهای مختلف

مدل‌ها	$LSCV$	$CV.KL$	BCV	PI	روش توافقی
$(-0/2, 0/2), (0/2, 0/4)$	27 (0/07)	27 (0/07)	28 (2)	26 (0/03028601) 26 (0/02876342)	21
$(-0/2, -0/3), (-0/2, 0/2)$	0 (0/07)	0 (0/07)	0 (2)	0 (0/03484548)	3
$(0/2, 0/4), (0/2, 0/6)$	25 (0/07)	25 (0/07)	28 (2)	29 (0/03028601) 29 (0/02876342)	27
$(-0/2, -0/5), (-0/2, 0/2)$	0 (0/07)	0 (0/07)	0 (2)	0 (0/02452202) 0 (0/0224183 0 (0/02758993	2
$(0/2, 0/6), (0/2, -0/6)$	0 (0/07)	0 (0/07)	0 (0/07) 0 (2)	2 (0/02928186)	1

مراجع

[1] Abramson, I. S. (1982). On bandwidth variation in kernel estimates-a square root law. *Ann. Statist,* **10**, No. 4, 1217-1223.

[2] Aldershof, B. (1991). Estimation of Integrated Squared Density Derivatives. Ph. D. thesis, University of North Carolina, Chapel Hill.

[3] Bowman, A. W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika,* **71**, 353-360.

[4] Breiman, L. Meisel, M. and Purcell, E. (1977). Variable kernel estimates of multivariate densities. *Technometrics,* **19**, No. 2, 135-144.

[5] Chan, H. T. (1991). Discriminant Analysis of Time Series. Ph. D thesis, Newcastle Upon Tyne University.

[6] Chinipardaz, R. and Cox, T. (1996). Discrimination of Time Series Data. Ph. D thesis, Newcastle Upon Tyne University.

[7] Chinipardaz, R. and Cox, T. (2004). Nonparametric discrimination of time series data. *Metrika,* **50**, 13-20.

[8] Epanechnikov, V. A. (1969). Nonparametric estimation of a multidimensional probability density. *Theor. Probab. Appl,* **14**, 153-158.

[9] Fix, E. and Hodges, J. L. (1951). Discrimination analysis, nonparametric estimation: consistency properties. Report. No. 4, Randolph Field, Texas: U. S. Air Force School of Aviation Medicine, (Reprint as pp. 261-279 of Agrawala, 1977).

[10] Hall, P. and Marron, J. S. (1987). Estimation of integrated squared density derivatives. *Statist. Probab. Lett,* **6**, 106-115.

[11] Hand, D. J. (1981). *Discrimination and Classification*. Chichester: John Wiley and Sons.

[12] Jones, M. C. and Sheather, S. J. (1991). Using non-stochastic terms to advantage in kernel-based estimation of integrated squared density derivatives. *Statist. Probab. Lett,* **11**, 511-514.

[13] Park, B. U. and Marron, J. S. (1992). On the use of pilot estimators in bandwidth selection. *J. Nonparam. Statist,* **1**, No. 3, 231-240.

[14] Penrose, L. S. (1947). Some notes on discrimination. *Ann. Eug*, **13**, 228-237.

[15] Rao, C. R. (1946). Tests with discriminant functions in multivariate analysis. *Sankhya*, **7**, 407-413.

[16] Rao, C. R. (1950). Statistical inference applied to classification problems. *Sankhya*, **10**, 229-256.

[17] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist,* No. 27, 832-837.

[18] Rudemo, M. (1982). Empirical choice of histograms and kernel density estimators. *Scand. J. Statist,* No. 9, 65-78.

[19] Scott, D. W. and Terrell, G. R. (1987). Baised and unbiased cross-validation in density estimation. *J. Amer. Statist Assoc,* **82**, 1131-1146.

[20] Sheather, S. J. (2004). Density estimation. *Inst. Math. Statist*, **19**, No. 4, 588-597.

[21] Shumway, R. H. (1982). Discriminant analysis for time series. In Hanbook of Statistics,, **2**, 1-46, eds. Krishnaiah, P. R: and Kanal, L. N. Amsterdam: North-Holland.

[22] Shumway, R. H. and Stoffer, D. S. (2011). *Time Series and its Application with r example*. 3nd ed, Springer, New York.

[23] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* London: Chapman and Hall.

[24] Van Kerm, P. (2003). Adaptive kernel density estimation. *J. Stata*, **3**, No. 2, 148-156.

[25] Wald, A. (1944). On a statistical problem arising in the classification of an individual into one of two groups. *Ann. Math. Statist,* **15**, 145-162.

[26] Wand, M. P. and Jones, M. C. (1995). *Kernel Smoothing*. London: Chapman and Hall.

[27] Welch, B. L. (1939). Note on discriminant functions. *Biometrika,* **31**, 218-220.