GAMMA KERNEL ESTIMATION OF MULTIVARIATE DENSITY AND ITS DERIVATIVE ON THE NONNEGATIVE SEMI-AXIS BY DEPENDENT DATA

L. Markovich\textsuperscript{1} and A. Dobrovidov\textsuperscript{2}

\textsuperscript{1,2} Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia.

\textsuperscript{1}Corresponding author e-mail: kimo1@mail.ru
\textsuperscript{2}dobrovid@ipu.ru.

Abstract

In this paper, we consider the nonparametric estimation of the multivariate probability density function and its partial derivative with a support on $[0, \infty)$. To this end we use the class of kernel estimators with asymmetric gamma kernel functions. The gamma kernels are nonnegative. They change their shape depending on the position on the semi-axis and are robust to the boundary bias problem. We investigate the mean integrated squared error (MISE) assuming dependent data with strong mixing and find the optimal bandwidth of the kernel as a minimum of the MISE. We derive the bias, the variance and the covariance of the density and of its partial derivative.

Keywords: Density derivative; Multivariate dependent data; Gamma kernel; Nonparametric estimation.

1 Introduction

Kernel density estimators were originally introduced for univariate independent identically distributed (iid) data in [?, ?] for symmetrical kernels. The latter approach was widely adopted for tasks where support of the underlying probability density function (pdf) is unbounded. In case when the pdf has a nonsymmetric support $[0, \infty)$ the problem of the large bias on the zero boundary appears. That leads to a bad quality of the estimates in this case [?]. Obviously, the boundary bias for the multivariate pdf estimation becomes even more solid. To overcome this problem one can use special approaches such as the data reflection [?], the boundary kernels [?], the hybrid method [?], the local linear estimator [?, ?] among others. Another solution is to use the asymmetrical kernel functions instead of symmetrical ones. For univariate nonnegative iid random variables (r.v.s), the estimators with gamma kernels were proposed in [?]. The gamma kernel estimator was developed for univariate dependent data in [?]. In [?] the gamma kernel estimator of the multivariate pdf for the nonnegative iid r.v.s was introduced.

The gamma kernel is nonnegative and flexible regarding the shape. Thus, the estimators constructed with the gamma kernels of the pdfs with the nonsymmetrical support $[0, \infty)$ have no boundary bias problem. Other asymmetrical kernel estimators like inverse Gaussian and reciprocal inverse Gaussian estimators were studied in [?]. The comparison of these asymmetric kernels with the gamma kernel is given in [?].

Along with the density estimation it is often necessary to estimate the derivative of the pdf. The estimation of the univariate density derivative by the gamma kernel estimator was proposed
by authors in [?] for independent data and in [?] for the strong mixing dependent data. Our procedure achieves the optimal MISE of the order $n^{-\frac{4}{7}}$ when the optimal bandwidth is of order $n^{-\frac{4}{7}}$. In [?] an optimal MISE of the kernel estimate of the first derivative of order $n^{-\frac{4}{7}}$ corresponding to the optimal bandwidth of order $n^{-\frac{4}{7}}$ for symmetrical kernels was indicated.

In this paper, we introduce the gamma product kernel estimator for the multivariate density with the nonnegative support and its partial derivative by the multivariate dependent data with a strong mixing. Such estimator has no weight at the negative axis, so it is robust to the boundary bias problem. The asymptotical behavior of the estimates and the optimal bandwidths in the sense of minimal MISE are obtained.

The pdf derivative could be useful to find the slope of the pdf curve, its local extremes, significant features in data as well as in regression analysis [?]. The probability density derivative also plays a key role in clustering via mode seeking [?]. Furthermore, we are able to write down the estimator for the logarithmic derivative of the nonnegative pdf for the multivariate dependent data. The logarithmic derivative of the pdf may be used in the generalized equation of the optimal filtering of signal processing [?].

The outline of the paper is as follows. In Section 2 we define the kernel estimators of the multivariate density and its partial derivative. In Section 3 we obtain the bias, the variance and the covariance for the pdf estimate. Using these results we derive the rate of the optimal MISE of the estimator for the logarithmic derivative of the nonnegative pdf for the multivariate dependent data. The logarithmic derivative of the pdf is defined on the bounded support and its partial derivative by the multivariate dependent data with a strong mixing. Such estimator has no weight at the negative axis, so it is robust to the boundary bias problem. The asymptotical behavior of the estimates and the optimal bandwidths in the sense of minimal MISE are obtained.

The outline of the paper is as follows. In Section 2 we define the kernel estimators of the multivariate density and its partial derivative. In Section 3 we obtain the bias, the variance and the covariance for the pdf estimate. Using these results we derive the rate of the optimal MISE and the corresponding bandwidth. In Section 4 we obtain the same for the estimate of the partial derivative of the pdf.

2 Gamma Kernel Estimation

Let $\{X_i = (X_{i1}, \ldots, X_{id})^T\}_{i=1}^n$ be a strongly stationary sequence of $d$-dimensional variables with an unknown probability density function $f(x_1^T, x_2^T, \ldots, x_n^T)$ which is defined on the bounded support $x_i^d \in \mathbb{R}^{+d}$. We assume that the sequence $\{X_s\}$ is $\alpha-$mixing with coefficient

$$\alpha(k) = \sup_{t} \sup_{A \in n(X_s \leq t)} |P(A \cap B) - P(A)P(B)|, \quad k \geq 1.$$  

Here $\alpha(k) \to 0$ as $k \to \infty$. For these sequences we will use a notation $\{X_s\} \in \mathcal{S}(\alpha)$. To estimate the unknown pdf the product gamma kernel estimator was proposed in [?]

$$\hat{f}(x_1^T) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d K_{\rho(x_j, b_j), b_j} (X_{ij}),$$

where $b \equiv \{b_j\}_{j=1}^d$ are the bandwidth parameters such that $b \to 0$ as $n \to \infty$. The gamma kernel

$$K_{\rho(x,b), b}(t) = \frac{t^{\rho(x,b)-1} \exp(-t/b)}{b^{\rho(x,b)} \Gamma(\rho(x,b))}$$

is used for each variable [?], where $\Gamma(\cdot)$ is a standard gamma function and parameter $\rho(x, b)$ is defined as

$$\rho(x,b) = \begin{cases} \rho_1(x, b) = x/b, & \text{if} \quad x \geq 2b, \\ \rho_2(x, b) = \left(\frac{x}{2b}\right)^2 + 1, & \text{if} \quad x \in [0, 2b). \end{cases}$$

Furthermore, we shall estimate the pdf $f(x_{n-\tau}^T)$, where $0 \leq \tau \leq n - 1$. Hence (1) can be rewritten as

$$\hat{f}(x_{n-\tau}^T) = \frac{1}{n} \sum_{i=1}^n \prod_{j=n-\tau} K_{\rho(x_j, b_j), b_j} (X_{ij}).$$
Using (3) the pdf estimator can be written as
\[
\hat{f}(x^n_{n-x}) = \begin{cases} \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{n} K_{\rho_1(x_j, b_j), b_j} (X_{ij}) , & \text{if } x \geq 2b, \\ \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{n} K_{\rho_2(x_j, b_j), b_j} (X_{ij}) , & \text{if } x \in [0, 2b). \end{cases}
\]
(4)

In [9], [10] the derivative of the univariate pdf was estimated just like the derivative of the gamma kernel estimator. For the multivariate case, we can analogically estimate any partial derivative of \(f(x^n_{n-x})\). For example,
\[
\hat{f}'_x(x^n_{n-x}) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{n} K_{\rho(x_j, b_j), b_j} (X_{ij})(K_{\rho_0(x_n, b_n), b_n} (X_{in}))'_{x_n}
\]
holds, where
\[
(K_{\rho(x,b), b}(t))'_{x_n} = \begin{cases} K'_{\rho_1(x,b), b}(t) = \frac{1}{b} K_{\rho_1(x,b), b}(t)L_1(t, x, b), & \text{if } x \geq 2b, \\ K'_{\rho_2(x,b), b}(t) = \frac{2b}{K_{\rho_2(x,b), b}(t)}L_2(t, x, b), & \text{if } x \in [0, 2b), \end{cases}
\]
(6)
is the partial derivative of (2) and
\[
L_i(t, x, b) = \ln t - \ln b - \Psi(\rho_i(x, b)), \quad i = 1, 2.
\]
(7)

Here, \(\Psi(x)\) denotes the Digamma function that is the logarithmic derivative of the gamma function. Formula (5) can be rewritten as
\[
\hat{f}'_x(x^n_{n-x}) = \begin{cases} \frac{1}{n} \sum_{i=1}^{n} \frac{1}{b_n} L_1(X_{in}, x_n, b_n) \prod_{j=1}^{n} K_{\rho_1(x_j, b_j), b_j} (X_{ij}) , & \text{if } x \geq 2b, \\ \frac{1}{n} \sum_{i=1}^{n} \frac{x_i}{2b_n} L_2(X_{in}, x_n, b_n) \prod_{j=1}^{n} K_{\rho_2(x_j, b_j), b_j} (X_{ij}) , & \text{if } x \in [0, 2b). \end{cases}
\]
(8)

We shall study asymptotic properties of the estimators (4) and (8). The unknown smoothing parameters for (4) and (8) are obtained as minima of the MISEs which are defined by
\[
MISE(\hat{f}(x^n_{n-x})) = \frac{1}{n} \int (f(x^n_{n-x}) - \hat{f}(x^n_{n-x}))^2 dx,
\]
(9)
\[
MISE(\hat{f}'_x(x^n_{n-x})) = \frac{1}{n} \int (f'_x(x^n_{n-x}) - \hat{f}'_x(x^n_{n-x}))^2 dx.
\]
(10)
The integrals in (9) can be split into two integrals \(\int_0^{2b}\) and \(\int_0^{2b}\). Further we shall do all the proofs for the case when \(x \geq 2b\) because the integral \(\int_0^{2b}\) tends to zero when \(b \to 0\). Hence, we omit the indices for \(\rho_1(x, b_j), L_1(X_{ik}, x_k, b_k)\) and use instead \(\rho(x, b_j), L(X_{ik}, x_k, b_k)\).

It is known that the mean squared error (MSE) for dependent r.v.s is determined as
\[
MSE(\hat{f}(x^n_{n-x})) = (Bias(\hat{f}(x^n_{n-x}))^2 + Var(\hat{f}(x^n_{n-x})),
\]
MSE(\hat{f}'_x(x^n_{n-x})) = (Bias(\hat{f}'_x(x^n_{n-x}))^2 + Var(\hat{f}'_x(x^n_{n-x})),
\]
where the variances are
\[ \text{Var}(\hat{f}(x_{n-\tau}^n)) = \text{Var}\left( \frac{1}{n} \sum_{i=1}^{n} \prod_{j=n-\tau} K_{\rho(x_j, b_j)}(X_{ij}) \right) \]
\[ = \frac{1}{n^2} \left( \sum_{i=1}^{n} \text{Var}\left( \tilde{K}(X_i, x, b) \right) + \sum_{i,k=1, i \neq k}^{n} \text{Cov}\left( \tilde{K}(X_i, x, b), \tilde{K}(X_k, x, b) \right) \right) \]
\[ = \frac{1}{n} \text{var}(\hat{f}(x_{n-\tau}^n)) + \frac{2}{n} \sum_{i=1}^{n} \left( 1 - \frac{i}{n} \right) \text{Cov}\left( \tilde{K}(X_i, x, b), \tilde{K}(X_{1+i}, x, b) \right), \]
\[ \text{Var}(\hat{f}_{x_n}(x_{n-\tau}^n)) = \text{Var}\left( \frac{1}{n} \sum_{i=1}^{n} \frac{1}{b_n} L(X_{in}, x_n, b_n) \prod_{j=n-\tau}^{n} K_{\rho(x_j, b_j)}(X_{ij}) \right) \]
\[ = \frac{1}{n^2 b_n^2} \left( \sum_{i=1}^{n} \text{Var}\left( \tilde{K}(X_i, x, b) \right) + \sum_{i,k=1, i \neq k}^{n} \text{Cov}\left( \tilde{K}(X_i, x, b), \tilde{K}(X_k, x, b) \right) \right) \]
\[ = \frac{1}{nb_n^2} \text{var}(\hat{f}_{x_n}(x_{n-\tau}^n)) + \frac{2}{n b_n^2} \sum_{i=1}^{n} \left( 1 - \frac{i}{n} \right) \text{Cov}\left( \tilde{K}(X_1, x, b), \tilde{K}(X_{1+i}, x, b) \right). \]

Here we use the following notations
\[ \tilde{K}(X_i, x, b) = \prod_{j=n-\tau}^{n} K_{\rho(x_j, b_j)}(X_{ij}), \]
\[ \tilde{K}(X_i, x, b) = L(X_{in}, x_n, b_n) \prod_{j=n-\tau}^{n} K_{\rho(x_j, b_j)}(X_{ij}). \]

3 Convergence rate of the density estimator

In this section we obtain the asymptotic properties of the estimator (4). To this end we derive the bias, the variance and the covariance determined in (10) in the following lemmas. All proofs hold assuming that all the bandwidths are different. For the practical use we give the simpler formulation of lemmas for the equal bandwidths \( b_1 = b_2 = \ldots = b_n = b \).

**Lemma 3.1.** If \( b_1 = b_2 = \ldots = b_n = b \) and \( b \to 0 \) as \( n \to \infty \), then the bias for the pdf estimate (4) is equal to
\[ \text{Bias}(\hat{f}(x_{n-\tau}^n)) = \frac{b}{2} \sum_{j=n-\tau}^{n} x_j \frac{\partial^2 f(x_{n-\tau}^n)}{\partial x_j^2} + o((\tau + 1)b). \]

**Lemma 3.2.** If \( b_1 = b_2 = \ldots = b_n = b \) and \( b \to 0, nb^{2+1} \to \infty \) as \( n \to \infty \), then the variance expansion for the pdf estimate (4) is equal to
\[ \text{Var}(\hat{f}(x_{n-\tau}^n)) = \frac{b^{(\tau+1)/2}}{n} \left( \prod_{j=n-\tau}^{n} x_j^{-1/2} x_{n-\tau}^{1/2} \right) \left( f(x_{n-\tau}^n) + bv_1(x) + b^2 v_2(x) \right) \]
\[ - \frac{1}{n} \left( f(x_{n-\tau}^n) + \frac{1}{2} \sum_{j=n-\tau}^{n} x_j b \frac{\partial^2 f}{\partial x_j^2} \right)^2 + o((\tau + 1)b^2). \]
where
\[ v_1(x) = \sum_{j=n-\tau}^{n} \left( -\frac{1}{2} \frac{\partial f(x_{n-\tau})}{\partial x_j} + \frac{x_j}{4} \frac{\partial^2 f(x_{n-\tau})}{\partial x_j^2} \right), \]
\[ v_2(x) = -\sum_{j=n-\tau}^{n} \sum_{i=n-\tau}^{n} \frac{x_j}{8} \frac{\partial^3 f(x_{n-\tau})}{\partial x_j^2 \partial x_i}. \]

Lemma 3.3. Let
1. \( \{X_j\}_{j\geq 1} \in S(\alpha) \) and \( \int_1^\infty \alpha(\tau)^v d\tau < \infty, \quad 0 < v < 1 \) hold,
2. \( f(x_{n-\tau}) \) be a twice continuously differentiable function,
3. \( b \to 0 \) and \( nb(\tau+1)^{\frac{v+1}{2}} \to \infty \) as \( n \to \infty \).

Then, the covariance is bounded by
\[
|C(\hat{f}(x_{n-\tau}))| \leq \frac{D(v, x)}{n} \frac{b^{-(\tau+1)^{\frac{v+1}{2}}}}{n} \left( bS(v, x) + f(x_{n-\tau}) \frac{3v-1}{2(v-1)} + o(b^2) \right)^{1-v} \int_1^\infty \alpha(\tau)^v d\tau,
\]
where
\[ S(v, x) = \sum_{i=n-\tau}^{n} \frac{v+1}{(v-1)^2 x_i} f(x_{n-\tau}) + \frac{v+1}{v-1} \frac{\partial f(x_{n-\tau})}{\partial x_i} + \frac{x_i}{2} \frac{\partial^2 f(x_{n-\tau})}{\partial x_i^2}, \]
\[ D(v, x) = 2(2\pi)^{-\frac{(v+1)^{\frac{v+1}{2}}}{2}} \left( \prod_{j=n-\tau}^{n} x_j^{\frac{v+1}{2}} \right). \]

The proofs of Lemmas 3.1 - 3.3 are given in Appendices A.1, A.2, B. Using the results of the latter the MISE (9) can be written. Hence, the following theorem can be proved.

Theorem 3.1. In the conditions of Lemmas 3.1 - 3.3 the optimal bandwidth that provides minimum of the MISE is
\[
b = \left( \frac{(\tau+1)(v+1)(3v-1)^{1-v}}{n(2v-2)^{1-v}} \int D(v, x) f(x_{n-\tau})^{1-v} dx_{n-\tau} \right)^{\frac{(\tau+1)^{\frac{v+1}{2}}}{2}} \int \alpha(\tau)^v d\tau \int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{1}\int \frac{dx_{n-\tau}}{
Lemma 4.1. If the bias, the variance and the covariance determined in (10)are derived in the following lemmas.

Similarly to Section 3 we obtain the asymptotic properties of the density derivative estimator (8).

Theorem 3.2. In conditions of Lemmas 3.1, 3.2, 3.4 the optimal bandwidth which provides minimum of the MISE is

\[ b = \left( \tau + 1 \right) \frac{\int_0^\infty \left( \prod_{j=n-\tau}^{\infty} x_j^{-\frac{1}{2}} \right) f(x_{n-\tau}) dx_{n-\tau}}{\int_0^\infty \left( \sum_{j=n-\tau}^{\infty} x_j \frac{\partial^2 f(x_{n-\tau})}{\partial x_j^2} \right)^2 dx_{n-\tau}} \]

The proof is given in Appendix B.3.

Corollary 3.1. The result of Theorem 3.2 coincides with the results of [?] for univariate pdf and iid case when \( \tau = 1 \).

4 Convergence rate of the density derivative estimator

Similarly to Section 3 we obtain the asymptotic properties of the density derivative estimator (8). The bias, the variance and the covariance determined in (10) are derived in the following lemmas.

Lemma 4.1. If \( b_1 = b_2 = \ldots = b_n = b \) and \( b \to 0 \) as \( n \to \infty \), then the bias for the density derivative estimate (8) is equal to

\[ \text{Bias}(\hat{f}_{x_n}^n(x_{n-\tau})) = bB_1(x) + b^2B_2(x) + o((\tau + 1)b), \]

where we denote

\[ B_1(x) = \frac{f(x_{n-\tau})}{12x_n^2} + \frac{1}{4x_n} \sum_{j=n-\tau}^{n} x_j \frac{\partial^2 f(x_{n-\tau})}{\partial x_j^2}, \]

\[ B_2(x) = \frac{1}{24x_n^2} \sum_{j=n-\tau}^{n} x_j \frac{\partial^2 f(x_{n-\tau})}{\partial x_j^2}. \]
Lemma 4.2. If $b_1 = b_2 = \ldots = b_n = b$ and $b \to 0$, $n b^{\frac{\tau+1}{\tau}} \to \infty$ as $n \to \infty$, then the variance of the density derivative estimate \ref{6} is equal to

$$\text{var}(\hat{f}_{x_n}(x_{n-\tau}^{\tau})) = \frac{b^{\frac{(\tau+1)}{2}}}{n} \left( \prod_{j=n-\tau}^{n} \frac{x_j^{-1/2}}{2\sqrt{\pi}} \right) \left( bV_1(x) + b^2V_2(x) + \frac{1}{b}V_3(x) + V_4(x) \right)$$

$$- \frac{1}{n} \left( b^2B_2^2(x) + \left( \frac{\partial f(x_{n-\tau}^{\tau})}{\partial x_n} \right)^2 + 2 \frac{\partial f(x_{n-\tau}^{\tau})}{\partial x_n} \left( bB_1(x) + b^2B_2(x) + o((\tau+1)b^2) \right) \right),$$

where

$$V_1(x) = -\frac{1}{24x_n^2} \frac{\partial f(x_{n-\tau}^{\tau})}{\partial x_n} + \sum_{j=n-\tau}^{n} \left( \frac{1}{8x_n} \frac{\partial^2 f(x_{n-\tau}^{\tau})}{\partial x_j^2} - \frac{1}{8x_n^2} \frac{\partial f(x_{n-\tau}^{\tau})}{\partial x_j} \right) + \frac{7}{48x_n^3} f(x_{n-\tau}^{\tau}),$$

$$V_2(x) = \frac{7}{576x_n^4} f(x_{n-\tau}^{\tau}) + \sum_{j=n-\tau}^{n} \left( \frac{1}{16x_n^2} \frac{\partial^2 f(x_{n-\tau}^{\tau})}{\partial x_j^2} - \frac{7}{96x_n^3} \frac{\partial f(x_{n-\tau}^{\tau})}{\partial x_j} + \frac{1}{48x_n^2} \frac{\partial^2 f(x_{n-\tau}^{\tau})}{\partial x_n \partial x_j} \right),$$

$$V_3(x) = \frac{f(x_{n-\tau}^{\tau})}{2x_n}, \quad V_4(x) = \frac{f(x_{n-\tau}^{\tau})}{4x_n^2} - \frac{1}{4x_n} \frac{\partial f(x_{n-\tau}^{\tau})}{\partial x_j}.$$

Lemma 4.3. Let

1. $\{X_j\}_{j \geq 1} \in S(\alpha)$ and $\int_1^{\infty} \alpha(\tau)^v d\tau < \infty, \quad 0 < v < 1$ hold,

2. $f(x_{n-\tau}^{\tau})$ be a twice continuously differentiable function,

3. $b \to 0$ and $nb^{(\tau+1)-\frac{\tau+1}{\tau}} \to \infty$ as $n \to \infty$.

Then the covariance of the estimate of the pdf partial derivative is bounded by

$$|C(\hat{f}_{x_n}(x_{n-\tau}^{\tau}))| \leq \frac{R(v,x)}{nb^{(\tau+1)-\frac{\tau+1}{2}}} \left( b^2V(v,x) + bW(v,x) + R(v,x) + o(b^2) \right)^{1-v} \int_1^{\infty} \alpha(\tau)^v d\tau,$$

where we denote

$$V(v,x) = \sum_{i=n-\tau}^{n} \left( \frac{(v+1)(3v-1)}{72(v-1)^3x_n^2} + \frac{v+1}{(v-1)^2x_i} - \frac{v(v+1)}{9(v-1)^4x_n^3} \right) f(x_{n-\tau}^{\tau})$$

$$+ \frac{v+1}{v-1} \frac{\partial f(x_{n-\tau}^{\tau})}{\partial x_i} - \frac{v(v+1)}{9(v-1)^3x_n} \frac{\partial f(x_{n-\tau}^{\tau})}{\partial x_n} + x_i \frac{\partial^2 f(x_{n-\tau}^{\tau})}{\partial x_i^2},$$

$$W(v,x) = \sum_{i=n-\tau}^{n} \left( \frac{3v-1}{4(v-1)x_n} + \frac{v+1}{(v-1)^2x_i} + \frac{2x_i(v+1)}{3(v-1)^3x_n^2} \right) f(x_{n-\tau}^{\tau})$$

$$+ \frac{v+1}{v-1} \frac{\partial f(x_{n-\tau}^{\tau})}{\partial x_i} + \frac{2(v+1)x_i}{3(v-1)^2x_n} \frac{\partial f(x_{n-\tau}^{\tau})}{\partial x_n} + x_i \frac{\partial^2 f(x_{n-\tau}^{\tau})}{\partial x_i^2},$$

$$L(v,x) = \sum_{i=n-\tau}^{n} \left( f(x_{n-\tau}^{\tau}) \frac{3v-1}{2(v-1)} + x_i \left( -\frac{4}{v-1} \frac{\partial f(x_{n-\tau}^{\tau})}{\partial x_n} - \frac{4}{(v-1)^2x_n} f(x_{n-\tau}^{\tau}) \right) \right).$$
\[ R(u, x) = \left( \prod_{j=n-\tau}^{n} x_j^{-\frac{u+1}{2}} \right) \frac{(2\pi)^{-\frac{(u+1)(u+2)}{2}}}{2x_n^u}. \]

The proofs of Lemmas 4.1 - 4.3 are given in Appendices [B.4, B.5, B.7].

Using the upper bound of the covariance \( C(\hat{f}_{x_n}(x_{n-\tau}^n)) \) we can obtain the upper bound of the MISE and find the optimal bandwidth \( b \).

**Theorem 4.1.** In the conditions of Lemmas 4.1 - 4.3 the optimal bandwidth that provides minimum of the MISE is

\[ b^* = \left( \frac{\tau + 3}{2^n \pi^{n+1}} \int_{0}^{\infty} \left( \frac{f(x_{n-\tau}^n)}{3x_n^2} + \frac{1}{x_n} \sum_{i=n-\tau}^{n} x_i \frac{\partial^2 f(x_{n-\tau}^n)}{\partial x_i^2} \right) dx_{n-\tau} \right)^{1/\tau} \]

and the corresponding MISE has the order \( \text{MISE} \sim n^{-\frac{2}{\tau+1}} \).

The proof is given in Appendix [B.7].

### A Proofs

#### A.1 Proof of Lemma [3.1]

For \( x \geq 2b \) the mathematical expectation is

\[ E_X(\hat{f}_{x_n}(x_{n-\tau}^n)) = \int K_{\rho_1(x_n),b_{n-\tau}}(t_{n-\tau}) \ldots K_{\rho_1(x_n),b_n}(t_n) f(x_{n-\tau}^n) dt_{n-\tau} \ldots dt_n = E_\xi(f(\xi_{n-\tau}^n)), \]

where the r.v’s \( \xi_j \) are i.i.d and gamma distributed \( G(\rho_1(x_j), b_j) \) with mathematical expectation \( \mu_j = \rho_1(x_j)b_j = x_j \) and variance \( \sigma_j^2 = \rho_1(x_j)b_j^2 = x_jb_j \). Using the second order Taylor expansion

\[ f(\xi_{n-\tau}^n) = f(\mu_{n-\tau}^n) + \sum_{j=n-\tau}^{n} (\xi_j - \mu_j) \frac{\partial f}{\partial x_j} + \frac{1}{2} \sum_{j=n-\tau}^{n} (\xi_j - \mu_j)^2 \frac{\partial^2 f}{\partial x_j^2} + \sum_{j \neq l} (\xi_j - \mu_j)(\xi_l - \mu_l) \frac{\partial^2 f}{\partial x_j \partial x_l} + o(b) \]

\[ E_\xi(f(\xi_{n-\tau}^n)) = f(x_{n-\tau}^n) + \frac{1}{2} \sum_{j=n-\tau}^{n} x_j b_j \frac{\partial^2 f(x_{n-\tau}^n)}{\partial x_j^2} + o(b) \]

\[ \text{Bias}(\hat{f}_{x_n}(x_{n-\tau}^n)) = \frac{1}{2} \sum_{j=n-\tau}^{n} x_j b_j \frac{\partial^2 f(x_{n-\tau}^n)}{\partial x_j^2} + o(b). \]

If \( b_1 = b_2 = \ldots = b_n = b \) then

\[ \text{Bias}(\hat{f}_{x_n}(x_{n-\tau}^n)) = \frac{b}{2} \sum_{j=n-\tau}^{n} x_j \frac{\partial^2 f(x_{n-\tau}^n)}{\partial x_j^2} + o(b). \]
A.2 Proof of Lemma 3.2

The variance for \( x \geq 2b \)

\[
\text{var}(\hat{f}(x^n_{n-\tau})) = \frac{1}{n} \text{Var}(K_{\rho_1(x_{n-\tau}),b_{n-\tau}}(X_{n-\tau}) \cdots K_{\rho_1(x_n),b_n}(X_n))
\]

\[
= \frac{1}{n} \left( \text{E} \left( (K_{\rho_1(x_{n-\tau}),b_{n-\tau}}(X_{n-\tau}) \cdots K_{\rho_1(x_n),b_n}(X_n))^2 \right) - \text{E}^2 \left( K_{\rho_1(x_{n-\tau}),b_{n-\tau}}(X_{n-\tau}) \cdots K_{\rho_1(x_n),b_n}(X_n) \right) \right)
\]

We can write immediately that

\[
\frac{1}{n} \text{E}_n^2(\hat{f}(x^n_{n-\tau})) = \left( f(x^n_{n-\tau}) + \frac{1}{2} \sum_{j=n-\tau} x_j b_j \frac{\partial^2 f}{\partial x_j^2} + o(b) + o(b^2) \right)^2
\]

\[
= \frac{1}{n} \int \left( \prod_{j=n-\tau}^{2n} \frac{2x_j - 2}{t_j^{\gamma}} \text{exp} \left( \frac{-2t_j}{t_j^2} \right) \right) f(t^n_{n-\tau})dt_{n-\tau} \cdots dt_n = \frac{1}{n} \left( \prod_{j=n-\tau} B(x_j, b_j) \right) \text{E}_n(f(\eta^n_{n-\tau}))
\]

\[
\text{E}_n(\eta_{n-\tau}) = f(\mu^n_{n-\tau}) + \frac{1}{2} \sum_{j=n-\tau} b_j^2 \frac{\partial^2 f(\mu^n_{n-\tau})}{\partial x_j^2} + o(b^2)
\]

Therefore, Taylor expanding the arguments of the functions \( f(\mu^n_{n-\tau}) \)

\[
f(\mu^n_{n-\tau}) = f(x^n_{n-\tau}) - \sum_{j=n-\tau} b_j \frac{\partial f(x^n_{n-\tau})}{\partial x_j} + \sum_{j=n-\tau} b_j^2 \frac{\partial^2 f(x^n_{n-\tau})}{\partial x_j^2} + o(b^2)
\]

we deduce that

\[
\frac{1}{n} \left( \prod_{j=n-\tau} B(x_j, b_j) \right) \left( f(x^n_{n-\tau}) - \sum_{j=n-\tau} b_j \frac{\partial f(x^n_{n-\tau})}{\partial x_j} + \sum_{j=n-\tau} b_j^2 \frac{\partial^2 f(x^n_{n-\tau})}{\partial x_j^2} \right)
\]

\[
+ \frac{1}{2} \sum_{j=n-\tau} \left( x_j b_j - b_j^2 \right) \left( \frac{\partial^2 f(x^n_{n-\tau})}{\partial x_j^2} - \sum_{i=\tau} b_i \frac{\partial^3 f(x^n_{n-\tau})}{\partial x_j^2 \partial x_i} \right)
\]

Combining (16) and (18) we can write the variance

\[
\text{var}(\hat{f}(x^n_{n-\tau})) = \frac{1}{n} \left( \prod_{j=n-\tau} \frac{b_j^{-1/2} x_j^{-1/2}}{2\sqrt{\pi}} \right) \left( f(x^n_{n-\tau}) - \sum_{j=n-\tau} b_j \frac{\partial f(x^n_{n-\tau})}{\partial x_j} + \sum_{j=n-\tau} b_j^2 \frac{\partial^2 f(x^n_{n-\tau})}{\partial x_j^2} \right)
\]

\[
+ \frac{1}{2} \sum_{j=n-\tau} \left( x_j b_j - b_j^2 \right) \left( \frac{\partial^2 f(x^n_{n-\tau})}{\partial x_j^2} - \sum_{i=\tau} b_i \frac{\partial^3 f(x^n_{n-\tau})}{\partial x_j^2 \partial x_i} \right)
\]

\[
- \frac{1}{n} \left( f(x^n_{n-\tau}) + \frac{1}{2} \sum_{j=n-\tau} x_j b_j \frac{\partial^2 f}{\partial x_j^2} \right)^2 + o(b^2)
\]
If \( b_1 = b_2 = \ldots = b_n = b \) then
\[
\text{var}(\hat{f}(x_{n-\tau}^n)) = \frac{b^{\frac{r+1}{2}}}{n} \left( \prod_{j=n-\tau}^n \frac{x_j^{1/2}}{2\sqrt{\pi}} \right) \left( f(x_{n-\tau}^n) - \sum_{j=n-\tau}^n \frac{b}{2} \frac{\partial f(x_{n-\tau}^n)}{\partial x_j} + \sum_{j=n-\tau}^n \frac{b^2}{8} \frac{\partial^2 f(x_{n-\tau}^n)}{\partial x_j^2} \right)
+ \frac{1}{2} \sum_{j=n-\tau}^n \left( \frac{x_j b}{2} - \frac{b^2}{4} \right) \left( \frac{\partial^2 f(x_{n-\tau}^n)}{\partial x_j^2} - \sum_{i=n-\tau}^n \frac{b}{2} \frac{\partial^3 f(x_{n-\tau}^n)}{\partial x_j^2 \partial x_i} \right)
- \frac{1}{n} \left( f(x_{n-\tau}^n) + \frac{1}{2} \sum_{j=n-\tau}^n x_j b \frac{\partial^2 f(x_{n-\tau}^n)}{\partial x_j^2} \right)^2 + o(b^2)
\]

Hence it can be rewritten in the following notations
\[
v_1(x) = \sum_{j=n-\tau}^n \left( -\frac{1}{2} \frac{\partial f(x_{n-\tau}^n)}{\partial x_j} + \frac{x_j}{4} \frac{\partial^2 f(x_{n-\tau}^n)}{\partial x_j^2} \right),
\]
\[
v_2(x) = -\sum_{j,n-\tau}^n \frac{x_j \partial^3 f(x_{n-\tau}^n)}{8 \partial x_j^2 \partial x_i}
\]
\[
\text{var}(\hat{f}(x_{n-\tau}^n)) = \frac{b^{\frac{r+1}{2}}}{n} \left( \prod_{j=n-\tau}^n \frac{x_j^{1/2}}{2\sqrt{\pi}} \right) \left( f(x_{n-\tau}^n) + b v_1(x) + b^2 v_2(x) \right)
- \frac{1}{n} \left( f(x_{n-\tau}^n) + \frac{1}{2} \sum_{j=n-\tau}^n x_j b \frac{\partial^2 f(x_{n-\tau}^n)}{\partial x_j^2} \right)^2 + o(b^2)
\]

**B Proof of Lemma 3.3**

\[
\| \tilde{K}_1(X_{1t}, x_1, b) \|_q = \left( \int \left( \tilde{K}_1(y, x_1, b) \right)^q f(y) dy \right)^{1/q}
\]
\[
= \left( \int \left( \prod_{j=n-\tau}^n K_{\rho_1(x_j), b_j} (t_j) \right)^q f(t_{n-\tau}^n) dt_{n-\tau} \ldots dt_n \right)^{1/q}
\]
\[
= \left( E \left( f(\xi_{n}^n) \prod_{j=n-\tau}^n K_{\rho_1(x_j), b_j} (\xi_j) \right) \right)^{1/q},
\]
where the kernel \( \prod_{j=n}^{n} K_{\rho_1(x_j),b_j}(\xi_j) \) was used as the density function and \( \xi_j \) is a Gamma\((\rho_1(x_j), b_j)\) random variable. Expectation is \( \mu_j = x_j \) and variance is \( \text{var} = x_j b_j \).

\[
\begin{align*}
\mathbb{E}\left( \prod_{j=n}^{n} K_{\rho_1(x_j),b_j}(\xi_j) f(\xi_{n-r}) \right) &= \prod_{j=n}^{n} K_{\rho_1(x_j),b_j}(\mu_j) f(\mu_{n-r}) \\
&+ \sum_{i=n-r}^{n} \text{var} \left( \prod_{j=n}^{n} K_{\rho_1(x_j),b_j}(\xi_j) f(\xi_{n-r}) \right) \left| \xi = \mu \right| + o(b^2) \\
&= \prod_{j=n}^{n} K_{\rho_1(x_j),b_j}(x_j) \left( f(x_{n-r}) \left( \frac{3-q}{2} \right) + \sum_{i=n}^{n} \frac{q(q-1)b_i}{2x_i} f(x_{n-r}) \right) \\
&- (q-1)b_i \frac{\partial f(x_{n-r})}{\partial x_i} + \frac{b_i x_i \partial^2 f(x_{n-r})}{2 \partial x_i^2} + o(b^2)
\end{align*}
\]

If \( b_1 = b_2 = \ldots = b_n = b \) then

\[
\begin{align*}
&\left\| \tilde{K}_1 (X_{1l}, x_l, b_l) \right\|_q = b^{(r+1)\frac{1-q}{2q}} \left( \prod_{j=n}^{n} (2\pi x_j)^{\frac{1-q}{2q}} \right) \left( f(x_{n-r}) \left( \frac{3-q}{2} \right) + \sum_{i=n-r}^{n} \frac{q(q-1)b_i}{2x_i} f(x_{n-r}) \right) \\
&\quad - (q-1) \frac{\partial f(x_{n-r})}{\partial x_l} + \frac{x_l \partial^2 f(x_{n-r})}{2 \partial x_l^2} \left( \frac{1}{q} \right) + o(b^2)
\end{align*}
\]

\[
\left| \text{Cov} \left( \tilde{K}_1 (X_{1l}, x_l, b_l), \tilde{K}_1 (X_{1+k,l}, x_l, b_l) \right) \right| \\
\leq 2\pi \alpha(k)^{1/r} \left( b^{(r+1)\frac{1-q}{2q}} \left( \prod_{j=n}^{n} (2\pi x_j)^{\frac{1-q}{2q}} \right) \left( f(x_{n-r}) \left( \frac{3-q}{2} \right) + \sum_{i=n-r}^{n} \frac{q(q-1)b_i}{2x_i} f(x_{n-r}) \right) \\
&- (q-1) \frac{\partial f(x_{n-r})}{\partial x_l} + \frac{x_l \partial^2 f(x_{n-r})}{2 \partial x_l^2} \left( \frac{2/q}{q} \right) + o(b^2) \right)
\]

Taking \( p = q = 2 + \delta, r = \frac{2+\delta}{\delta} \) it can be deduced that the covariance is

\[
\begin{align*}
\frac{2}{n} \sum_{i=1}^{n} \left( 1 - \frac{k}{n} \right) \text{Cov} \left( \tilde{K}_1 (X_{1l}, x_l, b_l), \tilde{K}_1 (X_{1+k,l}, x_l, b_l) \right) \\
\leq \frac{2b^{(r+1)\frac{1-q}{2q}}}{n} (2\pi)^{1-(r+1)\frac{1-q}{2q}} \left( \prod_{j=n}^{n} x_j^{\frac{1-q}{2q}} \right) \left( f(x_{n-r}) \left( 1 - \frac{\delta}{2} \right) + b \sum_{i=n-r}^{n} \frac{(\delta + 1)(\delta + 2)}{2x_i} f(x_{n-r}) \right) \\
&- (\delta + 1) \frac{\partial f(x_{n-r})}{\partial x_l} + \frac{x_l \partial^2 f(x_{n-r})}{2 \partial x_l^2} \left( \frac{2/q}{q} \right) + o(b^2) \left( \sum_{k=1}^{n} \left( 1 - \frac{k}{n} \right) \alpha(k)^{\frac{1}{2q}} \right)
\end{align*}
\]

11
\[ S(\delta, x) = \sum_{i=n-\tau}^{n} \frac{(\delta + 1)(\delta + 2)}{2x_i} f(x_{n-\tau}^n) - (\delta + 1) \frac{\partial f(x_{n-\tau}^n)}{\partial x_i} + \frac{x_i \partial^2 f(x_{n-\tau}^n)}{2} \]

\[ D(\delta, x) = 2(2\pi)^{1-(\tau + 1)\frac{\tau + 1}{2}} \left( \prod_{j=n-\tau}^{n} x_j^{\frac{\tau + 1}{2}} \right) \]

\[ \frac{2}{n} \sum_{i=1}^{n} \left( 1 - \frac{i}{n} \right) C_{\text{ov}} \left( \widetilde{K}_1(X_{1:t}, x_i), \widetilde{K}_1(X_{1+1:t}, x_i) \right) \leq \frac{D(\delta, x)}{b^{-(\tau + 1)\frac{\tau + 1}{2}}} \left( f(x_{n-\tau}^n) \frac{1 - \delta}{2} + bS(\delta, x) + o(b^2) \right)^{2/(\tau + 2)} \int_{1}^{\infty} \alpha(\tau)^{\frac{\tau}{2}} d\tau \]

Let us denote \( \frac{\delta}{2 + \theta} = \upsilon, 0 < \upsilon < 1 \). Then, in this notations, we get

\[ |C(\hat{f}(x_{n-\tau}^n))| \leq \frac{D(\upsilon, x)}{n} b^{-(\tau + 1)\frac{\tau + 1}{2}} \left( bS(\upsilon, x) + f(x_{n-\tau}^n) \frac{3\upsilon - 1}{2(\upsilon - 1)} + o(b^2) \right)^{1-\upsilon} \int_{1}^{\infty} \alpha(\tau)^{\upsilon} d\tau \]

Hence

\[ |C(\hat{f}(x_{n-\tau}^n))| \sim \frac{1}{n} b^{-(\tau + 1)\frac{\tau + 1}{2}} \]

### B.1 Proof of Theorem 3.1

\[ \text{MSE}(\hat{f}(x_{n-\tau}^n)) = \text{Bias}(\hat{f}(x_{n-\tau}^n))^2 + \text{var}(\hat{f}(x_{n-\tau}^n)) + C(\hat{f}(x_{n-\tau}^n)) \]

\[ = \left( \frac{b}{2} \sum_{j=n-\tau}^{n} x_j \frac{\partial^2 f(x_{n-\tau}^n)}{\partial x_j^2} \right)^2 + \frac{1}{n} \left( \prod_{j=n-\tau}^{n} b^{\frac{(\tau + 1)}{2}} x_j^{\frac{1}{2}} \right) \left( f(x_{n-\tau}^n) + b\upsilon_1(x) + b^2\upsilon_2(x) \right) \]

\[ - \frac{1}{n} \left( f(x_{n-\tau}^n) + \frac{1}{2} \sum_{j=n-\tau}^{n} x_j b \frac{\partial^2 f}{\partial x_j^2} \right)^2 \]

\[ + \frac{D(\upsilon, x)}{n} b^{-(\tau + 1)\frac{\tau + 1}{2}} \left( bS(\upsilon, x) + f(x_{n-\tau}^n) \frac{3\upsilon - 1}{2(\upsilon - 1)} \right)^{1-\upsilon} \int_{1}^{\infty} \alpha(\tau)^{\upsilon} d\tau + o(b^2) \]

\[ \frac{\tau + 1}{2} \leq \frac{(\tau + 1)(\upsilon + 1)}{2} \]

\[ b_{\text{MSE}} = \left( \frac{(\tau + 1)(\upsilon + 1)(3\upsilon - 1)\upsilon}{n(2\upsilon - 1)^{1-\upsilon}} \int_{1}^{\infty} \alpha(\tau)^{\upsilon} d\tau \right) \left( \int_{n}^{n-\tau} x_j \frac{\partial f(x_{n-\tau}^n)}{\partial x_j^2} dx_{n-\tau} \ldots dx_n \right)^{\frac{(\tau + 1)(\upsilon + 1)(3\upsilon - 1)\upsilon}{2}} \]
B.2 Proof of Lemma 3.4

Let us partition the covariance into two sums

\[ |C(x_n^{n-\tau})| \leq \frac{2}{n} \sum_{\tau=2}^{c(n)} \text{Cov} \left( \tilde{K}_1 (X_{1l}, x_l, b), \tilde{K}_1 (X_{\tau l}, x_l, b) \right) \]

\[ + \frac{2}{n} \sum_{\tau=c(n)+1}^{\infty} \text{Cov} \left( \tilde{K}_1 (X_{1l}, x_l, b), \tilde{K}_1 (X_{\tau l}, x_l, b) \right) = I_1 + I_2 \]

Using Lemma 4.3, the second sum in the latter equation has the following view

\[ I_2 = \frac{2}{n} \sum_{\tau=c(n)+1}^{\infty} \text{Cov} \left( \tilde{K}_1 (X_{1l}, x_l, b), \tilde{K}_1 (X_{1+i l}, x_l, b) \right) \]

\[ \leq \frac{D(\delta, x)}{n} b^{-(\tau+1)\frac{4+1}{2+2}} \left( f(x_{n-\tau}) \frac{1-\delta}{2} + bS(\delta, x) + O \left( (\tau+1)b^2 \right) \right) \sum_{\tau=2}^{\infty} \frac{1}{\tau} \alpha(\tau)^{\frac{2}{2+2}} \]

where we used the notations

\[ S(\delta, x) = \sum_{i=\tau=n-\tau}^{n} \frac{(\delta + 1)(\delta + 2)}{2x_i} f(x_{n-\tau}) - (\delta + 1) \frac{\partial f(x_{n-\tau})}{\partial x_i} + \frac{x_i}{2} \frac{\partial^2 f(x_{n-\tau})}{\partial x_i^2} \]

\[ D(\delta, x) = 2(2\pi)^{1-(\tau+1)\frac{4+1}{2+2}} \left( \prod_{j=\tau=n-\tau}^{n} x_j^{-\frac{6+1}{2+2}} \right) \]

\[ I_2 \leq \frac{D(\kappa, x)}{c(n)n} b^{-(\tau+1)\frac{2\kappa+1}{2\kappa}} \left( f(x_{n-\tau}) \frac{6\kappa - 1}{2(2\kappa - 1)} + bS(\kappa, x) + O \left( (\tau+1)b^2 \right) \right)^{1-2\kappa} \sum_{\tau=2}^{\infty} \frac{1}{\tau} \alpha(\tau)^{2\kappa} \]

The first sum we do using Lemma 4.3.

\[ I_1 \leq \frac{2M}{n} \sum_{\tau=2}^{c(n)} \left[ \int_{0}^{\infty} \prod_{j=\tau=n-\tau}^{n} K_{\rho(x_j, b_j, b_j)} (u) \left| du \right| \right]^2 = \frac{2Mc(n)}{n} \]

We want the rate of convergence in \( b \) of \( I_1 \) would be more then \( I_2 \). Then

\[ c(n) \leq b^{-\frac{1}{2+2} \frac{4+1}{2+2}} \]

Let as choose for example \( c(n) = b^{-\frac{1+1}{8+1}, \kappa = 1/4} \) then

\[ I_1 \leq \frac{2M}{nb^{1+1}} \]

\[ I_2 \leq \frac{D(\kappa, x)}{nb^{1+1} \kappa} \left( f(x_{n-\tau}) \frac{6\kappa - 1}{2(2\kappa - 1)} + bS(\kappa, x) + O \left( (\tau+1)b^2 \right) \right)^{1-2\kappa} \sum_{\tau=2}^{\infty} \frac{1}{\tau} \alpha(\tau)^{2\kappa} \]
B.3 Proof of Theorem 3.2

Using Lemmas 3.1, 3.2 and 3.4 let us write MSE of the estimate \( \hat{f}(x^n_{n-r}) \) and find the optimal bandwidth \( \theta \) that minimizes MISE. Using Bias, variance and covariance from Lemmas ??, ?? and 3.4 into (??) we get

\[
MSE(\hat{f}(x^n_{n-r})) = \left( \frac{b}{2} \sum_{j=n-r}^{n} x_j \frac{\partial^2 f(x^n_{n-r})}{\partial x_j^2} \right)^2 + \frac{1}{n} \left( \frac{b}{\sqrt{2\pi}} \prod_{j=n-r}^{n} x_j^{-1/2} \right)^2 
\]

\[
\cdot \left( f(x^n_{n-r}) + b v_1(x) + b^2 v_2(x) \right) - \frac{1}{n} \left( f(x^n_{n-r}) + \frac{1}{2} \sum_{j=n-r}^{n} x_j \frac{\partial f(x^n_{n-r})}{\partial x_j^2} \right)^2 
\]

\[
+ \frac{2M}{nb^2} + \frac{D(\kappa, x)}{nb^2} \left( f(x^n_{n-r}) \frac{6\kappa - 1}{2(2\kappa - 1)} + bS(\kappa, x) + O \left( (\tau + 1)b^2 \right) \right)^{1-2\kappa} 
\]

\[
\cdot \sum_{\tau=2}^{\infty} \tau \alpha(\tau) 2^{2\kappa} + O \left( (\tau + 1)b^2 \right) 
\]

As the rate of convergence in \( b \) of variance is more then of covariance we can neglect it. Integrating MSE and minimizing it we find the optimal bandwidth.

\[
b = \left( \frac{\int_{0}^{\infty} \left( \prod_{j=n-r}^{n} x_j^{-1/2} \right)^2 f(x^n_{n-r}) dx^n_{n-r} }{\int_{0}^{\infty} \sum_{j=n-r}^{n} x_j \frac{\partial^2 f(x^n_{n-r})}{\partial x_j^2} dx^n_{n-r} } \right)^{\frac{2}{n + \tau}} 
\]

\[
\cdot (\tau + 1) \int_{0}^{\infty} \left( \prod_{j=n-r}^{n} x_j^{-1/2} \right)^2 f(x^n_{n-r}) dx^n_{n-r} 
\]

B.4 Proof of Lemma 4.1

For \( x \geq 2b \) the mathematical expectation is

\[
E[x(\hat{f}(x^n_{n-r}))] = \frac{1}{b^n} E[K_{\rho_1(x_{n-r}),b_{n-r}}(X_{n-r}) \ldots K_{\rho_1(x_n),b_n}(X_n) L_1(X_n, x_n, b_n)] \tag{12}
\]

\[
= \frac{1}{b^n} \int K_{\rho_1(x_{n-r}),b_{n-r}}(t_{n-r}) \ldots K_{\rho_1(x_n),b_n}(t_n) L_1(t_n, x_n, b_n) f(t^n_{n-r}) dt_{n-r} \ldots dt_n 
\]

\[
= \frac{1}{b^n} E_k(f(\xi^n_{n-r}) L_1(\xi_n, x_n, b_n)) = \frac{1}{b^n} E_k(f(\xi^n_{n-r}) \ln(\xi_n)) 
\]

\[
- \frac{1}{b^n} E_k(f(\xi^n_{n-r})(\ln b_n + \Psi(\rho_1(x_n)))) 
\]

where the r.v’s \( \xi \) are i.i.d and gamma distributed \( G(\rho_1(x_j), b_j) \) with mathematical expectation \( \mu_j = \rho_1(x_j)b_j = x_j \) and variance \( \sigma_j^2 = \rho_1(x_j)b_j^2 = x_jb_j \). Using the second order Taylor expansion

\[
f(\xi^n_{n-r}) = f(\mu^n_{n-r}) + \sum_{j=n-r}^{n} (\xi_j - \mu_j) \frac{\partial f}{\partial x_j} 
\]

\[
+ \frac{1}{2} \sum_{j=n-r}^{n} (\xi_j - \mu_j)^2 \frac{\partial^2 f}{\partial x_j^2} + \sum_{j \neq l} (\xi_j - \mu_j)(\xi_l - \mu_l) \frac{\partial^2 f}{\partial x_j \partial x_l} + o(b) 
\]
We can write that

\[ E_\xi(f(x^n_{n-\tau})) = f(x^n_{n-\tau}) + \frac{1}{2} \sum_{j=n-\tau}^n x_j b_j \frac{\partial^2 f}{\partial x_j^2} + o(b) \]

\[ f(\xi^n_{n-\tau}) \ln(\xi_n) = f(\mu^n_{n-\tau}) \ln(\mu_n) + \sum_{j=n-\tau}^n (\xi_j - \mu_j) \frac{\partial (f(\xi^n_{n-\tau}) \ln(\xi_n))}{\partial x_j} \bigg|_\mu + \frac{1}{2} \sum_{j=n-\tau}^n (\xi_j - \mu_j)^2 \frac{\partial^2 (f(\xi^n_{n-\tau}) \ln(\xi_n))}{\partial x_j^2} \bigg|_\mu + o(b) \]

\[ + \frac{1}{2} \sum_{j=n-\tau}^n (\xi_j - \mu_j)^2 \frac{\partial^2 (f(\xi^n_{n-\tau}) \ln(\xi_n))}{\partial x_j^2} \bigg|_\mu + \sum_{j \neq l} (\xi_j - \mu_j)(\xi_l - \mu_l) \frac{\partial^2 (f(\xi^n_{n-\tau}) \ln(\xi_n))}{\partial x_j \partial x_l} \bigg|_\mu + o(b) \]

\[ E_\xi(f(\xi^n_{n-\tau}) \ln(\xi_n)) = f(x^n_{n-\tau}) \ln(x_n) \]

\[ + \frac{1}{2} \sum_{j=n-\tau}^n x_j b_j \frac{\partial^2 f}{\partial x_j^2} \ln(x_n) + \frac{1}{2} x_n b_n \left( \frac{2}{x_n} \frac{\partial f}{\partial x_n} - \frac{f(x^n_{n-\tau})}{x_n} \right) + o(b) \]

Using the approximation of the Digamma function when \( \rho \to \infty \)

\[ \Psi(\rho) = \ln \rho - \frac{1}{2\rho} - \frac{1}{12\rho^2} + \frac{1}{120\rho^4} - \frac{1}{252\rho^6} + O \left( \frac{1}{\rho^8} \right), \]

we can write that

\[ \ln b_n + \Psi(\rho_1(x_n)) = \ln x_n - \frac{b_n}{2x_n} - \frac{b_n^2}{12x_n^2} + o(b_n^2). \]

Hence (12) is the following

\[ E_X(\hat{f}_{x_n}(x^n_{n-\tau})) = f(x^n_{n-\tau}) \frac{b_n}{12x_n^2} + \frac{\partial f(x^n_{n-\tau})}{\partial x_n} + \frac{1}{2} \sum_{j=n-\tau}^n x_j b_j \frac{\partial^2 f(x^n_{n-\tau})}{\partial x_j^2} \left( \frac{1}{2x_n} + \frac{b_n}{12x_n^2} \right) + O(b) \]

When \( \tau = 0 \) the multivariate mathematical expectation \( E(13) \) coincides with it’s univariate variant in [7].

The leading term of bias expansion may be written as

\[ Bias(\hat{f}_{x_n}(x^n_{n-\tau})) = f(x^n_{n-\tau}) \frac{b_n}{12x_n^2} + \frac{1}{2} \sum_{j=n-\tau}^n x_j b_j \frac{\partial^2 f(x^n_{n-\tau})}{\partial x_j^2} \left( \frac{b}{x_n} + \frac{b^2}{6x_n^2} \right) + o(b). \]

If \( b_1 = b_2 = \ldots = b_n = b \) then

\[ Bias(\hat{f}_{x_n}(x^n_{n-\tau})) = f(x^n_{n-\tau}) \frac{b}{12x_n^2} + \frac{1}{4} \sum_{j=n-\tau}^n x_j \frac{\partial^2 f(x^n_{n-\tau})}{\partial x_j^2} \left( \frac{b}{x_n} + \frac{b^2}{6x_n^2} \right) + o(b) \]

\[ = bB_1(x) + b^2 B_2(x) + o(b). \]

\[ B_1(x) = f(x^n_{n-\tau}) \frac{1}{12x_n^2} + \frac{1}{4x_n} \sum_{j=n-\tau}^n x_j \frac{\partial^2 f(x^n_{n-\tau})}{\partial x_j^2}, \]

\[ B_2(x) = \frac{1}{24x_n^2} \sum_{j=n-\tau}^n x_j \frac{\partial^2 f(x^n_{n-\tau})}{\partial x_j^2}. \]
B.5 Proof of Lemma 4.2

The variance for $x \geq 2b$.

\[
\text{var}(\hat{\rho}_n(x_n^n)) = \frac{1}{n} \text{Var}(\frac{1}{b_n} K_{\rho_1(x_n^n),b_n}(X_{n^-}) \cdots K_{\rho_1(x_n^n),b_n}(X_{n^-}) L_1(X_n,x_n,b_n))
\]

\[
= \frac{1}{nb_n^2} \left( \mathbb{E} \left( (K_{\rho_1(x_n^n),b_n}(X_{n^-}) \cdots K_{\rho_1(x_n^n),b_n}(X_{n^-}) L_1(X_n,x_n,b_n))^2 \right) - \mathbb{E}^2(K_{\rho_1(x_n^n),b_n}(X_{n^-}) \cdots K_{\rho_1(x_n^n),b_n}(X_{n^-}) L_1(X_n,x_n,b_n)) \right)
\]

\[
\text{where the r.v's } \eta_j \text{ are iid and gamma distributed } G(\frac{2x_j}{b_j} - 1, \frac{b_j}{2}) \text{ with mathematical expectation } \mu_j = x_j - \frac{b_j}{2} \text{ and variance } \sigma_j^2 = \frac{x_j b_j^2}{4}.
\]

Using the property of the gamma function $\Gamma^2\left(\frac{x}{2}\right) + 1 = \left(\frac{x}{2}\right)^2 \Gamma^2\left(\frac{x}{2}\right)$, we get

\[
\int_{t_{n-\tau}}^{t_n} \left( \prod_{j=n-\tau}^{n} B(x_j,b_j) \right) \mathbb{E}_\eta \left( L_1(\eta_j,x_n,b_n)^2 f(\eta_j^n) \right) = \frac{1}{nb_n^2} \left( \prod_{j=n-\tau}^{n} B(x_j,b_j) \right) \frac{1}{nb_n^2} \left( \prod_{j=n-\tau}^{n} B(x_j,b_j) \right) \mathbb{E}_\eta \left( f(\eta_j^n) \ln^2 \eta_j \right)
\]

\[
- 2 \left( \ln b_n + \Psi \left( \frac{x_n}{b_n} \right) \right) \mathbb{E}_\eta \left( f(\eta_j^n) \ln \eta_j \right) + \left( \ln b_n + \Psi \left( \frac{x_n}{b_n} \right) \right)^2 \mathbb{E}_\eta \left( f(\eta_j^n) \right)
\]

where the r.v's $\eta_j$ are iid and gamma distributed $G(\frac{2x_j}{b_j} - 1, \frac{b_j}{2})$ with mathematical expectation $\mu_j = x_j - \frac{b_j}{2}$ and variance $\sigma_j^2 = \frac{x_j b_j^2}{4}$.
Let \( R(z) = \sqrt{2\pi} \exp(-z)z^{z+1/2}/\Gamma(z+1) \) for \( z \geq 0 \). So we can express gamma function as
\[
\Gamma^2 \left( \frac{x}{b} + 1 \right) = \left( \frac{\sqrt{2\pi} \exp\left( -\frac{x}{b} \right) \left( \frac{x}{b} \right)^{\frac{x}{b} + 1/2}}{R\left( \frac{x}{b} \right)} \right)^2.
\]
Using the properties of the gamma function
\[
\Gamma \left( \frac{2x}{b} - 1 \right) = \frac{\Gamma \left( \frac{2x}{b} + 1 \right)}{\Gamma \left( \frac{2x}{b} - 1 \right)} = \frac{\sqrt{2\pi} \exp\left( -\frac{2x}{b} \right) \left( \frac{2x}{b} \right)^{\frac{2x}{b} + 1/2}}{\frac{2x}{b} - 1} R\left( \frac{2x}{b} \right),
\]
we obtain
\[
B(x, b_j) = \frac{\left( \frac{b_j}{2} \right)^{-\frac{1}{2}x_j^2} R^2 \left( \frac{x_j}{b_j} \right)}{2\sqrt{\pi} R\left( \frac{2x_j}{b_j} \right)(1 - \frac{b_j}{2x_j})}.
\]
According to Lemma 3 of [?], \( R(z) \) is increasing function which converges to 1 as \( z \to \infty \) and \( R(z) < 1 \) for any \( z > 0 \). Then
\[
f(\eta_{n-\tau}) \ln(\eta_n) = f(\mu_{n-\tau}) \ln(\mu_n) + \sum_{j=n-\tau}^{n} (\eta_j - \mu_j) \frac{\partial f(\eta_{n-\tau}) \ln(\eta_n)}{\partial x_j}|_{\eta=\mu}
\]
\[
+ \frac{1}{2} \sum_{j=n-\tau}^{n-1} (\eta_j - \mu_j)^2 \frac{\partial^2 f(\eta_{n-\tau}) \ln(\eta_n)}{\partial x_j^2}|_{\eta=\mu} + \sum_{j \neq l} (\eta_j - \mu_j)(\eta_l - \mu_l) \frac{\partial^2 f(\eta_{n-\tau}) \ln(\eta_n)}{\partial x_j \partial x_l}|_{\eta=\mu} + o(b^2)
\]
\[
E_\eta(f(\eta_{n-\tau}) \ln(\eta_n)) = f(\mu_{n-\tau}) \ln(\mu_n)
\]
\[
+ \frac{1}{2} \sum_{j=n-\tau}^{n-1} \left( \frac{x_j b_n}{2} - \frac{b_n^2}{4} \right) \frac{\partial^2 f(\mu_{n-\tau})}{\partial x_j^2}|_{\eta=\mu} \ln(\mu_n) + \frac{1}{2} \left( \frac{x_n b_n}{2} - \frac{b_n^2}{4} \right) \frac{\partial^2 f}{\partial x_n^2}|_{\eta=\mu} \ln(\mu_n)
\]
\[
- \frac{1}{2} \left( \frac{x_n b_n}{2} - \frac{b_n^2}{4} \right) f(\mu_{n-\tau}) \mu_n^2 + \frac{1}{2} \left( \frac{x_n b_n}{2} - \frac{b_n^2}{4} \right) \frac{\partial f}{\partial x_n}|_{\eta=\mu} \frac{2}{\mu_n} + o(b^2)
\]
\[
= f(\mu_{n-\tau}) \left( \ln(\mu_n) - \frac{1}{2} \left( \frac{x_n b_n}{2} - \frac{b_n^2}{4} \right) \frac{1}{\mu_n^2} \right)
\]
\[
+ \frac{1}{2} \sum_{j=n-\tau}^{n} \left( \frac{x_j b_n}{2} - \frac{b_n^2}{4} \right) \frac{\partial^2 f(\mu_{n-\tau})}{\partial x_j^2}|_{\eta=\mu} \ln(\mu_n) + \frac{1}{2} \left( \frac{x_n b_n}{2} - \frac{b_n^2}{4} \right) \frac{\partial f}{\partial x_n}|_{\eta=\mu} \frac{2}{\mu_n} + o(b^2)
\]
\[
= f(\mu_{n-\tau}) \left( \ln \mu_n - \frac{b_n}{4\mu_n} \right) + \sum_{j=n-\tau}^{n} \left( \frac{x_j b_j}{2} - \frac{b_j^2}{4} \right) \frac{\partial^2 f(\mu_{n-\tau})}{\partial x_j^2} \ln(\mu_n) + \frac{b_n}{2} \frac{\partial f(\mu_{n-\tau})}{\partial x_n} + o(b^2).
\]
\[
E_n(f(\eta_{n-\tau}^n \ln^2(\eta_n))) = f(\mu_{n-\tau}^n) \ln^2(\mu_n)
\]
\[+ \frac{1}{2} \sum_{j=n-\tau}^{n-1} \left( \frac{x_j b_j}{2} - \frac{b_j^2}{4} \right) \frac{\partial^2 f}{\partial x_j^2} \bigg|_{\eta=\mu} \ln^2(\mu_n) + \frac{1}{2} \left( \frac{x_n b_n}{2} - \frac{b_n^2}{4} \right) \left( \frac{\partial^2 f}{\partial x_n^2} \bigg|_{\eta=\mu} \ln^2(\mu_n) \right)
\]
\[+ \frac{1}{2} \left( \frac{x_n b_n}{2} - \frac{b_n^2}{4} \right) 2f(\mu_{n-\tau}^n) \left( 1 - \ln \mu_n \right) + \frac{1}{2} \left( \frac{x_n b_n}{2} - \frac{b_n^2}{4} \right) \frac{\partial f}{\partial x_n} \bigg|_{\eta=\mu} 4 \ln \mu_n \frac{\mu_n}{\mu_n} + o(b^2)
\]
\[= f(\mu_{n-\tau}^n) \left( \ln^2(\mu_n) + \frac{b_n^2}{2 \mu_n} (1 - \ln \mu_n) \right) + \ln \mu_n \frac{\partial f(\mu_{n-\tau}^n)}{\partial x_n} + o(b^2)
\]
\[+ \frac{\ln^2 \mu_n}{2} \sum_{j=n-\tau}^{n} \left( \frac{x_j b_j}{2} - \frac{b_j^2}{4} \right) \frac{\partial^2 f(\mu_{n-\tau}^n)}{\partial x_j^2} + b_n \ln \mu_n \frac{\partial f(\mu_{n-\tau}^n)}{\partial x_n} + o(b^2).
\]

Using Taylor expansion
\[\ln \mu_n = \ln x_n - \frac{b_n}{2x_n} - \frac{b_n^2}{8x_n^2} + o(b^2),\]
\[\frac{1}{\mu_n} = \frac{1}{x_n} + \frac{b_n}{2x_n^2} + \frac{b_n^2}{4x_n^3} + o(b^2),\]
we can rewrite (17) as following
\[
\frac{1}{n} \left( \prod_{j=n-\tau}^{\tau} B(x_j, b_j) \right) \left( f(\mu_{n-\tau}^n) \left( \frac{1}{4x_n^2} + \frac{1}{2x_n b_n} + \frac{7b_n}{48x_n^3} + \frac{7b_n^2}{576x_n^4} + \frac{b_n^3}{192x_n^5} \right) \right)
\]
\[+ \sum_{j=n-\tau}^{n} \left( \frac{x_j b_j}{2} - \frac{b_j^2}{4} \right) \frac{\partial^2 f(\mu_{n-\tau}^n)}{\partial x_j^2} \left( \frac{b_n^2}{1152x_n^4} - \frac{b_n}{24x_n^2} \frac{\partial f(\mu_{n-\tau}^n)}{\partial x_n} + o(b^2) \right).
\]
Therefore, Taylor expanding the arguments of the functions \(f(\mu_{n-\tau}^n)\)
\[f(\mu_{n-\tau}^n) = f(x_n^\tau) - \sum_{j=n-\tau}^{n} \frac{b_j}{2} \frac{\partial f(x_n^\tau)}{\partial x_j} + \sum_{j=n-\tau}^{n} \frac{b_j^2}{4} \frac{\partial^2 f(x_n^\tau)}{\partial x_j^2} + o(b^2)
\]
we deduce that (17) is

\[
\frac{1}{n} \left( \prod_{j=n-\tau}^n B(x_j, b_j) \right) \left( \frac{b_n}{24x_n^2} \frac{\partial f(x_{n-\tau})}{\partial x_n} - \sum_{j=n-\tau}^n \frac{b_j \partial^2 f(x_{n-\tau})}{4 \partial x_n \partial x_j} + \sum_{j=n-\tau}^n \frac{b_j^2 \partial^3 f(x_{n-\tau})}{4 \partial x_n \partial x_j^2} \right) (18)
\]

\[+ \frac{b_n^2}{1152x_n^4} \sum_{j=n-\tau}^n \left( \frac{x_jb_j}{2} - \frac{b_j^2}{4} \right) \left( \frac{b_j^2}{1152x_n^4} + \frac{b_j^2}{4} \left( \frac{1}{4x_n^2} + \frac{1}{2x_n b_n} + \frac{7}{48x_n^6} + \frac{7}{576x_n^4} + \frac{b_n^2}{192x_n^6} \right) \right) \]

\[+ \left( f(x_{n-\tau}) - \sum_{j=n-\tau}^n b_j \frac{\partial f(x_{n-\tau})}{\partial x_j} \right) \left( \frac{1}{4x_n^2} + \frac{1}{2x_n b_n} + \frac{7b_n}{48x_n^4} + \frac{7b_n^2}{576x_n^4} + \frac{b_n^3}{192x_n^6} \right) + o(b^2) \]

Combining (16) and (18) we can write the variance

\[
\text{var}(\hat{f}_{x_n}(x_{n-\tau})) = \frac{1}{n} \left( \prod_{j=n-\tau}^n \frac{b_j^{-1/2} x_j^{-1/2}}{2\sqrt{\pi}} \right) \left( \frac{b_n}{24x_n^2} \frac{\partial f(x_{n-\tau})}{\partial x_n} - \sum_{j=n-\tau}^n \frac{b_j \partial^2 f(x_{n-\tau})}{4 \partial x_n \partial x_j} + \sum_{j=n-\tau}^n \frac{b_j^2 \partial^3 f(x_{n-\tau})}{4 \partial x_n \partial x_j^2} \right)
\]

\[+ \frac{b_n^2}{1152x_n^4} \sum_{j=n-\tau}^n \left( \frac{x_jb_j}{2} - \frac{b_j^2}{4} \right) \left( \frac{b_j^2}{1152x_n^4} + \frac{b_j^2}{4} \left( \frac{1}{4x_n^2} + \frac{1}{2x_n b_n} + \frac{7}{48x_n^4} + \frac{7}{576x_n^4} + \frac{b_n^2}{192x_n^6} \right) \right) \]

\[+ \left( f(x_{n-\tau}) - \sum_{j=n-\tau}^n b_j \frac{\partial f(x_{n-\tau})}{\partial x_j} \right) \left( \frac{1}{4x_n^2} + \frac{1}{2x_n b_n} + \frac{7b_n}{48x_n^4} + \frac{7b_n^2}{576x_n^4} + \frac{b_n^3}{192x_n^6} \right) \]

\[= \frac{1}{n} \left( \prod_{j=n-\tau}^n \frac{x_j^{-1/2}}{2\sqrt{\pi}} \right) \left( \frac{b_n}{24x_n^2} \frac{\partial f(x_{n-\tau})}{\partial x_n} - \sum_{j=n-\tau}^n \frac{b_j \partial^2 f(x_{n-\tau})}{4 \partial x_n \partial x_j} \right)
\]

\[+ \sum_{j=n-\tau}^n \frac{\partial^2 f(x_{n-\tau})}{\partial x_j} \left( \frac{b_n}{16x_n^2} + \frac{b_n^2}{8x_n^2} \right) + f(x_{n-\tau}) \left( \frac{1}{4x_n^2} + \frac{1}{2x_n b_n} + \frac{7b_n}{48x_n^4} + \frac{7b_n^2}{576x_n^4} \right) \]

\[+ \frac{1}{n} \left( \text{Bias}^2(\hat{f}_{x_n}(x_{n-\tau})) + \left( \frac{\partial f(x_{n-\tau})}{\partial x_n} \right)^2 + 2 \frac{\partial f(x_{n-\tau})}{\partial x_n} \text{Bias}(\hat{f}_{x_n}(x_{n-\tau})) + o(b^2) \right) \]

If \( b_1 = b_2 = \ldots = b_n = b \) then

\[
\text{var}(\hat{f}_{x_n}(x_{n-\tau})) = \frac{b^{(r+1)}}{n} \left( \prod_{j=n-\tau}^n \frac{x_j^{-1/2}}{2\sqrt{\pi}} \right) \left( \frac{b_n}{24x_n^2} \frac{\partial f(x_{n-\tau})}{\partial x_n} \right)
\]

\[+ \sum_{j=n-\tau}^n \frac{\partial^2 f(x_{n-\tau})}{\partial x_j} \left( \frac{b_n}{16x_n^2} + \frac{b_n^2}{8x_n^2} \right) + f(x_{n-\tau}) \left( \frac{1}{4x_n^2} + \frac{1}{2x_n b_n} + \frac{7b_n}{48x_n^4} + \frac{7b_n^2}{576x_n^4} \right) \]

\[+ \frac{1}{n} \left( \text{Bias}^2(\hat{f}_{x_n}(x_{n-\tau})) + \left( \frac{\partial f(x_{n-\tau})}{\partial x_n} \right)^2 + 2 \frac{\partial f(x_{n-\tau})}{\partial x_n} \text{Bias}(\hat{f}_{x_n}(x_{n-\tau})) + o(b^2) \right) \]

Hence it can be rewritten in the following notations

\[
V_1(x) = -\frac{1}{24x_n^2} \frac{\partial f(x_{n-\tau})}{\partial x_n} + \sum_{j=n-\tau}^n \left( \frac{1}{8x_n} \frac{\partial^2 f(x_{n-\tau})}{\partial x_j^2} - \frac{1}{8x_n^2} \frac{\partial f(x_{n-\tau})}{\partial x_j} \right) + \frac{7}{48x_n^6} f(x_{n-\tau}),
\]
\[ V_2(x) = \frac{7}{576x_n^4} f(x_n^{n-t}) + \sum_{j=n-t}^n \left( \frac{1}{16x_n^2} \frac{\partial^2 f(x_n^{n-t})}{\partial x_j^2} - \frac{7}{96x_n^3} \frac{\partial f(x_n^{n-t})}{\partial x_j} + \frac{1}{48x_n^2} \frac{\partial^2 f(x_n^{n-t})}{\partial x_n \partial x_j} \right), \]

\[ V_3(x) = \frac{f(x_n^{n-t})}{2x_n}, \quad V_4(x) = \frac{f(x_n^{n-t})}{4x_n^2} - \frac{1}{4x_n} \frac{\partial f(x_n^{n-t})}{\partial x_j}, \]

\[ \text{var}(\tilde{p}_{x_n}(x_n^{n-t})) = \]

\[ = \frac{b^{-(\tau+1)/2}}{n} \left( \prod_{j=n-\tau}^n \frac{x_j^{-1/2}}{2\sqrt{\pi}} \right) \left( bV_1(x) + b^2V_2(x) + \frac{1}{b} V_3(x) + V_4(x) + o(b^2) \right) \]

\[ - \frac{1}{n} \left( (bB_1(x) + b^2B_2(x))^2 + \left( \frac{\partial f(x_n^{n-t})}{\partial x_n} \right)^2 + 2 \frac{\partial f(x_n^{n-t})}{\partial x_n} (bB_1(x) + b^2B_2(x)) + o(b^2) \right) \]

B.6 Proof of Lemma 4.3

Furthermore we apply Davydov’s inequality

\[ |\text{Cov} \left( \tilde{K}(X_1, x, b), \tilde{K}(X_{1+i}, x, b) \right) | \leq 2\pi \alpha(i)^{1/\tau} \| \tilde{K}(X_1, x, b) \|_q \| \tilde{K}(X_{1+i}, x, b) \|_p, \]

where \( p^{-1} + q^{-1} + r^{-1} = 1 \), too, \([?]\).
Using Tailor series enpension we can write

\[
E \left( \sum_{i=n-\tau}^{n} \frac{(\xi_i - \mu_i)^2}{2} \frac{\partial^2}{\partial \xi_i^2} \left( L(\xi_n, x_n, b_n)^q f(\xi_n^{n-\tau}) \prod_{j=n-\tau}^{n} K_{\rho(x_j, b_j), b_j}^{q-1}(\xi_j) \right) \right)_{\xi = \mu = x}
\]

\[
= \left( \prod_{j=n-\tau}^{n} K_{\rho_j(x_j, b_j), b_j}^{q-1}(x_j) \right) \sum_{i=n-\tau}^{n} \frac{x_i b_i}{2} \left( \frac{\partial^2 f(x^{n-\tau}_i)}{\partial x_i^2} - f(x^{n-\tau}_i) \frac{q(1 - q)}{x_i} + \frac{\partial \left( \frac{\partial f(x^{n-\tau}_i)}{\partial x_i} \right) f(x^{n-\tau}_i) - \frac{\partial f(x^{n-\tau}_i)}{\partial x_i} \frac{2(1 - q)}{x_i} \right) + \frac{\partial f(x^{n-\tau}_i)}{\partial x_i} \cdot \frac{q f(x^{n-\tau}_i)}{x_i} \right)_{\xi = \mu = x}
\]

Using Tailor series enpension we can write

\[
L_1(x_n, x_n, b_n) = \ln x_n - \ln b_n - \Psi \left( \frac{x_n}{b_n} \right) = \frac{b_n}{2x_n} + \frac{b_n^2}{12x_n^2} + o(b^2), \quad (19)
\]

\[
L_1^q(x_n, x_n, b_n) = \left( \frac{b_n}{2x_n} \right)^q \left( 1 + \frac{b_n}{6x_n} + \frac{q(q - 1)}{2} \left( \frac{b_n}{6x_n} \right)^2 \right)
\]

\[
\frac{\partial}{\partial \xi_i} L_1(\xi_n, x_n, b_n) = \frac{1}{\xi_n} \bigg|_{\xi_n = x_n} = \frac{1}{x_n}, \quad \frac{\partial^2}{\partial \xi_i^2} L_1(\xi_n, x_n, b_n) = -\frac{1}{x_n^2}
\]

Using Stirling’s formula

\[
\Gamma(z) = \sqrt{2\pi} \left( \frac{z}{e} \right)^z \left( 1 + O \left( \frac{1}{z} \right) \right)
\]

and as \( \mu_j = x_j \) we can rewrite the kernel function as

\[
\prod_{j=n-\tau}^{n} K_{\rho_j(x_j, b_j), b_j}(x_j) = \prod_{j=n-\tau}^{n} \frac{x_j^{x_j^{-1}} \exp \left( \frac{-x_j}{b_j} \right) \Gamma \left( \frac{x_j}{b_j} \right)}{\sqrt{2\pi}(1 + O(b_j/x_j))}
\]

21
Hence, its upper bound is

$$\prod_{j=n^{-\tau}}^{n} K_{b_{i}(x),b_{j}}^{2^{-1}}(x) \leq \prod_{j=n^{-\tau}}^{n} \frac{1}{(2\pi x_{j} b_{j})^{2^{-\tau}}}.$$  \hspace{1cm} (20)

If \(b_{1} = b_{2} = \ldots = b_{n} = b\) then

$$\| \widetilde{K}_{1}(X_{1},x_{1},b_{i}) \|_{q} \leq \left( \prod_{j=n^{-\tau}}^{n}(2\pi x_{j})^{\frac{1-q}{2}} \right) \frac{b^{(r+1)\frac{1-q}{2}}} {2x_{n}} \left( \sum_{i=n^{-\tau}}^{n} \left( 1 + \frac{qb}{6x_{n}} + \frac{q(q-1)b^{2}}{72x_{n}^{2}} \right) \right) \cdot \left( f(x_{n,t}) \frac{3-q}{2} + b(1-q) \frac{\partial f(x_{n,t})}{\partial x_{i}} + \frac{x_{i}b}{2} \frac{\partial^{2} f(x_{n,t})}{\partial x_{i}^{2}} - \frac{q(1-q)b}{2x_{n}} f(x_{n,t}) \right) + x_{i} \left( 1 + \frac{(q-1)b}{6x_{n}} + \frac{(q-1)(q-2)b^{2}}{72x_{n}^{2}} \right) \left( 2q \frac{\partial f(x_{n,t})}{\partial x_{n}} - \frac{q^{2}}{x_{n}} f(x_{n,t}) \right)^{1/q} + o(b^{2}).$$

The same for \(\| \widetilde{K}_{1}(X_{1},x_{1},b_{i}) \|_{p}\). Let \(p = q\) then

$$|\text{Cov}(\widetilde{K}_{1}(X_{1},x_{1},b_{i}), \widetilde{K}_{1}(X_{1+k},x_{1},b_{i}))| \leq 2\pi \alpha(k)^{1/r} \left( \prod_{j=n^{-\tau}}^{n}(2\pi x_{j})^{\frac{1-q}{2}} \right) \frac{b^{(r+1)\frac{1-q}{2}}}{4x_{n}^{2}} \left( \sum_{i=n^{-\tau}}^{n} \left( 1 + \frac{qb}{6x_{n}} + \frac{q(q-1)b^{2}}{72x_{n}^{2}} \right) \right) \cdot \left( f(x_{n,t}) \frac{3-q}{2} + b(1-q) \frac{\partial f(x_{n,t})}{\partial x_{i}} + \frac{x_{i}b}{2} \frac{\partial^{2} f(x_{n,t})}{\partial x_{i}^{2}} - \frac{q(1-q)b}{2x_{n}} f(x_{n,t}) \right) + x_{i} \left( 1 + \frac{(q-1)b}{6x_{n}} + \frac{(q-1)(q-2)b^{2}}{72x_{n}^{2}} \right) \left( 2q \frac{\partial f(x_{n,t})}{\partial x_{n}} - \frac{q^{2}}{x_{n}} f(x_{n,t}) \right)^{2/q} + o(b^{2}).$$

Taking \(p = q = 2 + \delta, r = \frac{2+\delta}{\delta}\) it can be deduced that the covariance is

$$\frac{2}{nb^{2}} \sum_{k=1}^{n} \left( 1 - \frac{k}{n} \right) \text{Cov}(\widetilde{K}_{1}(X_{1},x_{1},b), \widetilde{K}_{1}(X_{1+k},x_{1},b)) \leq \left( \prod_{j=n^{-\tau}}^{n} x_{j}^{\frac{1-q}{2}} \right) \frac{(2\pi)^{\frac{1-q}{2}}}{2x_{n}^{2}} b^{-(r+1)\frac{1-q}{2}n^{-\tau}} \left( \sum_{i=n^{-\tau}}^{n} \left( f(x_{n,t}) \frac{1-\delta}{2} + x_{i} \left( 2(2 + \delta) \frac{\partial f(x_{n,t})}{\partial x_{n}} \right) \right) \right) - \frac{(2 + \delta)^{2}}{x_{n}} f(x_{n,t}) + b \left( \frac{1-\delta}{4x_{n}} + \frac{(\delta+1)(\delta+2)}{2x_{n}} f(x_{n,t}) \right) - \frac{(\delta+1)\frac{\partial f(x_{n,t})}{\partial x_{i}}}{3x_{n}} + \frac{(\delta+1)(\delta+2)x_{i} \frac{\partial f(x_{n,t})}{\partial x_{n}}}{144x_{n}^{3}} f(x_{n,t}) - \frac{\partial^{2} f(x_{n,t})}{2x_{n}^{2}} \left( \frac{1}{144x_{n}^{3}} \right) f(x_{n,t}) - \frac{\delta(\delta+1)(\delta+2)}{72x_{n}^{3}} f(x_{n,t}) \right.$$

$$+ \frac{\delta(\delta+1)(\delta+2)}{36x_{n}^{2}} \frac{x_{i} \frac{\partial f(x_{n,t})}{\partial x_{n}}}{\partial x_{i}} \right)^{2/(\delta+2)} \left( \sum_{k=1}^{n} \left( 1 - \frac{k}{n} \right) \frac{\partial f(x_{n,t})}{\partial x_{i}} \right)^{2/(\delta+2)} + o(b^{2}).$$
Let us introduce the notations

\[ V(\delta, x) = \sum_{i=n-\tau}^n \left( \left( -\frac{(\delta - 1)(\delta + 1)(\delta + 2)}{144x_i^2} + \frac{(\delta + 1)(\delta + 2)}{2x_i} - \frac{\delta(\delta + 1)(\delta + 2)^2}{72x_i^3} \right) f(x_i^n) \right) \]

\[ - (\delta + 1) \frac{\partial f(x_{n-\tau})}{\partial x_i} + \frac{\delta(\delta + 1)(\delta + 2)}{36x_i^2} \frac{\partial f(x_{n-\tau})}{\partial x_n} + \frac{x_i \partial^2 f(x_{n-\tau})}{2} \frac{\partial x_i^2}{\partial x_i^2} , \]

\[ W(\delta, x) = \sum_{i=n-\tau}^n \left( \left( -\frac{1 - \delta}{4x_i} + \frac{(\delta + 1)(\delta + 2)}{2x_i} - \frac{x_i(\delta + 1)(\delta + 2)^2}{6x_i^2} \right) f(x_i^n) \right) \]

\[ - (\delta + 1) \frac{\partial f(x_{n-\tau})}{\partial x_i} + \frac{\delta(\delta + 1)(\delta + 2)x_i \partial f(x_{n-\tau})}{3x_n} + \frac{x_i \partial^2 f(x_{n-\tau})}{2} \frac{\partial x_i^2}{\partial x_i^2} , \]

\[ L(\delta, x) = \sum_{i=n-\tau}^n \left( f(x_i^n) \frac{1 - \delta}{2} + x_i \left( 2 + \delta \right) \frac{\partial f(x_{n-\tau})}{\partial x_n} - \frac{(2 + \delta)^2}{x_n} f(x_{n-\tau}) \right) , \]

\[ R(\delta, x) = \left( \prod_{j=n-\tau}^n \frac{x_j^{\frac{2}{n} + 1}}{2x_j^n} \right) \frac{2}{nb^2} \sum_{k=1}^n \left( 1 - \frac{k}{n} \right) Cov \left( \tilde{K}_1(X_{1l}, x_l, b), \tilde{K}_1(X_{1+k,l}, x_l, b) \right) \]

\[ \leq \frac{R(\delta, x)}{b^{(\tau + 1)\frac{\tau + 2}{2}}} \left( \int_1^{\infty} (\tau)^{\frac{\tau + 1}{2}} d\tau + (2 + \delta)^2 \int_1^{\infty} (\tau)^{\frac{\tau + 2}{2}} d\tau \right) \]

Let us denote \( \frac{\delta}{x_{n-\tau}} = \nu, 0 < \nu < 1 \). Then, in this notations, we get

\[ |C(\tilde{f}_{x_n}(x_{n-\tau}^n))| \leq \frac{R(\nu, x)}{nb^{(\tau + 1)\frac{\tau + 2}{2}}} \left( \int_1^{\infty} (\tau)^{\frac{\tau + 2}{2}} d\tau \right) \]

Hence

\[ |C(\tilde{f}_{x_n}(x_{n-\tau}^n))| \sim \frac{1}{nb^{(\tau + 1)\frac{\tau + 2}{2}}} . \]

**B.7 Proof of Theorem 4.1**

\[ MSE(\tilde{f}_{x_n}(x_{n-\tau}^n)) = (Bias(\tilde{f}_{x_n}(x_{n-\tau}^n))^2 + var(\tilde{f}_{x_n}(x_{n-\tau}^n))^2 + C(\tilde{f}_{x_n}(x_{n-\tau}^n)) \]
\[
MSE(\hat{f}_{x_n}(x_{n-\tau})) = \left( bB_1(x) + b^2B_2(x) \right)^2 \left( 1 - \frac{1}{n} \right) \\
+ \frac{b^{-\tau+1}}{n} \left( \prod_{j=n-\tau}^{n} \frac{x_j^{-1/2}}{2\sqrt{\pi}} \right) \left( bV_1(x) + b^2V_2(x) + \frac{1}{b}V_3(x) + V_4(x) + o(b^2) \right) \\
- \frac{1}{n} \left( \frac{\partial f(x_{n-\tau})}{\partial x_n} \right)^2 + 2 \frac{\partial f(x_{n-\tau})}{\partial x_n} \left( bB_1(x) + b^2B_2(x) \right) \\
+ \frac{R(v,x)}{n} b^{-(\tau+1)\frac{\tau+1}{2}} \left( b^2V(v,x) + bW(v,x) + L(v,x) + o(b^2) \right) \int_{1}^{\infty} \alpha(\tau)^v \, d\tau,
\]

As \(0 < \nu < 1\) then

\[
\frac{\tau + 1}{2} + 1 > \frac{(\tau + 1)(\nu + 1)}{2},
\]

\(0 < \nu < \frac{2}{\tau + 1}\).

We can neglect some terms and find optimal bandwidth

\[
b = \left( \frac{(\tau + 3)V_3(x)}{4B_1^2(x)} \left( \prod_{j=n-\tau}^{n} \frac{x_j^{-1/2}}{2\sqrt{\pi}} \right) \right)^{-\frac{2}{\tau + 7}} n^{-\frac{2}{\tau + 7}}
\]

The optimal bandwidth for MISE is

\[
b_{MISE} = \left( \frac{\tau + 3}{2\pi} \frac{\tau + 7}{2} \int_{0}^{\infty} \frac{f(x_{n-\tau})}{3x_n} \prod_{j=n-\tau}^{n} x_j^{-1/2} \, dx_{n-\tau} \ldots dx_n \right)^{-\frac{2}{\tau + 7}} n^{-\frac{2}{\tau + 7}}
\]