APPLICATION OF MAZUR–ORLICZ’S THEOREM IN AMISE CALCULATION

Abstract. An approximation error and an asymptotic formula are given for shift invariant operators of polynomial order $q$. Density estimators based on shift invariant operators are introduced and AMISE is calculated.

1. Asymptotic formulas. We assume that $F, G : \mathbb{R}^d \to \mathbb{R}$ are functions such that there are constants $C > 0$ and $0 < q < 1$ such that for all $x \in \mathbb{R}^d$,

$$|F(x)| < C q^{|x|} \quad \& \quad |G(x)| < C q^{|x|},$$

where $|x|^2 = x \cdot x$ and $x \cdot x$ is the scalar product in $\mathbb{R}^d$. Consider the operator given by

$$Qf(x) = \int_{\mathbb{R}^d} K(x,y)f(y)dy,$$

where

$$K(x,y) = \sum_{\alpha \in \mathbb{Z}^d} F(y - \alpha)G(x - \alpha).$$

For $h > 0$, define

$$Q_h = \sigma_h \circ Q \circ \sigma_{1/h},$$

where

$$\sigma_h f(x) = f(x/h).$$

We call the operators with kernel of type (3) shift invariant. Examples of such operators are:

- spline operators: the Ciesielski–Durrmeyer operator (see [C]), a quasi-projection (see [Dz1]), an orthogonal projection (see [BD2], [BHR]),


\textbf{Key words and phrases:} shift invariant operators, asymptotic formula, density estimators, central limit, AMISE, asymptotic mean integral square error.
an orthogonal projection based on multiresolution approximation [M],
operators based on shift invariant spaces (see [JZ] and [BDR]; in particular
shift invariant spaces constructed by a function which satisfies the Strang–
Fix conditions, see [SF]).

Let $W^r_p$ be a Sobolev space (for details see [M]). Let $C^r_0$ be the space of
$r$-differentiable functions with compact support. Set
\[
|f|_{r,p} = \sum_{|\beta|=r} \|D^\beta f\|_p, \quad \|f\|_p = \left( \int_{\mathbb{R}^d} |f|^p \right)^{1/p},
\]
\[
D^\beta f = \frac{\partial^{|\beta|} f}{\partial x_1^{\beta_1} \cdots \partial x_d^{\beta_d}}, \quad \beta = (\beta_1, \ldots, \beta_d), \quad |\beta| = \beta_1 + \ldots + \beta_d.
\]
Assume the operator $Q$ reproduces all polynomials of degree less than $r$, i.e.
$Q(P) = P$ provided $\text{deg } P < r$. We then say that $Q$ has polynomial order $r$.
The following theorem is a generalization of [BHR, Proposition 4, p. 63].

**Theorem 1.1.** Let $1 \leq p < \infty$. Assume that $Q$ has polynomial order $r$. Then there is a constant $C(p) > 0$ such that for all $f \in W^r_p(\mathbb{R}^d)$,
\[
\|Q_h f - f\|_p \leq C(p) h^r |f|_{r,p}, \tag{5}
\]

**Proof.** Since the operators $Q_h$ are bounded from $L^p$ to $L^p$ it is sufficient
to prove (5) for $f \in C^r_0$. Let $f \in C^r_0$. Let $P_x$ be the Taylor polynomial of $f$ of
degree $r-1$ at $x$. Note that $f(x) = P_x(x)$ and $Q_h f(x) - f(x) = Q_h (f - P_x)(x)$. Now Lemma 1.1 below yields (5) for $1 \leq p < \infty$. $\blacksquare$

An easy computation shows the assertion for $p = \infty$ (see proof of [Dz4, Theorem 9.7]).

In statistics we need an asymptotic formula for the error in shift invariant
operators. Such a formula was proved in [BD3], [BD4] for an interpolation
operator and an orthogonal projection. Those proofs are based on a general-
ization of Mazur–Orlicz’s theorem (see [BD3]). This theorem goes back to
L. Fejér. Recall that a function $g$ defined on $\mathbb{R}^d$ is called $\mathbb{Z}^d$-periodic if for
all $x \in \mathbb{R}^d$,
\[
g(x) = g(x + \alpha) \quad \text{for all } \alpha \in \mathbb{Z}^d. \tag{6}
\]

**Theorem 1.2** (Mazur–Orlicz [MO]). If for $j = 1, \ldots, m$, $g_j$ are mea-
surable, bounded, $\mathbb{Z}^d$-periodic functions and $f_j$ are measurable functions with
\[
\int_{\mathbb{R}^d} |f_j(x)|^p \, dx < \infty
\]
for some $1 \leq p < \infty$, then
\[
\int_{\mathbb{R}^d} \left| \sum_{j=1}^m f_j(x) g_j(x/h) \right|^p \, dx \to \int_{[0,1]^d} \int_{\mathbb{R}^d} \left| \sum_{j=1}^m f_j(t) g_j(x) \right|^p \, dt \, dx \quad \text{as } h \to 0. \tag{7}
\]
Earlier results concerning the asymptotic formula can be found in [C], [Dz1] for spline operators, and in [DU], [BD2] for an orthogonal projection in $L^2$. See also [DLP]. Let $\beta(x) = x_1^{\beta_1} \ldots x_d^{\beta_d}$. We present a new and simpler proof of the asymptotic formula for the error in shift invariant operators.

**Theorem 1.3.** Assume that $Q$ has maximal polynomial order $q$. Let $1 \leq p < \infty$ and $f \in W_p^q(\mathbb{R}^d)$. Then

$$\lim_{h \to 0^+} \left\| \frac{Q_h f - f}{h} \right\|^p_p = \int_{\mathbb{R}^d} \left( \int_{[0,1]^d} \left| \sum_{|\beta| = q} \frac{1}{\beta!} D^\beta f(t)(Q(\beta)(x) - x^\beta) \right|^p \, dx \right) \, dt. \tag{8}$$

**Proof.** It is sufficient to prove (8) for the dense subset $C_0^{q+1}$ of $W_p^q(\mathbb{R}^d)$ since

$$\left\| \frac{Q_h f - f}{h} \right\|^p_p \leq C|f|_{q,p}.$$

Fix $f \in C_0^{q+1}$. Let $P_x$ be the Taylor polynomial of degree $q$ of $f$ at $x$. By the triangle inequality (we take $F(x) = Q_h(P_x)(x) \neq P_x(x)$)

$$\left\| \frac{Q_h f - f}{h} \right\|^p_p \leq \left\| \frac{Q_h(f - P_x)}{h} \right\|^p_p + \left\| \frac{Q_h P_x - P_x}{h} \right\|^p_p$$

and

$$\left\| \frac{Q_h P_x - P_x}{h} \right\|^p_p \leq \left\| \frac{Q_h f - f}{h} \right\|^p_p + \left\| \frac{Q_h(f - P_x)}{h} \right\|^p_p.$$

If we prove that there is $C$ such that for all $f \in C_0^{q+1}$,

$$\|Q_h(f - P_x)\|_p \leq C h^{q+1} |f|_{q+1,p}, \tag{9}$$

then the proof of (8) is completed by showing that

$$\lim_{h \to 0^+} \left\| \frac{Q_h P_x - P_x}{h} \right\|^p_p = \int_{\mathbb{R}^d} \left( \int_{[0,1]^d} \left| \sum_{|\beta| = q} \frac{1}{\beta!} D^\beta f(t)(Q(\beta)(x) - x^\beta) \right|^p \, dx \right) \, dt. \tag{10}$$

The technical proof of (9) is postponed to Lemma 1.1. Let

$$P_x = T_x + R_x,$$

where $T_x$ is homogeneous of degree $q$ and $\deg R_x < q$. Since $Q(R_x) = R_x$ we have

$$\frac{Q_h(P_x)(t) - P_x(t)}{h} = \frac{Q_h(T_x)(t) - T_x(t)}{h} = Q(T_x)(t/h) - T_x(t/h)$$

$$= \sum_{|\beta| = q} \frac{1}{\beta!} D^\beta f(x)(Q(\beta)(t/h) - (t/h)^{\beta}). \tag{11}$$
Consequently, from (11) we get
\[
\left\| \frac{Q_h P - P}{h^\theta} \right\|_p^p = \int_{\mathbb{R}^d} \left| \sum_{|\beta| = \theta} \frac{1}{\beta!} D^\beta f(x)(Q(\beta)(x/h) - (x/h)^\beta) \right|_p^p \, dx.
\]

An easy calculation shows (cf. [Dz3, Lemma 3.3]) that the functions
\[
Q(\beta)(x) - x^\beta = (-1)^{|\beta|} \sum_{\alpha \in \mathbb{Z}^d} \int (x - y)^\beta F(y - \alpha) \, dy G(x - \alpha)
\]
are \(\mathbb{Z}^d\)-periodic. Now the Mazur–Orlicz Theorem (7) implies (10).

**Lemma 1.1.** Let \(1 \leq p < \infty\). Let \(P_x\) be the Taylor polynomial of degree \(k - 1\) of a function \(f\). There is \(C\) such that for all \(f \in C^k_0\),
\[
(12) \quad \|Q_h(f - P)\|_p \leq C h^k |f|_{k,p}.
\]

**Proof.** By Taylor’s formula,
\[
\left\| Q_h(f - P) \right\|_p \leq \int_{\mathbb{R}^d} \left| \sum_{\alpha \in \mathbb{Z}^d} \sum_{\beta, |\beta| = k} \frac{1}{\beta!} D^\beta f(x + s(hy - x))(1 - s)^{k-1} ds
\]
\[
\times |h y - x|^\beta F(y - \alpha) \, dy G(x/h - \alpha) \right|_p^p \, dx.
\]

To prove (12), using assumption (1), it is sufficient to estimate
\[
J_\beta = \int_{\mathbb{R}^d} \left| \sum_{\alpha \in \mathbb{Z}^d} \sum_{\beta, |\beta| = k} D^\beta f(x + s(hy - x)) ds
\]
\[
\times |h y - x|^\beta |q|^{y-\alpha} dy q^{|x/h-\alpha|} \right|_p^p \, dx.
\]

We apply Jensen’s inequality three times:
\[
\left( \int_0^1 g(s) \, ds \right)^p \leq \int_0^1 |g(x)|^p dx,
\]
\[
\left( \sum_{\alpha \in \mathbb{Z}^d} |a_\alpha| |q|^{x-\alpha} \right)^p \leq C_1 \sum_{\alpha \in \mathbb{Z}^d} |a_\alpha| |q|^{x-\alpha},
\]
where \(C_1\) is independent of \(x\), i.e. \(C_1 = \max_x (\sum_{\alpha \in \mathbb{Z}^d} q^{x-\alpha})^{p-1}\),
\[
\left( \int_{\mathbb{R}^d} |g(y)| |q|^{y-\alpha} \, dy \right)^p \leq C_2 \int_{\mathbb{R}^d} |g(y)| |q|^{y-\alpha} \, dy,
\]
where \(C_2\) is independent of \(\alpha\). Consequently,
\[
J_\beta \leq C \sum_{\alpha \in \mathbb{Z}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left| D^\beta f(x + s(hy - x)) \right|^p ds
\]
\[
\times |h y - x|^\beta |q|^{y-\alpha} \, dy q^{|x/h-\alpha|} \, dx.
\]
Letting \( x/h - \alpha = u \) yields

\[
J_\beta \leq \left( \sum_{\alpha \in \mathbb{Z}^d} \int \int_0^1 \left| D^\beta f(hu + h\alpha + s(hy - hu - h\alpha)) \right|^p \, ds \times |hy - hu - h\alpha|^p \right) q^{y-\alpha} \, dy q^{|u|} \, du
\]

and by obvious changes of variables

\[
J_\beta \leq \left( \sum_{\alpha \in \mathbb{Z}^d} \int \int_0^1 \left| D^\beta f(hu + h\alpha + sh(z - u)) \right|^p \, ds \times h^{pk} |z - u|^p q^{|z|} \, dz q^{|u|} \, du
\]

\[
= Ch^{d+pk} \sum_{\alpha \in \mathbb{Z}^d} \int \int_0^1 \left| D^\beta f(hu + h\alpha + shv) \right|^p \, ds |v|^p q^{u+v} \, dv q^{|u|} \, du.
\]

Let us split the integrals:

\[
J_\beta \leq Ch^{d+pk} \sum_{n=1}^{\infty} \sum_{\alpha \in \mathbb{Z}^d} \int_0^1 \int_0^1 \left| D^\beta f(hu + h\alpha + shv) \right|^p \, ds |v|^p q^{u+v} \, dv q^{n-1} \, du
\]

\[
\leq Ch^{d+pk} \sum_{n=1}^{\infty} \sum_{\alpha \in \mathbb{Z}^d} \int_0^1 \int_0^1 |D^\beta f(hu + h\alpha + shv)|^p \, ds |v|^p q^{u+v} \, dv q^{n-1} \, du
\]

Note that if \(|v + u| < j\) and \(|u| < n\) then

\(|v| < |v + u| + |u| < j + n\).

Thus

\[
J_\beta \leq Ch^{d+pk} \sum_{n=1}^{\infty} q^{n-1} \sum_{\alpha \in \mathbb{Z}^d} \int_0^1 \int_0^1 |D^\beta f(hu + h\alpha + shv)|^p \, ds |v|^p q^{u+v} \, dv q^{n-1} \, du.
\]

Changing the order of the integrations we get

\[
J_\beta \leq Ch^{d+pk} \sum_{n=1}^{\infty} q^{n-1} \sum_{j=1}^{\infty} q^{j-1} \\
\times \int_0^1 \int_0^1 \int_0^1 |D^\beta f(hu + h\alpha + shv)|^p \, du \, ds |v|^p q^{u+v} \, dv.
\]
Note that if $|v| < j + n$ then
\[
    h^d \sum_{\alpha \in \mathbb{Z}^d} \int_{|u| < n} |D^\beta f(hu + h\alpha + shv)|^p \, du
\]
\[
    \leq \sum_{\alpha \in \mathbb{Z}^d} \int_{|x - h\alpha| < h(2n + j)} |D^\beta f(x)|^p \, dx
\]
and moreover
\[
    \int_{|v| < j + n} |v|^{pk} \, dv = C(j + n)^{pk + d}.
\]
Consequently,
\[
    J_\beta \leq C \sum_{n=1}^{\infty} q^{n-1} h^{pk} \sum_{j=1}^{\infty} q^{j-1} (4n + 2j)^d (j + n)^{pk + d} \int_{\mathbb{R}^d} |D^\beta f(x)|^p \, dx
\]
\[
    \leq Ch^{pk} \int_{\mathbb{R}^d} |D^\beta f|^p.
\]
This finishes the proof of the lemma. ■

Let $X_1, \ldots, X_n$ be a random sample from a distribution with density $f \in W_2^\beta$. We define a density estimator based on the kernel $K$ by
\[
    f_{h,n}(x) = \frac{1}{n} \sum_{j=1}^{n} K_h(x, X_j),
\]
where
\[
    K_h(x, y) = (1/h)^d K(x/h, y/h).
\]
Note that
\[
    Ef_{h,n} = Q_h f.
\]
As usual we consider the estimation error given by
\[
    \text{MISE}(f, h) = E \left[ \int_{\mathbb{R}^d} [f_{h,n} - f]^2 \right].
\]
It is known that
\[
    \text{MISE}(f, h) = E \left[ \int_{\mathbb{R}^d} [f_{h,n} - Q_h f]^2 \right] + \int_{\mathbb{R}^d} [Q_h f - f]^2.
\]
The asymptotic formula for the second factor in (16) is given in (8). We prove that

**Theorem 1.4.** Assume that $Q$ has maximal polynomial order $q > 0$. If $nh^d \to \infty$, $h \to 0$ then
\[
    \lim_{nh^d \to \infty} nh^d E \left[ \int_{\mathbb{R}^d} [f_{h,n} - Q_h f]^2 \right] = \int_{\mathbb{R}^d} \left[ \int_{[0,1]^d} K^2(x, y) \, dy \right] dx,
\]
where
\begin{equation}
(18) \quad \int_{\mathbb{R}^d} \left[ \int_{[0,1]^d} K^2(x, y) \, dy \right] \, dx = \sum_{\alpha \in \mathbb{Z}^d} \eta(\alpha) \xi(\alpha)
\end{equation}
and
\[ \eta = G \ast \tilde{G}, \quad \xi = F \ast \tilde{F}, \quad \tilde{G}(x) = G(-x), \quad \tilde{F}(x) = F(-x). \]

**Proof.** Note that
\[
\begin{align*}
E \left[ \int_{\mathbb{R}^d} [f_{h,n} - Ef_{h,n}]^2 \right] &= \frac{1}{n^2} \sum_{j=1}^n \int_{\mathbb{R}^d} E[K_h(x, X_j) - EK_h(x, X_j)]^2 \, dx \\
&= \frac{1}{n^2} \sum_{j=1}^n \left( E[K_h^2(x, X_j)] - [EK_h(x, X_j) ]^2 \right) \, dx.
\end{align*}
\]

If \( h \to 0 \) then by (5),
\[
\int_{\mathbb{R}^d} [EK_h(x, X_j)]^2 \, dx = \int_{\mathbb{R}^d} (Q_h f)^2 \to \int_{\mathbb{R}^d} f^2.
\]

On the other hand
\[
\int_{\mathbb{R}^d} EK_h^2(x, X_j) \, dx = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_h^2(x, y) f(y) \, dy \, dx.
\]

From Fubini’s theorem
\[
\int_{\mathbb{R}^d} EK_h^2(x, X_j) \, dx = \frac{1}{h^d} \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} K^2(u, y/h) \, du \right] f(y) \, dy.
\]

Note that for all \( \alpha \in \mathbb{Z}^d \),
\begin{equation}
(19) \quad \int_{\mathbb{R}^d} K^2(x, y + \alpha) \, dx = \int_{\mathbb{R}^d} K^2(x, y) \, dx.
\end{equation}

From Mazur–Orlicz’s theorem we get
\[
\lim_{h \to 0} \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} K^2(u, y/h) \, du \right] f(y) \, dy = \int_{[0,1]^d} \left[ \int_{\mathbb{R}^d} K^2(u, y) \, du \right] dy \int_{\mathbb{R}^d} f(y) \, dy.
\]

We thus get (17). A simple calculation leads to (18). \( \blacksquare \)

**Remarks.** 1. From (16)–(8) we get
\[
\text{MISE}(f, h) \sim \text{AMISE} := \frac{1}{nh^d} \int_{\mathbb{R}^d} \left[ \int_{[0,1]^d} K^2(x, y) \, dy \right] \, dx \\
+ h^2 \int_{\mathbb{R}^d} \left( \int_{[0,1]^d} \left[ \sum_{|\beta| = \alpha} 1 \frac{1}{\beta!} D^\beta f(t)Q(\beta)(x) - x^\beta \right]^2 \, dx \right) \, dt.
\]
So the best choice of $h > 0$ which minimizes (16) is

$$h \sim n^{-1/(2\alpha+d)}.$$  

2. Using the methods of [Dz2] one can prove the central limit theorem. This theorem generalizes the results for wavelet estimators [DL1]–[DL2] in $\mathbb{R}^d$ and box spline estimators [Dz2]. These results are motivated by the result for the Rosenblatt–Parzen estimator [H].

References


