PENALIZED LEAST SQUARES FITTING

Manfred von Golitschek and Larry L. Schumaker

Communicated by P. P. Petrushev

Dedicated to the memory of our colleague Vasil Popov
January 14, 1942 – May 31, 1990

Abstract. Bounds on the error of certain penalized least squares data fitting methods are derived. In addition to general results in a fairly abstract setting, more detailed results are included for several particularly interesting special cases, including splines in both one and several variables.

1. Introduction. We begin with an abstract definition of what we mean by a penalized least squares fit. Suppose $X$, $Y$ and $S$ are linear spaces over $\mathbb{R}$, where $S \subseteq Y \subseteq X$. Let $\| \cdot \|_X : X \to \mathbb{R}$ and $\| \cdot \|_Y : Y \to \mathbb{R}$ be semi-norms induced by semi-definite inner products $\langle \cdot, \cdot \rangle$ on $X$ and $[\cdot, \cdot]$ on $Y$, respectively. Given $f \in X$ and $\lambda > 0$, suppose there exists $s_\lambda(f, S)$ in $S$ such that

$$\Phi(s_\lambda(f, S)) = \min_{u \in S} \Phi(u),$$

2000 Mathematics Subject Classification: Primary 41A15; Secondary 41A25.
Key words: Penalized least squares, Tikhonov regularization, Spline approximation.
*Supported by the Army Research Office under grant DAAD-19-02-10059.
where

\[ \Phi(u) := \|f - u\|_X^2 + \lambda \|u\|_Y^2. \]

Then we call \( s_\lambda(f, S) \) a penalized least squares fit of \( f \) corresponding to \( \lambda \). If there exists a unique minimizer of \( \Phi \) in \( S \) for each \( f \in X \), then \( Q_\lambda : X \to S \) defined by \( Q_\lambda f := s_\lambda(f, S) \) defines a linear operator which is not in general a linear projection. Our aim in this paper is to investigate the behavior of \( \|f - Q_\lambda f\|_X \) as a function of both \( \lambda \) and the approximation properties of \( S \).

The paper is organized as follows. In Section 2 we give a detailed treatment of the penalized least squares problem for approximation by trigonometric polynomials. The next section is devoted to Tikhonov regularization (with weights) for univariate functions defined on an interval. The results include error bounds for certain classes of functions. In Section 4 we derive an explicit formula for the Tikhonov regularization of arbitrary functions in \( C[-\pi, \pi] \). Our main \( L_2 \) error bounds for penalized least squares fits of univariate functions are contained in Section 5. General penalized least squares is treated in Section 6. We conclude the paper by outlining two typical applications: univariate splines and bivariate splines on triangulations.

2. Penalized Fourier series. Throughout this section, we take \( X = L_2[-\pi, \pi] \) with the inner product \( \langle f, g \rangle = \int_{-\pi}^{\pi} fg \), and set \( Y = W_2^2[\mathbb{T}] \) with the semi-definite inner product \( \langle f, g \rangle = \int_{-\pi}^{\pi} f''g'' \), where \( \mathbb{T} \) denotes the circle. We let \( S = T_n \) be the set of all real-valued trigonometric polynomials of degree at most \( n \), with \( n \in \mathbb{N} \) fixed. Clearly, for each \( f \in L_2[-\pi, \pi] \) and \( \lambda > 0 \), there exists a unique minimal solution \( s_{\lambda,n} := s_{\lambda,n}(f) \) of the penalized least squares problem

\[ \min_{T \in T_n} \left\{ \int_{-\pi}^{\pi} (f - T)^2 \, dx + \lambda \int_{-\pi}^{\pi} (T'')^2 \, dx \right\}. \]

Our first theorem compares \( s_{\lambda,n} \) with the \( n \)-th partial sum \( s_n(f) := s_{0,n}(f) \) of the Fourier series of \( f \).

**Theorem 2.1.** Suppose \( f \in L_2[-\pi, \pi] \), and that

\[ f(x) = a_0 + \sum_{k=1}^{\infty} \left( a_k \cos(kx) + b_k \sin(kx) \right) \]
is its Fourier series. Then for each $\lambda > 0$ and $n \in \mathbb{N}$, $s_{\lambda,n} = s_\lambda(f, T_n)$ is given by

$$s_{\lambda,n}(x) = a_0 + \sum_{k=1}^{n} \frac{1}{1 + \lambda k^4} \left( a_k \cos(kx) + b_k \sin(kx) \right),$$

or equivalently,

$$s_n(f) = s_{\lambda,n} + \lambda s^{(4)}_{\lambda,n}.$$  \hspace{1cm} (2.2)

**Proof.** The minimal solution $s_{\lambda,n}$ is characterized by the orthogonality relations

$$\int_{-\pi}^{\pi} (f - s_{\lambda,n})T = \lambda \int_{-\pi}^{\pi} s'_{\lambda,n}T'', \quad \text{for all } T \in T_n. \hspace{1cm} (2.3)$$

A simple computation shows that the function $s_{\lambda,n}$ given by (2.1) is the only function in $T_n$ which satisfies (2.3). This implies that $s_{\lambda,n}$ is the unique minimal solution. The proof of (2.2) is equally simple. \qed

Taking the limit as $n \to \infty$ in Theorem 2.1, we get

**Corollary 2.2.** Let $f$ be as in Theorem 2.1. Then for each $\lambda > 0$, the function

$$s_{\lambda,\infty}(x) = a_0 + \sum_{k=1}^{\infty} \frac{1}{1 + \lambda k^4} \left( a_k \cos(kx) + b_k \sin(kx) \right), \quad x \in \mathbb{R},$$

lies in $W^4_2(\mathbb{T})$, and

$$f = s_{\lambda,\infty} + \lambda s^{(4)}_{\lambda,\infty}.$$  \hspace{1cm} \text{Moreover, } s_{\lambda,\infty} \text{ is the unique solution of the minimum problem}

$$\min \left\{ \int_{-\pi}^{\pi} (f - g)^2 dx + \lambda \int_{-\pi}^{\pi} (g'')^2 dx, \quad g \in W^4_2(\mathbb{T}) \right\}. \hspace{1cm}$$

Let $\| \cdot \|_2$ and $\| \cdot \|_\infty$ denote the $L_2$ and uniform norms on $[-\pi, \pi]$, respectively. Since $\| f - s_n(f) \|_2 \leq \| f - T \|_2$ for all $T \in T_n$, we have $\| f - s_{\lambda,n} \|_2 \geq \| f - s_n(f) \|_2$. In the next theorem we will see that the $L_2$-norm $\| f - s_{\lambda,n}(f) \|_2$ changes very little for increasing $n > \lambda^{-1/4}$. Hence, for numerical purposes, it is not worthwhile to compute $s_{\lambda,n}(f)$ for $n$ larger than $\lambda^{-1/4}$. 

Theorem 2.3. Let \( f \in L^2[-\pi, \pi] \) and \( 0 < \lambda \leq 1 \). Let \( N_\lambda = \lfloor \lambda^{-1/4} \rfloor \) be the largest integer smaller than or equal to \( \lambda^{-1/4} \). Then

\[
\frac{1}{2} \sqrt{\varepsilon_{\lambda,n}} \leq \| f - s_{\lambda,n}(f) \|_2 \leq \sqrt{\varepsilon_{\lambda,n}},
\]

where \( \varepsilon_{\lambda,n} := \varepsilon_{\lambda,n}(f) \) is defined by

\[
\varepsilon_{\lambda,n} := \begin{cases} 
\| f - s_n(f) \|_2^2 + \lambda^2 \| s_n^{(4)}(f) \|_2^2, & \text{if } n \leq N_\lambda, \\
\varepsilon_{\lambda,N_\lambda}, & \text{if } n > N_\lambda.
\end{cases}
\]

Proof. We apply the orthogonality relation

\[
\int_{-\pi}^{\pi} (f - s_n(f)) T = 0, \quad T \in \mathcal{T}_n,
\]

for \( T = s_n^{(4)} = s_n^{(4)}(f) \). This gives \( \int_{-\pi}^{\pi} (f - s_n(f)) s_{\lambda,n} = 0 \), and (2.2) implies

\[
\| f - s_{\lambda,n} \|_2^2 = \| f - s_n(f) \|_2^2 + \| f - s_n(f) \|_2^2 + \lambda^2 \| s_{\lambda,n} \|_2^2.
\]

It follows from (2.1) that

\[
\| s_{\lambda,n} \|_2^2 = \left\| \sum_{k=1}^{n} \frac{k^4}{1 + \lambda k^4} (a_k \cos (kx) + b_k \sin (kx)) \right\|_2^2 = \pi \sum_{k=1}^{n} \frac{k^8 (a_k^2 + b_k^2)}{(1 + \lambda k^4)^2}.
\]

If \( n \leq N_\lambda \), this implies

\[
\| s_{\lambda,n} \|_2^2 \leq \pi \sum_{k=1}^{n} k^8 (a_k^2 + b_k^2) = \| s_n^{(4)} \|_2^2
\]

\[
\| s_{\lambda,n} \|_2^2 \geq \frac{\pi}{4} \sum_{k=1}^{n} k^8 (a_k^2 + b_k^2) = \frac{1}{4} \| s_n^{(4)} \|_2^2.
\]

If \( n > N_\lambda \), then

\[
\| s_{\lambda,n} \|_2^2 = \| s_{\lambda,N_\lambda} \|_2^2 + \pi \sum_{k=N_\lambda+1}^{n} \frac{k^8 (a_k^2 + b_k^2)}{(1 + \lambda k^4)^2} \leq \| s_{\lambda,N_\lambda} \|_2^2 + \pi \sum_{k=N_\lambda+1}^{n} \frac{a_k^2 + b_k^2}{\lambda^2} = \| s_{\lambda,N_\lambda} \|_2^2 + \lambda^{-2} \left( \| f - s_{N_\lambda} \|_2^2 - \| f - s_n \|_2^2 \right).
\]
Moreover, if \( n > N_\lambda \), then
\[
\| s^{(4)}_{\lambda,n} \|_2^2 = \| s^{(4)}_{\lambda,N_\lambda} \|_2^2 + \pi \sum_{k=N_\lambda+1}^n \frac{k^8 (a_k^2 + b_k^2)}{(1 + \lambda k^4)^2} \\
\geq \| s^{(4)}_{\lambda,N_\lambda} \|_2^2 + \pi \sum_{k=N_\lambda+1}^n \frac{a_k^2 + b_k^2}{4\lambda^2} \\
= \| s^{(4)}_{\lambda,N_\lambda} \|_2^2 + \frac{1}{4\lambda^2} \left( \| f - s_{N_\lambda} \|_2^2 - \| f - s_n \|_2^2 \right).
\]

Inserting these estimates into (2.5) yields (2.4).

3. Tikhonov regularization on \([a, b]\) with weights. Suppose \( w : [a, b] \rightarrow \mathbb{R} \) is a piecewise continuous nonnegative function on \([a, b]\) with \( \int_a^b w > 0 \).

In this section we take \( X = L_2(a, b) \) with the inner product \( \langle f, g \rangle = \int_a^b w fg \), and set \( Y = S = W_2^2[a, b] \) with the semi-definite inner product \( \langle f, g \rangle = \int_a^b f'^{\prime\prime}g'' \).

**Definition 3.1.** Let \( f \in L_2(a, b) \) and \( \lambda > 0 \). We call \( \phi_{\lambda,w} = \phi_{\lambda}(f, w) \in W_2^2[a, b] \) the (nonperiodic) Tikhonov regularization of \( f \) corresponding to \( \lambda \) and \( w \) provided it minimizes
\[
\int_a^b w(x)(f - u)^2 \, dx + \lambda \int_a^b (u'')^2 \, dx 
\]
with respect to all \( u \in W_2^2[a, b] \).

For the weight \( w = 1 \), Ragozin [5] proved that the Tikhonov regularization \( \phi_{\lambda} \) satisfies
\[
\| f^{(q)} - \phi_{\lambda}^{(q)} \|_2 \leq C_q \lambda^{(2-q)/4} \| f'' \|_2, \quad q = 0, 1, 2,
\]
for some absolute constants \( C_q > 0 \). The next theorem states, among other things, that \( C_1 \leq 1 \) and \( C_2 \leq 2 \) for the weight \( w = 1 \).

**Theorem 3.2.** Let \( f \in C[a, b] \) and \( \lambda > 0 \). Let \( \phi_{\lambda} := \phi_{\lambda}(f, w) \) be the Tikhonov regularization of \( f \) corresponding to \( \lambda, w \). Then \( \phi_{\lambda} \in C^3[a, b] \), \( \phi_{\lambda}^{(4)} \) is piecewise continuous on \([a, b]\), and
1) \( w(x)[f(x) - \phi_\lambda(x)] = \lambda \phi_\lambda^{(4)}(x), \) for all \( x \in [a, b] \) where \( w \) is continuous,
2) \( (\phi_\lambda)^{(4)}(a) = (\phi_\lambda)^{(4)}(b) = \phi_\lambda^{(3)}(a) = \phi_\lambda^{(3)}(b) = 0. \)

If \( f \in W_2^2[a, b] \), then \( \| (\phi_\lambda)^{''} \|_2 \leq \| f'' \|_2 \) and

\[
\left( \int_a^b w(f - \phi_\lambda)^2 \right)^{1/2} \leq \lambda^{1/2} \| f'' \|_2.
\]

**Proof.** Let \( f \in C[a, b] \). The orthogonality relations for (3.1) are

\[
(3.2) \quad \int_a^b w(f - \phi_\lambda)u - \lambda \int_a^b (\phi_\lambda)^{''}u'' = 0, \quad \text{for all } u \in W_2^2[a, b].
\]

We define the function \( F \) by

\[
F(x) := \int_a^x \int_a^t w(s)(f(s) - \phi_\lambda(s))ds\, dt, \quad a \leq x \leq b.
\]

Then, \( F(a) = F'(a) = 0 \). Applying (3.2) for \( u = 1 \) and \( u(x) = x \) yields

\[
\int_a^b w(f - \phi_\lambda) dx = \int_a^b (f(x) - \phi_\lambda(x)) x dx = 0,
\]

and thus \( F'(b) = 0 \) and \( F(b) = 0 \). Therefore, integrating by parts gives

\[
\int_a^b w(f - \phi_\lambda)u = \int_a^b Fu''.
\]

Hence (3.2) is equivalent to

\[
\int_a^b (F - \lambda(\phi_\lambda)^{''})u'' = 0, \quad \text{for all } u \in W_2^2[a, b],
\]

and thus

\[
\int_a^b (F - \lambda(\phi_\lambda)^{''})g = 0, \quad \text{for all } g \in L_2[a, b].
\]

This implies that

\[
F(x) = \lambda(\phi_\lambda)^{''}(x), \quad a \leq x \leq b.
\]

Differentiating twice yields

\[
w(x)(f(x) - \phi_\lambda(x)) = \lambda \phi_\lambda^{(4)}(x)
\]
at all points $x \in [a,b]$ where $w$ is continuous. Then $F = \lambda(\phi_\lambda)''$ and $F(a) = F(b) = 0$ imply $(\phi_\lambda)''(a) = (\phi_\lambda)''(b) = 0$, while $F' = \lambda(\phi_\lambda)'$ and $F'(a) = F'(b) = 0$ imply $\phi_\lambda^{(3)}(a) = \phi_\lambda^{(3)}(b) = 0$.

Let $f \in W_2^4[a,b]$. We denote the minimal value in (3.1) by $M$. Then (3.1) and (3.2) imply that

$$M = \int_a^b w(f - \phi_\lambda)(f - u) + \lambda \int_a^b (\phi_\lambda)''u''$$

for all $u \in W_2^4[a,b]$. Taking $u = f$, it follows that

$$M = \lambda \int_a^b (\phi_\lambda)'' f'' \leq \lambda \|(\phi_\lambda)''\|_2 \|f''\|_2.$$ 

Therefore,

$$\int_a^b w(f - \phi_\lambda)^2 + \lambda \|(\phi_\lambda)''\|_2^2 \leq \lambda \|(\phi_\lambda)''\|_2 \|f''\|_2,$$

and thus

$$\|(\phi_\lambda)''\|_2 \leq \|f''\|_2; \quad \|\sqrt{w}(f - \phi_\lambda)\|_2 \leq \lambda^{1/2}\|f''\|_2.$$ 

This concludes the proof. □

**Theorem 3.3.** Let $w = 1$. Let $f \in W_2^4[a,b]$ and $f''(a) = f''(b) = f^{(3)}(a) = f^{(3)}(b) = 0$. Then the Tikhonov regularization $\phi_\lambda$ of $f$ corresponding to $\lambda > 0$ satisfies

$$\|f - \phi_\lambda\|_2 \leq \lambda\|f^{(4)}\|_2,$$

$$\|f'' - (\phi_\lambda)''\|_2 \leq \lambda^{1/2}\|f^{(4)}\|_2,$$

$$\|\phi_\lambda^{(4)}\|_2 \leq \|f^{(4)}\|_2.$$ 

**Proof.** We write $U := \{u \in W_2^4[a,b] : u''(a) = u''(b) = u^{(3)}(a) = u^{(3)}(b) = 0\}$. Clearly, $f \in U$ by assumption and $\phi_\lambda \in U$ by Theorem 3.2. For each $u \in U$ we obtain by integration by parts that

$$\int_a^b (f - \phi_\lambda)u^{(4)} = \int_a^b (f'' - (\phi_\lambda)''u'' = 0.$$ 

Penalized Least Squares Fitting 335
Inserting the identity $f - \phi_\lambda = \lambda \phi_\lambda^{(4)}$, it follows that

\begin{equation}
\int_a^b (f'' - (\phi_\lambda)''u'' - \lambda \int_a^b \phi_\lambda^{(4)}u^{(4)}) = 0, \quad \text{for all } u \in U.
\end{equation}

But (3.3) are the orthogonality relations for the minimum problem

\begin{equation}
M^* := \min_{u \in U} \|f'' - u''\|_2^2 + \lambda \|u^{(4)}\|_2^2,
\end{equation}

and so $\phi_\lambda$ is the minimal solution of (3.4). Since $f \in U$, it follows that $M^* \leq \lambda \|f^{(4)}\|_2^2$, and thus

\[ \|\phi_\lambda^{(4)}\|_2 \leq \|f^{(4)}\|_2, \quad \|f'' - (\phi_\lambda)''\|_2 \leq \lambda^{1/2} \|f^{(4)}\|_2. \]

Finally, we have

\[ \|f - \phi_\lambda\|_2 = \lambda \|\phi_\lambda^{(4)}\|_2 \leq \lambda \|f^{(4)}\|_2, \]

which concludes the proof of the theorem. \(\square\)

4. An explicit formula for the Tikhonov regularization

In this section we consider the inner product spaces $X = L_2(a, b)$ and $Y = S = W_2^2[a, b]$ of Section 3 with the weight function $w = 1$ on $[a, b]$, that is, $\langle f, g \rangle = \int_a^b fg$, $[f, g] = f_a^b f''''g''$. In Theorem 2.1 we have given an explicit formula for the periodic Tikhonov regularization. In this section we want to apply the methods of Section 2 to the nonperiodic case. We may assume for convenience that the interval is $[a, b] = [-\pi, \pi]$. Otherwise, under the linear transformation

\[ x(t) = a + \frac{(b - a)(t + \pi)}{2\pi}, \quad -\pi \leq t \leq \pi, \]

of the interval $[-\pi, \pi]$ onto $[a, b]$, the Tikhonov regularizations $\phi_\lambda(x)$ on $[a, b]$ of $f$ and $\psi_\mu(t)$ on $[-\pi, \pi]$ of $g(t) = f(x(t))$ are related by

\begin{equation}
\psi_\mu(t) = \phi_\lambda(x(t)), \quad \lambda = \left(\frac{b - a}{2\pi}\right)^4 \mu.
\end{equation}

To prove (4.1), we simply compare the orthogonality relations

\begin{equation}
\int_a^b (f(x) - \phi_\lambda(x))u(x)\, dx = \lambda \int_a^b (\phi_\lambda''(x)u''(x))\, dx, \quad u \in W_2^2[a, b],
\end{equation}
and

\begin{equation}
\int_{-\pi}^{\pi} (g(t) - \psi(t))v(t) \, dt = \mu \int_{-\pi}^{\pi} \psi''(t)v''(t) \, dt, \quad v \in W_2^2[-\pi, \pi],
\end{equation}

where the equivalence of (4.2) and (4.3) follows from the bijection \( v(t) = u(x(t)) \) of \( W_2^2[a, b] \) onto \( W_2^2[-\pi, \pi] \).

The main idea of this section is contained in the next lemma.

**Lemma 4.1** Each function \( v \in C^1[-\pi, \pi] \) has a unique representation of the form

\begin{equation}
v(t) = c_1 t + c_2 t^2 + w(t), \quad c_1 \in \mathbb{R}, \ c_2 \in \mathbb{R}, \ w \in C^1(\mathbb{T}).
\end{equation}

In other words, (4.4) defines a bijection from \( \mathbb{R}^2 \times C^1(\mathbb{T}) \) onto \( C^1[-\pi, \pi] \).

**Proof.** Given \( c_1, c_2 \in \mathbb{R} \) and \( w \in C^1(\mathbb{T}) \), the function \( v \) in (4.4) lies in \( C^1[-\pi, \pi] \). Conversely, for a function \( v \in C^1[-\pi, \pi] \), the condition \( w(-\pi) = w(\pi) \) is equivalent to

\[
v(-\pi) + c_1 \pi - c_2 \pi^2 = v(\pi) - c_1 \pi - c_2 \pi^2,
\]

and thus to

\[
c_1 = \frac{v(\pi) - v(-\pi)}{2\pi}.
\]

The condition \( w'(-\pi) = w'(\pi) \) is equivalent to \( v'(-\pi) + 2c_2 \pi = v'(\pi) - 2c_2 \pi \), and thus to

\[
c_2 = \frac{v'(\pi) - v'(-\pi)}{4\pi}.
\]

This concludes the proof of the lemma. \( \square \)

As a corollary, each function \( f \in C^1[-\pi, \pi] \) has a unique representation

\begin{equation}
f(t) = c_1 t + c_2 t^2 + a_0 + \sum_{k=1}^{\infty} (a_k \cos(kt) + b_k \sin(kt)),
\end{equation}

where

\[
c_1 = \frac{f(\pi) - f(-\pi)}{2\pi}, \quad c_2 = \frac{f'(\pi) - f'(-\pi)}{4\pi}.
\]
Theorem 4.2. Let \( f \in C[-\pi, \pi] \) be given by (4.5). If \( f \not\in C^1[-\pi, \pi] \), then we take \( c_2 = 0 \). Then the nonperiodic Tikhonov regularization \( \phi_\lambda \) of \( f \) corresponding to \( \lambda \) is given by

\[
\phi_\lambda(t) = \gamma_1 t + \gamma_2 t^2 + \alpha_0 + \sum_{k=1}^{\infty} (\alpha_k \cos(kt) + \beta_k \sin(kt))
\]

with

\[
\gamma_1 = c_1 - \frac{1}{2B(\lambda)} \sum_{k=1}^{\infty} \frac{(-1)^k k^3 b_k}{1 + \lambda k^4},
\]

\[
\gamma_2 = \frac{1}{2 + 4A(\lambda)} \left( 4A(\lambda) c_2 + \sum_{k=1}^{\infty} \frac{(-1)^k k^2 a_k}{1 + \lambda k^4} \right),
\]

\[
\alpha_0 = a_0 + \frac{(c_2 - \gamma_2) \pi^2}{3},
\]

\[
\alpha_k = \frac{1}{1 + \lambda k^4} \left( a_k + \frac{4(-1)^{k-1}(c_2 - \gamma_2)}{k^2} \right), \quad k \geq 1,
\]

\[
\beta_k = \frac{1}{1 + \lambda k^4} \left( b_k + \frac{2(-1)^{k-1}(c_1 - \gamma_1)}{k} \right), \quad k \geq 1,
\]

where

\[
A(\lambda) := \sum_{k=1}^{\infty} \frac{1}{1 + \lambda k^4}, \quad B(\lambda) := \sum_{k=1}^{\infty} \frac{k^2}{1 + \lambda k^4}.
\]

Proof. The Fourier series of \( t \) and \( t^2 \) on \((-\pi, \pi)\) are

\[
t = \sum_{k=1}^{\infty} \frac{2(-1)^{k-1} \sin(kt)}{k}, \quad -\pi < t < \pi,
\]

\[
t^2 = \frac{\pi^2}{3} + \sum_{k=1}^{\infty} \frac{4(-1)^k \cos(kt)}{k^2}, \quad -\pi \leq t \leq \pi,
\]

with pointwise convergence. From Theorem 3.2, we know that \( f = \phi_\lambda + \lambda \phi_\lambda^{(4)} \). Comparing the Fourier coefficients of \( f \) and \( \phi_\lambda + \lambda \phi_\lambda^{(4)} \) in the expansions

\[
f(t) = c_1 \sum_{k=1}^{\infty} \frac{2(-1)^{k-1} \sin(kt)}{k} + c_2 \left( \frac{\pi^2}{3} + \sum_{k=1}^{\infty} \frac{4(-1)^k \cos(kt)}{k^2} \right)
\]

\[+ a_0 + \sum_{k=1}^{\infty} (a_k \cos(kt) + b_k \sin(kt))
\]
Penalized Least Squares Fitting

\[ \phi_\lambda(t) + \lambda \phi_\lambda^{(4)}(t) = \gamma_1 \sum_{k=1}^{\infty} \frac{2(-1)^{k-1} \sin(k t)}{k} + \gamma_2 \left( \frac{\pi^2}{3} + \sum_{k=1}^{\infty} \frac{4(-1)^k \cos(k t)}{k^2} \right) + \alpha_0 + \sum_{k=1}^{\infty} (1 + \lambda k^4)(\alpha_k \cos(k t) + \beta_k \sin(k t)), \]

we obtain the equalities

\[ \frac{c_2 \pi^2}{3} + a_0 = \frac{\gamma_2 \pi^2}{3} + \alpha_0 \]
\[ \frac{4c_2 (-1)^k}{k^2} + a_k = \frac{4\gamma_2 (-1)^k}{k^2} + (1 + \lambda k^4)\alpha_k, \quad k \geq 1, \]
\[ \frac{2c_1 (-1)^{k-1}}{k} + b_k = \frac{2\gamma_1 (-1)^{k-1}}{k} + (1 + \lambda k^4)\beta_k, \quad k \geq 1, \]

which are equivalent to (4.8)–(4.10). From Theorem 3.2 we know that \((\phi_\lambda)''(\pi) = 0\) and \(\phi_\lambda^{(3)}(\pi) = 0\), and thus,

\[ (\phi_\lambda)''(\pi) = 2\gamma_2 - \sum_{k=1}^{\infty} (-1)^k k^2 \alpha_k = 0, \]
\[ (4.11) \]
\[ \phi_\lambda^{(3)}(\pi) = -\sum_{k=1}^{\infty} (-1)^k k^3 \beta_k = 0. \]

Inserting (4.8)–(4.10) into (4.11) yields

\[ 2\gamma_2 = \sum_{k=1}^{\infty} \frac{(-1)^k k^2 \alpha_k}{1 + \lambda k^4} + 4(c_2 - \gamma_2) \sum_{k=1}^{\infty} \frac{1}{1 + \lambda k^4} \]

and

\[ 0 = \sum_{k=1}^{\infty} (-1)^k k^3 \beta_k = \sum_{k=1}^{\infty} \frac{(-1)^k k^3 \beta_k}{1 + \lambda k^4} = 2(c_1 - \gamma_1) \sum_{k=1}^{\infty} \frac{k^2}{1 + \lambda k^4}. \]

These two conditions are equivalent to

\[ 2\gamma_2 = \sum_{k=1}^{\infty} \frac{(-1)^k k^2 \alpha_k}{1 + \lambda k^4} + 4(c_2 - \gamma_2) A(\lambda) \]
\[ (4.12) \]

and

\[ 2(c_1 - \gamma_1) B(\lambda) = \sum_{k=1}^{\infty} \frac{(-1)^k k^3 \beta_k}{1 + \lambda k^4}. \]
\[ (4.13) \]
Now (4.6) follows from (4.12), and (4.7) follows from (4.13). □

5. $L_2$-error bounds for penalized least squares fitting. In this section we consider the inner product spaces $X = L_2(a, b)$ and $Y = W^2_2[a, b]$ of Section 4 with the same inner products $\langle f, g \rangle = \int_a^b fg$ and $[f, g] = \int_a^b f''g''$, but we now take $S$ to be a proper subspace of $Y$. For $f \in L_2[a, b]$ and $\lambda > 0$, we compare the penalized least squares fit $\phi_\lambda \in Y$ of Section 3 (with weight function $w = 1$) satisfying

$$
\int_a^b (f - \phi_\lambda)^2 + \lambda \int_a^b (\phi''_\lambda)^2 = \min_{u \in Y} \left\{ \int_a^b (f - u)^2 + \lambda \int_a^b (u'')^2 \right\}
$$

with the penalized least squares fit $s_\lambda = Q_\lambda(f)$ from $S$ satisfying

$$
\int_a^b (f - s_\lambda)^2 + \lambda \int_a^b (s''_\lambda)^2 = \min_{u \in S} \left\{ \int_a^b (f - u)^2 + \lambda \int_a^b (u'')^2 \right\}.
$$

**Theorem 5.1.** Let $f \in W^2_2[a, b]$ and $\lambda > 0$. Then $s_\lambda$ minimizes the expression

$$(5.1) \quad \|\phi_\lambda - u\|_2^2 + \lambda \|(\phi_\lambda)' - u''\|_2^2$$

with respect to all $u \in S$.

**Proof.** $s_\lambda$ is characterized by the orthogonality relations

$$
\int_a^b (f - s_\lambda)u - \lambda \int_a^b (s''_\lambda)u'' = 0, \quad \text{for all } u \in S.
$$

(5.2)

Since $f = \phi_\lambda + \lambda \phi_\lambda^{(4)}$ by Theorem 3.2, (5.2) is equivalent to

$$
\int_a^b (\phi_\lambda - s_\lambda)u + \lambda \int_a^b (\phi_\lambda^{(4)}u - (s_\lambda)''u'') = 0, \quad \text{for all } u \in S,
$$

and since $\phi''_\lambda(a) = \phi''_\lambda(b) = \phi^{(3)}_\lambda(a) = \phi^{(3)}_\lambda(b) = 0$ by Theorem 3.2, integration by parts shows that (5.2) is equivalent to

$$
(5.3) \quad \int_a^b (\phi_\lambda - s_\lambda)u + \lambda \int_a^b (\phi''_\lambda - (s_\lambda)''u'') = 0, \quad \text{for all } u \in S.
$$

But (5.3) are the orthogonality relations for the minimum problem (5.1). □
Corollary 5.2. Let $S \subset W^2_2[a,b]$ be a linear space. Suppose that for each $g \in W^2_2[a,b]$ there exists $u \in S$ such that

$$
\|g - u\|_2 \leq Ch^2\|g''\|_2, \\
\|g'' - u''\|_2 \leq C\|g''\|_2,
$$

for positive numbers $C$ and $h$ independent of $g$. Then the penalized least squares fit $s_\lambda$ of $f$ in $S$ corresponding to $\lambda$ satisfies

$$
\|f - s_\lambda\|_2 \leq \lambda^{1/2}\left(1 + C\sqrt{h^4\lambda^{-1} + 1}\right)\|f''\|_2
$$

and

$$
\|(s_\lambda)''\|_2 \leq \left(1 + C\sqrt{h^4\lambda^{-1} + 1}\right)\|f''\|_2.
$$

Proof. By Theorem 3.2, the Tikhonov regularization $\phi_\lambda$ of $f$ satisfies

$$
\|f - \phi_\lambda\|_2 \leq \lambda^{1/2}\|f''\|_2.
$$

Let $M_0$ be the minimum in (5.1). Applying (5.4) for $g := \phi_\lambda$, it follows that

$$
M_0 \leq C^2\left(h^4 + \lambda\right)\|(\phi_\lambda)''\|_2^2.
$$

Then, since $\|(\phi_\lambda)''\|_2 \leq \|f''\|_2$ by Theorem 3.2,

$$
\|f - s_\lambda\|_2 \leq \|f - \phi_\lambda\|_2 + \|\phi_\lambda - s_\lambda\|_2 \\
\leq \lambda^{1/2}\|f''\|_2 + M_0^{1/2} \\
\leq \left(\lambda^{1/2} + C\sqrt{h^4 + \lambda}\right)\|f''\|_2,
$$

which yields (5.5). Since $\|(\phi_\lambda)''\|_2 \leq \|f''\|_2$ by Theorem 3.2, while

$$
\lambda\|(\phi_\lambda)'' - (s_\lambda)''\|_2^2 \leq M_0,
$$

it follows that

$$
\|(s_\lambda)''\|_2 \leq \|(\phi_\lambda)''\|_2 + \lambda^{-1/2}M_0^{1/2} \\
\leq \left(1 + C\left[h^4\lambda^{-1} + 1\right]^{1/2}\right)\|f''\|_2,
$$

which proves (5.6). \[\square\]
Corollary 5.3. Let $S \subset W_2^2[a,b]$ be a linear space. Suppose that for each $g \in W_4^2[a,b]$ there exists $u \in S$ such that

$$
\|g - u\|_2 \leq C h^4 \|g^{(4)}\|_2,
$$

$$
\|g'' - u''\|_2 \leq C h^2 \|g^{(4)}\|_2,
$$

for positive numbers $C$ and $h$ independent of $g$. Let $f \in W_2^4[a,b]$ and $f''(a) = f''(b) = f^{(3)}(a) = f^{(3)}(b) = 0$. Then the penalized least squares fit $s_\lambda$ of $f$ in $S$ corresponding to $\lambda$ satisfies

$$
\|f - s_\lambda\|_2 \leq \left(\lambda + C [\lambda + h^4]\right) \|f^{(4)}\|_2,
$$

(5.9)

$$
\|f'' - (s_\lambda)''\|_2 \leq \left(\lambda^{1/2} + Ch^2 h^{4\lambda^{-1} + 1} + 1\right)^{1/2} \|f^{(4)}\|_2,
$$

(5.10)

Proof. By Theorem 3.3, the Tikhonov regularization $\phi_\lambda$ of $f$ satisfies

$$
\|f - \phi_\lambda\|_2 \leq \lambda \|f^{(4)}\|_2.
$$

Let $M_0$ be the minimum in (5.1). Applying (5.8) for $g := \phi_\lambda$, it follows that

$$
M_0 \leq C^2 \left(h^8 + \lambda h^4\right) \|\phi_\lambda^{(4)}\|_2.
$$

Then, since $\|\phi_\lambda^{(4)}\|_2 \leq \|f^{(4)}\|_2$ by Theorem 3.3,

$$
\|f - s_\lambda\|_2 \leq \|f - \phi_\lambda\|_2 + \|\phi_\lambda - s_\lambda\|_2
\leq \lambda \|f^{(4)}\|_2 + M_0^{1/2}
\leq \left(\lambda + Ch^2 \sqrt{h^4 + \lambda}\right) \|f^{(4)}\|_2,
$$

which implies (5.9).

Since $\|f'' - (\phi_\lambda)''\|_2 \leq \lambda^{1/2} \|f^{(4)}\|_2$ by Theorem 3.3, using (5.7) gives

$$
\|f'' - (s_\lambda)''\|_2 \leq \lambda^{1/2} \|f^{(4)}\|_2 + \lambda^{-1/2} M_0^{1/2},
$$

and (5.10) follows. □

6. A general penalized least squares problem. Keeping in mind the problems and results of Sections 2–5, we now investigate the more general
Penalized Least Squares Fitting

setting described in the introduction. Let \( X, Y \) and \( S \) be linear spaces on \( \mathbb{R} \), where \( S \subseteq Y \subseteq X \); \( \| \cdot \|_X : X \to \mathbb{R} \) and \( \| \cdot \|_Y : Y \to \mathbb{R} \) are semi-norms induced by semi-definite inner products \( \langle \cdot, \cdot \rangle \) on \( X \) and \( [\cdot, \cdot] \) on \( Y \), respectively. For simplicity, we assume throughout the rest of this paper that \( S \) has dimension \( n \) and that

the restriction \( \| \cdot \|_X \) onto \( S \) is a norm on \( S \).

This implies that for any \( f \in X \) and \( \lambda > 0 \), there exists a unique \( s_\lambda := s_\lambda(f, S) := Q_\lambda f \) in \( S \) with the property

\[
\| f - s_\lambda \|_X^2 + \lambda \| s_\lambda \|_Y^2 = \inf_{u \in S} \left\{ \| f - u \|_X^2 + \lambda \| u \|_Y^2 \right\}.
\]

We denote the nonpenalized least squares fit in \( S \) by \( s_0 \), i.e.,

\[
\| f - s_0 \|_X^2 = \min_{u \in S} \| f - u \|_X^2.
\]

The penalized least squares approximation \( s_\lambda \) of \( f \) is characterized by the orthogonality relations

\[
\langle f - s_\lambda, u \rangle = \lambda [s_\lambda, u], \quad \text{for all } u \in S,
\]

while \( s_0 \) is characterized by

\[
\langle f - s_0, u \rangle = 0, \quad \text{for all } u \in S.
\]

By (6.1)–(6.2), we have

\[
\langle s_0 - s_\lambda, u \rangle = \lambda [s_\lambda, u], \quad \text{for all } u \in S,
\]

which proves that \( s_\lambda = Q_\lambda s_0 \), that is

\[
\| s_0 - s_\lambda \|_X^2 + \lambda \| s_\lambda \|_Y^2 = \min_{u \in S} \left\{ \| s_0 - u \|_X^2 + \lambda \| u \|_Y^2 \right\}.
\]

In what follows, we will discuss the error function \( s_0 - s_\lambda \) which indicates the consequences of the penalty, and is easier to analyse than the error function \( f - s_\lambda \) itself. This suffices since

\[
\| f - s_0 \|_X \leq \| f - s_\lambda \|_X \leq \| f - s_0 \|_X + \| s_0 - s_\lambda \|_X.
\]

In analogy with Theorem 3.2, we have
**Theorem 6.1.** For each $f \in X$ and $\lambda > 0$, $s_\lambda := s_\lambda(f,S)$ satisfies

$$
\|s_\lambda\|_Y \leq \|s_0\|_Y
$$

and

$$
\|s_0 - s_\lambda\|_X \leq \sqrt{\lambda} \|s_0\|_Y.
$$

**Proof.** Because of (6.3),

$$
M(s_0, \lambda) := \|s_0 - s_\lambda\|_X^2 + \lambda \|s_\lambda\|_Y^2 = \langle s_0 - s_\lambda, s_0 - s_\lambda \rangle + \lambda [s_\lambda, s_\lambda], \quad \text{for all } u \in S.
$$

Inserting $u = s_0$ yields

$$
M(s_0, \lambda) = \lambda [s_\lambda, s_0] \leq \lambda \|s_\lambda\|_Y \|s_0\|_Y,
$$

and inequalities (6.4) and (6.5) follow. □

Our next result provides an improvement of (6.5) for small $\lambda$ which implies that

$$
\|s_0 - s_\lambda\|_X = \mathcal{O}(\lambda) \quad \text{as } \lambda \to 0.
$$

**Theorem 6.2.** For each $f \in X$ and $\lambda > 0$, $s_\lambda := s_\lambda(f,S)$ satisfies

$$
\|s_0 - s_\lambda\|_X \leq K_S \lambda \|s_0\|_Y.
$$

where

$$
K_S := \sup \left\{ \frac{\|u\|_Y}{\|u\|_X} : u \in S, \; u \neq 0 \right\}.
$$

**Proof.** Coupling the orthogonality relation (6.3) for $u = s_0 - s_\lambda$ with the Cauchy-Schwarz inequality and (6.4), we have

$$
\|s_0 - s_\lambda\|_X^2 = \langle s_0 - s_\lambda, s_0 - s_\lambda \rangle = \lambda [s_\lambda, s_0 - s_\lambda] \\
\leq \lambda \|s_\lambda\|_Y \|s_0 - s_\lambda\|_Y \leq \lambda \|s_0\|_Y K_S \|s_0 - s_\lambda\|_X,
$$

which implies (6.6). □

We conclude this section by examining the special case where $X,Y,S$ are function spaces on some set $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$. Our next result gives a bound on $\|s_0 - s_\lambda\|_{L_\infty(\Omega)}$ which implies that

$$
\|s_0 - s_\lambda\|_{L_\infty(\Omega)} = \mathcal{O}(\lambda) \quad \text{as } \lambda \to 0.
$$
Theorem 6.3. Suppose $X \subseteq L_\infty(\Omega)$, and let

$$\kappa_S := \sup \left\{ \frac{\|u\|_{L_\infty(\Omega)}}{\|u\|_X} : u \in S, \ u \neq 0 \right\} < \infty.$$ 

Then

$$\|s_0 - s_\lambda\|_{L_\infty(\Omega)} \leq \kappa_S \sqrt{\lambda} \|s_0\|_Y \min \left\{ 1, K_S \sqrt{\lambda} \right\}.$$ 

Proof. The assertion follows directly from (6.5), Theorem 6.2, and the definition of $\kappa_S$. □

7. Applications.

7.1. Cubic splines on an interval. Let $X = C[a,b]$, $Y = W^2_2[a,b]$, and suppose $S := S_4(\Delta) \subseteq C^2[a,b]$ is the linear space of cubic splines with simple knots $\Delta : a = x_0 < x_1 < \cdots < x_n = b$. Suppose that $D := \{t_i\}_{i=1}^m$ are distinct points in $[a,b]$ so that the restriction of $S_4(\Delta)$ onto $D$ has full dimension $n + 3$. We can now apply the results of Section 6 for the inner products

$$\langle f, g \rangle = \frac{1}{m} \sum_{i=1}^m f(t_i)g(t_i), \quad f, g \in C[a,b]$$

and

$$[f, g] = \int_a^b f''g'', \quad f, g \in W^2_2[a,b].$$

In the notation of Sect. 6,

$$\|u\|_X = \left( \frac{1}{m} \sum_{i=1}^m u(t_i)^2 \right)^{1/2}, \quad u \in S.$$ 

We write $\| \cdot \|_2$ for the $L_2$-norm on $[a,b]$ and suppose that $D$ and $\Delta$ are such that

$$\delta := \sup \left\{ \frac{\|u\|_2}{\|u\|_X} : u \in S, \ u \neq 0 \right\}$$

is finite. Let $h_0 := h_0(\Delta) := \min\{x_{j+1} - x_j : j = 0, \ldots, n-1\}$. Since $\|u\|_2 \leq \delta \|u\|_X$, $u \in S$, it follows that for some absolute constant $C_0$,

$$\|u\|_Y = \|u''\|_2 \leq C_0 h_0^{-2} \|u\|_2 \leq C_0 \delta h_0^{-2} \|u\|_X,$$
and thus

\[ K_S \leq C_0 \delta h_0^{-2}. \]

Moreover, for some other absolute constant \( C_1 \),

\[ \kappa_S \leq C_1 \delta h_0^{-1/2}. \]

For given real numbers \( \{y_i\}_{i=1}^m \), let \( s_0 \in S \) be such that

\[
\frac{1}{m} \sum_{i=1}^m (y_i - s_0(t_i))^2 = \min_{u \in S} \frac{1}{m} \sum_{i=1}^m (y_i - u(t_i))^2.
\]

Recall that the penalized least squares spline \( s_\lambda \) minimizes the expression

\[
\frac{1}{m} \sum_{i=1}^m (y_i - u(t_i))^2 + \lambda \int_a^b (u''(t))^2 dt.
\]

The following result follows immediately from Theorems 6.1–6.3.

**Theorem 7.1.** For all \( \lambda > 0 \),

\[
\|s_0 - s_\lambda\|_2 \leq \delta \sqrt{\lambda} \min\{1, K_S \sqrt{\lambda}\} \|s''_0\|_2.
\]

Moreover,

\[
\|s_0 - s_\lambda\|_{L_\infty[a,b]} \leq C_1 \delta h_0^{-1/2} \sqrt{\lambda} \min\{1, C_0 \delta h_0^{-2} \sqrt{\lambda}\} \|s''_0\|_2.
\]

### 7.2. Bivariate spline spaces with stable local bases.

Let \( \Omega \) be a bounded set in \( \mathbb{R}^2 \) with polygonal boundary, and suppose that \( X \subseteq L_\infty(\Omega) \), \( Y = W_2^2(\Omega) \). Let \( S \subseteq X \) be a linear space of polynomial splines defined on a regular triangulation \( \triangle \) of \( \Omega \). For results on the approximation properties of these types of spaces, see \([1, 2, 4]\).

Suppose that \( D = \{t_i\}_{i=1}^m \) are distinct points in \( \Omega \) so that the restriction of \( S \) onto \( D \) has the same dimension as \( S \). We apply the results of Section 6 for the inner products

\[
\langle f, g \rangle = \frac{1}{m} \sum_{i=1}^m f(t_i)g(t_i), \quad f, g \in X,
\]
Penalized Least Squares Fitting

and

\[ [f, g] := \int_\Omega (f_{xx}g_{xx} + 2f_{xy}g_{xy} + f_{yy}g_{yy}) \, dx \, dy, \quad f, g \in Y. \]

For given real numbers \((z_i)_{i=1}^m, s_0 \in S\) is the unique function in \(S\) which satisfies

\[
\frac{1}{m} \sum_{i=1}^{m} (z_i - s_0(t_i))^2 = \min_{u \in S} \frac{1}{m} \sum_{i=1}^{m} (z_i - u(t_i))^2.
\]

In the notation of Sect. 6,

\[ \|u\|_X = \left( \frac{1}{m} \sum_{i=1}^{m} u(t_i)^2 \right)^{1/2}, \quad u \in S. \]

We have investigated properties of \(s_0 \in S\) in [3].

Let \(\| \cdot \|_2\) denote the \(L_2\)-norm on \(\Omega\), and suppose that \(D\) and \(\triangle\) are such that

\[ \delta := \sup \left\{ \|u\|_2 \|u\|_X : u \in S, \, u \neq 0 \right\} \]

is finite. Let \(h_0 := h_0(\triangle)\) be the minimum side length of the triangles in \(\triangle\). Since \(\|u\|_2 \leq \delta \|u\|_X, u \in S\), it follows that for some absolute constant \(C_0\)

\[ \|u\|_Y \leq C_0 h_0^{-2} \|u\|_2 \leq C_0 \delta h_0^{-2} \|u\|_X \]

so that

\[ K_S \leq C_0 \delta h_0^{-2}. \]

Moreover, for some other absolute constant \(C_1\),

\[ \kappa_S \leq C_1 \delta h_0^{-1}. \]

As for univariate splines, we can now apply Theorems 6.1–6.3 to get bounds on \(s_0 - s_\lambda\) in both the \(L_2\) and uniform norms on \(\Omega\).

REFERENCES


Manfred von Golitschek  
*Institut für Angewandte Mathematik und Statistik  
der Universität Würzburg, 97074 Würzburg, Germany  
e-mail: goli@mathematik.uni-wuerzburg.de*

Larry L. Schumaker  
*Department of Mathematics  
Vanderbilt University, Nashville  
TN 37240, USA  
e-mail: s@mars.cas.vanderbilt.edu*  
*Received September 19, 2002*