Optimal Function-on-Function Regression with Interaction between Functional Predictors

Honghe Jin, Xiaoxiao Sun and Pang Du

University of Georgia, University of Arizona and Virginia Polytechnic Institute and State University

Supplementary Material

In the online Supplementary Material, we introduce an example of the roughness penalty in S1 and the corresponding reproducing kernels in S2. S3 provides the proof of Theorem 1. The simulation results with random errors from heavy-tailed distributions are shown in S4. In S5, we list the gene names in the real-data example. A link to the R code of the proposed method can be found in S6.

S1 Penalty Functional

The roughness penalty functional $J(\beta)$ has the form

$$
\begin{aligned}
J(\beta) = & \int_0^1 \left([\int_0^1 \int_0^1 \frac{\partial^2\beta}{\partial t^2} drds]^2 + [\int_0^1 \int_0^1 \frac{\partial^3\beta}{\partial t^2 \partial s} drds]^2 + [\int_0^1 \int_0^1 \frac{\partial^3\beta}{\partial t^2 \partial r} drds]^2 \right. \\
& \left. + [\int_0^1 \int_0^1 \frac{\partial^4\beta}{\partial t^2 \partial s \partial r} drds]^2 \right) dt \\
& + \int_0^1 \left([\int_0^1 \int_0^1 \frac{\partial^2\beta}{\partial r^2} dtds]^2 + [\int_0^1 \int_0^1 \frac{\partial^3\beta}{\partial r^2 \partial s} dtds]^2 + [\int_0^1 \int_0^1 \frac{\partial^3\beta}{\partial r^2 \partial t} dtds]^2 \right. \\
& \left. + [\int_0^1 \int_0^1 \frac{\partial^4\beta}{\partial r^2 \partial s \partial t} dtds]^2 \right) dr \\
& + \int_0^1 \left([\int_0^1 \int_0^1 \frac{\partial^2\beta}{\partial s^2} drdt]^2 + [\int_0^1 \int_0^1 \frac{\partial^3\beta}{\partial s^2 \partial t} drdt]^2 + [\int_0^1 \int_0^1 \frac{\partial^3\beta}{\partial s^2 \partial r} drdt]^2 \right. \\
& \left. + [\int_0^1 \int_0^1 \frac{\partial^4\beta}{\partial s^2 \partial t \partial r} drdt]^2 \right) ds \\
& + \int_0^1 \int_0^1 \left([\int_0^1 \frac{\partial^4\beta}{\partial t^2 \partial s^2} dr]^2 + [\int_0^1 \frac{\partial^5\beta}{\partial t^2 \partial s^2 \partial r} dr]^2 \right) dsdt \\
& + \int_0^1 \int_0^1 \left([\int_0^1 \frac{\partial^4\beta}{\partial t^2 \partial r^2} ds]^2 + [\int_0^1 \frac{\partial^5\beta}{\partial t^2 \partial r^2 \partial s} ds]^2 \right) drdt \\
& + \int_0^1 \int_0^1 \left([\int_0^1 \frac{\partial^4\beta}{\partial t^2 \partial s^2} dt]^2 + [\int_0^1 \frac{\partial^5\beta}{\partial r^2 \partial s^2 \partial t} dt]^2 \right) dsdr \\
& + \int_0^1 \int_0^1 \int_0^1 [\frac{\partial^6\beta}{\partial r^2 \partial s^2 \partial t^2}]^2 drdsdt.
\end{aligned}
$$

Each component in the penalty represents the roughness of the trivariate function β at a different direction (as indicated by the variables included in the partial derivative) and a different level (as indicated by the order of the derivative). For example, in the first two lines of the definition, the first term represents the 2nd level of roughness of β at the direction of t, the second term represents the joint roughness of β of the 2nd level at the t direction and of the 1st level at the s direction, and so on. The last term on the last line of the definition represents the joint roughness of the trivariate function β of the 2nd level at all the three directions. Note that because the highest derivative with respect to any individual variable in J is of the 2nd order, the splines defined by J is multivariate cubic splines. In addition, the mechanism for obtaining these components actually roots from the construction of the reproducing kernel Hilbert spaces on the domain of each variable and their tensor product space. The details of such construction are referred to Wahba (1990) and Gu (2013).

S2 Reproducing Kernels and Inner Products

Let $k_i(x)$ be the ith Bernoulli polynomial, and denote $K_2(x_1, x_2) = k_2(x_1)k_2(x_2) - k_4(|x_1 - x_2|)$. Note that the null space $\mathcal{H}_0^*$ can be further decomposed as $\mathcal{H}_0^* = \mathcal{H}_{00}^* \oplus \mathcal{H}_{01}^*$, with the corresponding reproducing kernels $K_0(\cdot, \cdot) = 1 + k_1(\cdot)k_1(\cdot) = K_{00}(\cdot, \cdot) + K_{01}(\cdot, \cdot)$, where $*$ runs through the index set $\{x, y, z\}$. Then, the reproducing kernel Hilbert space tensor product decomposition is listed in Table 1.

Subspace	Kernel	Inner Product
$\mathcal{H}_{00x} \otimes \mathcal{H}_{00z} \otimes \mathcal{H}_{00y}$	1	$(\int\int\int f\,dxdydz)(\int\int\int g\,dxdydz)$
$\mathcal{H}_{01x} \otimes \mathcal{H}_{00z} \otimes \mathcal{H}_{00y}$	$k_1(x_1)k_1(x_2)$	$(\int\int\int \frac{\partial f}{\partial x}\,dxdydz)(\int\int\int \frac{\partial g}{\partial x}\,dxdydz)$
$\mathcal{H}_{00x} \otimes \mathcal{H}_{01z} \otimes \mathcal{H}_{00y}$	$k_1(z_1)k_1(z_2)$	$(\int\int\int \frac{\partial f}{\partial z}\,dxdydz)(\int\int\int \frac{\partial g}{\partial z}\,dxdydz)$
$\mathcal{H}_{00x} \otimes \mathcal{H}_{00z} \otimes \mathcal{H}_{01y}$	$k_1(y_1)k_1(y_2)$	$(\int\int\int \frac{\partial f}{\partial y}\,dxdydz)(\int\int\int \frac{\partial g}{\partial y}\,dxdydz)$
$\mathcal{H}_{01x} \otimes \mathcal{H}_{01z} \otimes \mathcal{H}_{00y}$	$k_1(x_1)k_1(x_2)k_1(z_1)k_1(z_2)$	$(\int\int\int \frac{\partial^2 f}{\partial x\partial z}\,dxdydz)(\int\int\int \frac{\partial^2 g}{\partial x\partial z}\,dxdydz)$
$\mathcal{H}_{01x} \otimes \mathcal{H}_{00z} \otimes \mathcal{H}_{01y}$	$k_1(x_1)k_1(x_2)k_1(y_1)k_1(y_2)$	$(\int\int\int \frac{\partial^2 f}{\partial x\partial y}\,dxdydz)(\int\int\int \frac{\partial^2 g}{\partial x\partial y}\,dxdydz)$
$\mathcal{H}_{00x} \otimes \mathcal{H}_{01z} \otimes \mathcal{H}_{01y}$	$k_1(z_1)k_1(z_2)k_1(y_1)k_1(y_2)$	$(\int\int\int \frac{\partial^2 f}{\partial z\partial y}\,dxdydz)(\int\int\int \frac{\partial^2 g}{\partial z\partial y}\,dxdydz)$
$\mathcal{H}_{01x} \otimes \mathcal{H}_{01z} \otimes \mathcal{H}_{01y}$	$k_1(x_1)k_1(x_2)k_1(z_1)k_1(z_2)k_1(y_1)k_1(y_2)$	$(\int\int\int \frac{\partial^3 f}{\partial x\partial z\partial y}\,dxdydz)(\int\int\int \frac{\partial^3 g}{\partial x\partial z\partial y}\,dxdydz)$
$\mathcal{H}_{1x} \otimes \mathcal{H}_{00z} \otimes \mathcal{H}_{00y}$	$K_2(x_1,x_2)$	$\int(\int\int \frac{\partial^2 f}{\partial x^2}dzdy)(\int\int \frac{\partial^2 g}{\partial x^2}dzdy)\,dx$
$\mathcal{H}_{00x} \otimes \mathcal{H}_{1z} \otimes \mathcal{H}_{00y}$	$K_2(z_1,z_2)$	$\int(\int\int \frac{\partial^2 f}{\partial z^2}dxdy)(\int\int \frac{\partial^2 g}{\partial z^2}dxdy)\,dz$
$\mathcal{H}_{00x} \otimes \mathcal{H}_{00z} \otimes \mathcal{H}_{1y}$	$K_2(y_1,y_2)$	$\int(\int\int \frac{\partial^2 f}{\partial y^2}dxdz)(\int\int \frac{\partial^2 g}{\partial y^2}dxdz)\,dy$
$\mathcal{H}_{1x} \otimes \mathcal{H}_{01z} \otimes \mathcal{H}_{00y}$	$K_2(x_1,x_2)k_1(z_1)k_1(z_2)$	$\int(\int\int \frac{\partial^3 f}{\partial x^2\partial z}dzdy)(\int\int \frac{\partial^3 g}{\partial x^2\partial z}dzdy)\,dx$
$\mathcal{H}_{1x} \otimes \mathcal{H}_{00z} \otimes \mathcal{H}_{01y}$	$K_2(x_1,x_2)k_1(y_1)k_1(y_2)$	$\int(\int\int \frac{\partial^3 f}{\partial x^2\partial y}dzdy)(\int\int \frac{\partial^3 g}{\partial x^2\partial y}dzdy)\,dx$
$\mathcal{H}_{01x} \otimes \mathcal{H}_{1z} \otimes \mathcal{H}_{00y}$	$K_2(z_1,z_2)k_1(x_1)k_1(x_2)$	$\int(\int\int \frac{\partial^3 f}{\partial z^2\partial x}dxdy)(\int\int \frac{\partial^3 g}{\partial z^2\partial x}dxdy)\,dz$
$\mathcal{H}_{00x} \otimes \mathcal{H}_{1z} \otimes \mathcal{H}_{01y}$	$K_2(z_1,z_2)k_1(y_1)k_1(y_2)$	$\int(\int\int \frac{\partial^3 f}{\partial z^2\partial y}dxdy)(\int\int \frac{\partial^3 g}{\partial z^2\partial y}dxdy)\,dz$
$\mathcal{H}_{01x} \otimes \mathcal{H}_{00z} \otimes \mathcal{H}_{1y}$	$K_2(y_1,y_2)k_1(x_1)k_1(x_2)$	$\int(\int\int \frac{\partial^3 f}{\partial y^2\partial x}dxdz)(\int\int \frac{\partial^3 g}{\partial y^2\partial x}dxdz)\,dy$
$\mathcal{H}_{00x} \otimes \mathcal{H}_{01z} \otimes \mathcal{H}_{1y}$	$K_2(y_1,y_2)k_1(z_1)k_1(z_2)$	$\int(\int\int \frac{\partial^3 f}{\partial y^2\partial z}dxdz)(\int\int \frac{\partial^3 g}{\partial y^2\partial z}dxdz)\,dy$
$\mathcal{H}_{1x} \otimes \mathcal{H}_{01z} \otimes \mathcal{H}_{01y}$	$K_2(x_1,x_2)k_1(z_1)k_1(z_2)k_1(y_1)k_1(y_2)$	$\int(\int\int \frac{\partial^4 f}{\partial x^2\partial z\partial y}dzdy)(\int\int \frac{\partial^4 g}{\partial x^2\partial z\partial y}dzdy)\,dx$
$\mathcal{H}_{01x} \otimes \mathcal{H}_{1z} \otimes \mathcal{H}_{01y}$	$K_2(z_1,z_2)k_1(x_1)k_1(x_2)k_1(y_1)k_1(y_2)$	$\int(\int\int \frac{\partial^4 f}{\partial z^2\partial x\partial y}dzdy)(\int\int \frac{\partial^4 g}{\partial z^2\partial x\partial y}dzdy)\,dx$
$\mathcal{H}_{01x} \otimes \mathcal{H}_{01z} \otimes \mathcal{H}_{1y}$	$K_2(y_1,y_2)k_1(x_1)k_1(x_2)k_1(z_1)k_1(z_2)$	$\int(\int\int \frac{\partial^4 f}{\partial y^2\partial x\partial z}dzdy)(\int\int \frac{\partial^4 g}{\partial y^2\partial x\partial z}dzdy)\,dx$
$\mathcal{H}_{1x} \otimes \mathcal{H}_{1z} \otimes \mathcal{H}_{00y}$	$K_2(x_1,x_2)K_2(z_1,z_2)$	$\int\int(\int \frac{\partial^4 f}{\partial x^2\partial z^2}dy)(\int \frac{\partial^4 g}{\partial x^2\partial z^2}dy)\,dxdz$
$\mathcal{H}_{1x} \otimes \mathcal{H}_{00z} \otimes \mathcal{H}_{1y}$	$K_2(x_1,x_2)K_2(y_1,y_2)$	$\int\int(\int \frac{\partial^4 f}{\partial x^2\partial y^2}dz)(\int \frac{\partial^4 g}{\partial x^2\partial y^2}dz)\,dxdy$
$\mathcal{H}_{00x} \otimes \mathcal{H}_{1z} \otimes \mathcal{H}_{1y}$	$K_2(z_1,z_2)K_2(y_1,y_2)$	$\int\int(\int \frac{\partial^4 f}{\partial z^2\partial y^2}dx)(\int \frac{\partial^4 g}{\partial z^2\partial y^2}dx)\,dzdy$
$\mathcal{H}_{1x} \otimes \mathcal{H}_{1z} \otimes \mathcal{H}_{01y}$	$K_2(x_1,x_2)K_2(z_1,z_2)k_1(y_1)k_1(y_2)$	$\int\int(\int \frac{\partial^5 f}{\partial x^2\partial z^2\partial y}dy)(\int \frac{\partial^5 g}{\partial x^2\partial z^2\partial y}dy)\,dxdz$
$\mathcal{H}_{1x} \otimes \mathcal{H}_{01z} \otimes \mathcal{H}_{1y}$	$K_2(x_1,x_2)K_2(y_1,y_2)k_1(z_1)k_1(z_2)$	$\int\int(\int \frac{\partial^5 f}{\partial x^2\partial y^2\partial z}dz)(\int \frac{\partial^5 g}{\partial x^2\partial y^2\partial z}dz)\,dxdy$
$\mathcal{H}_{01x} \otimes \mathcal{H}_{1z} \otimes \mathcal{H}_{1y}$	$K_2(z_1,z_2)K_2(y_1,y_2)k_1(x_1)k_1(x_2)$	$\int\int(\int \frac{\partial^5 f}{\partial z^2\partial y^2\partial x}dx)(\int \frac{\partial^5 g}{\partial z^2\partial y^2\partial x}dx)\,dzdy$
$\mathcal{H}_{1x} \otimes \mathcal{H}_{1z} \otimes \mathcal{H}_{1y}$	$K_2(x_1,x_2)K_2(z_1,z_2)K_2(y_1,y_2)$	$\int\int\int \frac{\partial^6 f}{\partial x^2\partial y^2\partial z^2}\frac{\partial^6 g}{\partial x^2\partial y^2\partial z^2}dxdzdy$

Table 1: Reproducing kernels and inner products

S3 Proof of Theorem 1

Proof. For brevity, we shall write $C_n = \frac{1}{n}\sum_{i=1}^n U_i(r_1)V_i(s_1)U_i(r_2)V_i(s_2)$ and $g_\epsilon = \frac{1}{n}\sum_{i=1}^n \epsilon_i(t)U_i(r)V_i(s)$. Given the assumption that $\mathcal{H}$ is dense in $\mathcal{L}_2$, there exist uniquely defined $f_0, \hat{f} \in \mathcal{L}_2$ such that $\beta_0 = L_{K^{1/2}} f_0$ and $\hat{\beta}_{n\lambda} = L_{K^{1/2}}\hat{f}_\lambda$. The optimal solution of (4) can be written as $\hat{f}_\lambda = (\lambda I + T_n)^{-1}(T_n f_0 + g_n)$, where $T_n = L_{K^{1/2}} L_{C_n} L_{K^{1/2}}$, $g_n = L_{K^{1/2}} g_\epsilon$, and I is the identity operator. The excess risk can be decomposed into

$$\left\|T^{1/2}(\hat{f}_\lambda - f_0)\right\|_{\mathcal{L}_2} = \left\|T^{1/2}(f_\lambda - f_0)\right\|_{\mathcal{L}_2} + \left\|T^{1/2}(\hat{f}_\lambda - f_\lambda)\right\|_{\mathcal{L}_2}, \qquad \text{(S3.1)}$$

where $f_\lambda = (\lambda I + T)^{-1} T f_0$. The proof can be completed in two steps. We now bound the first term on the right-hand side of (S3.1) in the first step. Write $f_0 = \sum_{k=1}^\infty a_k \zeta_k$, where ζ_k are eigenfunctions of T. Then, $f_\lambda = \sum_{k=1}^\infty \frac{\rho_k a_k}{\lambda + \rho_k}\zeta_k$. Thus,

$$f_\lambda - f_0 = -\sum_{k=1}^\infty \frac{\lambda a_k}{\lambda + \rho_k}\zeta_k,$$

and

$$\left\|T^{1/2}(f_\lambda - f_0)\right\|_{\mathcal{L}_2}^2 = \sum_{k=1}^\infty \left(\frac{\lambda a_k}{\lambda + \rho_k}\right)^2 \rho_k \le \max_{k \ge 1} \frac{\lambda^2 \rho_k}{(\lambda + \rho_k)^2} \sum_{k=1}^\infty a_k^2 \le \frac{\lambda ||f_0||_{\mathcal{L}_2}^2}{4}.$$

We then appeal to bound the second term on the right-hand side of (S3.1) in the second step. It is easy to show that

$$f_\lambda - \hat{f}_\lambda = (\lambda I + T)^{-1}(\lambda I + T_n)(f_\lambda - \hat{f}_\lambda) + (\lambda I + T)^{-1}(T - T_n)(f_\lambda - \hat{f}_\lambda).$$

Note that

$$(\lambda I + T_n)\hat{f}_\lambda = T_n f_0 - g_n.$$

Therefore,

$$\begin{aligned}
f_\lambda - \hat{f}_\lambda &= (\lambda I + T)^{-1} T_n (f_\lambda - f_0) + \lambda(\lambda I + T)^{-1} f_\lambda + (\lambda I + T)^{-1} g_n \\
&+ (\lambda I + T)^{-1}(T - T_n)(f_\lambda - \hat{f}_\lambda) \\
&= (\lambda I + T)^{-1} T_n (f_\lambda - f_0) + \lambda T f_0 + (\lambda I + T)^{-1} g_n \\
&+ (\lambda I + T)^{-1}(T - T_n)(f_\lambda - \hat{f}_\lambda) \\
&= (\lambda I + T)^{-1} T (f_\lambda - f_0) + \lambda T f_0 + (\lambda I + T)^{-1} g_n \\
&+ (\lambda I + T)^{-1}(T - T_n)(f_\lambda - f_0) + (\lambda I + T)^{-1}(T - T_n)(f_\lambda - \hat{f}_\lambda)
\end{aligned}$$

By triangular inequality,

$$\begin{aligned}
\left\| T^{1/2}(f_\lambda - \hat{f}_\lambda) \right\|_{\mathcal{L}_2} \leq & \left\| T^{1/2}(\lambda I + T)^{-1} T (f_\lambda - f_0) \right\|_{\mathcal{L}_2} \\
& + \left\| T^{1/2}(\lambda I + T)^{-1}(T - T_n)(f_\lambda - f_0) \right\|_{\mathcal{L}_2} \\
& \lambda \left\| T^{3/2} f_0 \right\|_{\mathcal{L}_2} + \left\| T^{1/2}(\lambda I + T)^{-1} g_n \right\|_{\mathcal{L}_2} \\
& \left\| T^{1/2}(\lambda I + T)^{-1}(T - T_n)(f_\lambda - \hat{f}_\lambda) \right\|_{\mathcal{L}_2}.
\end{aligned} \tag{S3.2}$$

Next, we will bound the five terms on the right-hand side of (S3.2).

Following the similar discussion in the proof of Theorem 2 in Sun et al. (2018), we have

$$\left\| T^{1/2}(\lambda I + T)^{-1} T (f_\lambda - f_0) \right\|_{\mathcal{L}_2} \leq \frac{1}{2} \lambda^{1/2} \left\| f_0 \right\|_{\mathcal{L}_2},$$

and

$$\lambda \left\| T^{3/2} f_0 \right\|_{\mathcal{L}_2} = O(\lambda).$$

We apply Lemma 1 to obtain

$$\left\| T^{1/2}(\lambda I + T)^{-1} g_n \right\|_{\mathcal{L}_2} = O((n\lambda^{1/2\omega})^{-1/2}).$$

Based on the usual operator inequality, we have

$$\left\| T^{1/2}(\lambda I + T)^{-1}(T - T_n)(f_\lambda - f_0) \right\|_{\mathcal{L}_2} \leq \left\| T^{1/2}(\lambda I + T)^{-1}(T - T_n)T^{-\nu} \right\|_{op} \left\| T^{\nu}(f_\lambda - f_0) \right\|_{\mathcal{L}_2}.$$

Using Assumption (C2) and following the proof in Theorem 2 in Sun et al. (2018), we obtain

$$\begin{aligned} \left\| T^{1/2}(\lambda I + T)^{-1}(T - T_n)T^{-\nu} \right\|_{op} \left\| T^{\nu}(f_\lambda - f_0) \right\|_{\mathcal{L}_2} \leq & O((n\lambda^{1/2\omega})^{-1/2}) \left\| T^{\nu}(f_\lambda - f_0) \right\|_{\mathcal{L}_2} \\ \leq & O((n\lambda^{1/2\omega})^{-1/2}\lambda^{\nu}), \end{aligned}$$

where $\nu > 0$ such that $2\omega(1 - 2\nu) > 1$.

The last term can be bounded by the following inequality

$$\begin{aligned} \left\| T^{1/2}(\lambda I + T)^{-1}(T - T_n)(f_\lambda - \hat{f}_\lambda) \right\|_{\mathcal{L}_2} \leq & \left\| T^{1/2}(\lambda I + T)^{-1}(T - T_n)T^{-\nu} \right\|_{op} \left\| T^{\nu}(f_\lambda - \hat{f}_\lambda) \right\|_{\mathcal{L}_2} \\ \leq & O((n\lambda^{1/2\omega})^{-1/2}) \left\| T^{\nu}(f_\lambda - \hat{f}_\lambda) \right\|_{\mathcal{L}_2}. \end{aligned}$$

We combine this inequality with (S3.2) and use the bounds developed previously to conclude that

$$\left\| T^{1/2}(\lambda I + T)^{-1}(T - T_n)(f_\lambda - \hat{f}_\lambda) \right\|_{\mathcal{L}_2} = o((n\lambda^{1/2\omega})^{-1/2}),$$

provided that $c_1 n^{-2\omega/(2\omega+1)} \leq \lambda \leq c_2 n^{-2\omega/(2\omega+1)}$ for some constants $0 < c_1 < c_2 < \infty$. We now conclude that

$$\left\| T^{1/2}(f_\lambda - \hat{f}_\lambda) \right\|_{\mathcal{L}_2} = O(n^{-2\omega/(2\omega+1)}).$$

Lemma 1. If Assumptions (C1), (C3), (C4) and (C5) hold,

$$\left\| T^{1/2}(\lambda I + T)^{-1} g_n \right\|_{\mathcal{L}_2} = O((n\lambda^{1/2\omega})^{-1/2}).$$

Proof of Lemma 1

We have that,

$$\begin{aligned}
\left\| T^{1/2}(\lambda I + T)^{-1} g_n \right\|_{\mathcal{L}_2}^2 &= \sum_{k=1}^{\infty} \langle g_n, (T + \lambda I)^{-1} T^{1/2} \zeta_k \rangle_{\mathcal{L}_2}^2 \\
&= \sum_{k=1}^{\infty} \langle g_n, \frac{\rho_k^{1/2}}{\lambda + \rho_k} \zeta_k \rangle_{\mathcal{L}_2}^2 \\
&= \sum_{k=1}^{\infty} \frac{\rho_k}{(\lambda + \rho_k)^2} \langle g_n, \zeta_k \rangle_{\mathcal{L}_2}^2 .
\end{aligned}$$

Recall that $U_i(\cdot) = X_i(\cdot) + \tau_i(\cdot)$ and $V_i(\cdot) = Z_i(\cdot) + \upsilon_i(\cdot)$. Then we have that

$$\begin{aligned}
g_n =& \frac{1}{n} \sum_{i=1}^{n} L_{K^{1/2}} \epsilon_i(t) U_i(r) V_i(s) \\
=& \frac{1}{n} \sum_{i=1}^{n} L_{K^{1/2}} \epsilon_i(t) (X_i(r) Z_i(s) + X_i(r) \upsilon_i(s) + \tau_i(r) Z_i(s) + \tau_i(r) \upsilon_i(s)).
\end{aligned}$$

Thus,

$$
\begin{aligned}
\left\|T^{1/2}(\lambda I+T)^{-1}g_n\right\|_{\mathcal{L}_2}^2 \leq \sum_{k=1}^{\infty}\frac{4\rho_k}{(\lambda+\rho_k)^2}\Big(&\langle\frac{1}{n}\sum_{i=1}^{n}\epsilon_i(t)X_i(r)Z_i(s),L_{K^{1/2}}\zeta_k\rangle_{\mathcal{L}_2}^2\\
&+\langle\frac{1}{n}\sum_{i=1}^{n}\epsilon_i(t)X_i(r)\upsilon_i(s),L_{K^{1/2}}\zeta_k\rangle_{\mathcal{L}_2}^2\\
&+\langle\frac{1}{n}\sum_{i=1}^{n}\epsilon_i(t)\tau_i(r)Z_i(s),L_{K^{1/2}}\zeta_k\rangle_{\mathcal{L}_2}^2\\
&+\langle\frac{1}{n}\sum_{i=1}^{n}\epsilon_i(t)\tau_i(r)\upsilon_i(s),L_{K^{1/2}}\zeta_k\rangle_{\mathcal{L}_2}^2\Big)
\end{aligned}
$$

Taking expectation and using Assumptions (C4) and (C5) and the assumed independence between ϵ and surrogate variables, we obtain

$$
\mathbb{E}\left\|T^{1/2}(\lambda I+T)^{-1}g_n\right\|_{\mathcal{L}_2}^2 \leq \frac{1}{n}\sum_{k=1}^{\infty}\frac{16M\rho_k^2}{(\lambda+\rho_k)^2}.
$$

Applying Lemma 5 in Reimherr et al. (2018), we obtain

$$
\mathbb{E}\left\|T^{1/2}(\lambda I+T)^{-1}g_n\right\|_{\mathcal{L}_2}^2 \leq \frac{c}{n\lambda^{\frac{1}{2\omega}}},
$$

where c is a constant. The proof is completed by applying Markov inequality.

□

S4 Simulation Results with Heavy-tailed Errors

We presented the simulation results with heavy-tailed random errors in Table 2. The simulation settings are the same as those in Section 5.1

except for the random errors. Let $\tau(t)$ be an i.i.d. random function such that $\tau(t)$ follows the t-distribution with 5 degrees of freedom for each fixed time point. The random error $\epsilon(t)$ was generated by $\epsilon(t) = k\tau(t)$, where k was chosen to yield the desired signal-to-noise ratios (SNR).

	SNR	*QFFR*	*pffr*	*FRegSig*
Scenario 1	1	2.717(1.682)	**1.200**(1.115)	5.001(11.85)
	5	1.298(0.722)	**0.401**(0.215)	0.790(0.998)
	10	0.585(0.400)	**0.190**(0.145)	0.503(0.616)
Scenario 2	1	**0.208**(0.160)	0.735(1.056)	1.622(1.843)
	5	**0.441**(0.317)	421.2(244.1)	3.397(2.99)
	10	**0.240**(0.421)	365.7(282.9)	1.045(1.710)
Scenario 3	1	**0.420**(0.971)	6.321(3.982)	2.200(2.966)
	5	**0.0501**(0.0727)	0.782(2.003)	0.591(0.903)
	10	**0.0237**(0.0555)	0.410(0.399)	0.382(0.175)

Table 2: The averages (standard deviations) of the MISEs ($\times 10^{-3}$) under the three scenarios with heavy-tailed errors. The best result is shown in boldface.

S5 List of Selected Genes

The Ensembl gene IDs of the selected 59 genes for our study are: ENSG00000183615 ENSG00000228217 ENSG00000186364 ENSG00000213047 ENSG00000271399

ENSG00000153187 ENSG00000237512 ENSG00000240527 ENSG00000107581

ENSG00000186104 ENSG00000270510 ENSG00000228607 ENSG00000255225

ENSG00000205885 ENSG00000271142 ENSG00000120833 ENSG00000089250

ENSG00000258289 ENSG00000188869 ENSG00000277864 ENSG00000259704

ENSG00000260395 ENSG00000225442 ENSG00000263477 ENSG00000237854

ENSG00000118271 ENSG00000127445 ENSG00000130734 ENSG00000267291

ENSG00000267879 ENSG00000252436 ENSG00000231062 ENSG00000228848

ENSG00000277901 ENSG00000201021 ENSG00000234271 ENSG00000283440

ENSG00000232360 ENSG00000253874 ENSG00000185686 ENSG00000273899

ENSG00000231711 ENSG00000226567 ENSG00000226621 ENSG00000199594

ENSG00000145103 ENSG00000240048 ENSG00000070193 ENSG00000250955

ENSG00000164308 ENSG00000254246 ENSG00000229282 ENSG00000112799

ENSG00000216636 ENSG00000219699 ENSG00000219532 ENSG00000201794

ENSG00000235379 ENSG00000200713

S6 R Code

The R code is available on GitHub (`https://github.com/honghe1994/QFFR.git`).

Bibliography

Gu, C. (2013). *Smoothing spline ANOVA models*, Volume 297. Springer Science & Business Media.

Reimherr, M., B. Sriperumbudur, B. Taoufik, et al. (2018). Optimal prediction for additive function-on-function regression. *Electronic Journal of Statistics 12*(2), 4571–4601.

Sun, X., P. Du, X. Wang, and P. Ma (2018). Optimal penalized function-on-function regression under a reproducing kernel Hilbert space framework. *Journal of the American Statistical Association*, 1–11.

Wahba, G. (1990). *Spline models for observational data*, Volume 59. Philadelphia: SIAM.