Functional linear regression with functional response

David Benatia Marine Carrasco
University of Montreal
Jean-Pierre Florens
Toulouse School of Economics

Montreal

Workshop on Inverse Problems - October 2016

Illustrative example

- $y_i(t)$ hourly electricity consumption for Ontario for day i
- $z_i(t)$ hourly average temperature for day i

$$\begin{aligned}\begin{bmatrix} y_i(1) \\ \vdots \\ y_i(24) \end{bmatrix} &= \underset{24\times 72}{A} \begin{bmatrix} z_i(-24) \\ \vdots \\ z_i(48) \end{bmatrix} + \begin{bmatrix} u_i(1) \\ \vdots \\ u_i(24) \end{bmatrix}, \\ Y_i &= AZ_i + U_i.\end{aligned}$$

Data from January 1, 2010 to September 30, 2014.

- If the frequency increases, the correlation between elements increases, yielding a singular $Z'Z$.
- Even at fixed frequency, inverting $Z'Z$ is problematic.

Model

$$Y_i(s) = \int \Pi(s,t) Z_i(t) \, dt + U_i(s), \; i = 1, 2, ..., n$$

where Y and Z are functions, elements of Hilbert spaces.

- Object of interest: estimation of Π.
- s could be the time or a vector, for instance (longitude, latitude).
- $Y_i(s)$ depends not only on $Z_i(s)$ but also on $Z_i(t)$, $t \neq s$.

Example of application: technological spillover

Our model could be used to study the effect of technological spillover on productivity, as in Manresa (2015).

$$Y_i(s) = \log \text{ of output of firms with industry code } s,$$

$$Z_i(s) = \text{ log of R\&D expenditures of firms with industry code } s.$$

Here i could be the time.
Industry code goes from 1 to 6 digits $\Rightarrow$ thinner and thinner grid.
Manresa uses Lasso to identify a network of firms.

Related literature

- Linear functional regression:

$$Y_i = \varphi(Z_i) + U_i,\ i = 1, 2, ..., n$$

  where the dependent variable Y_i is a scalar and the object of interest is the estimation of the function φ.
  See Cardo, Ferraty, and Sarda (2003), Hall and Horowitz (2007), Horowitz and Lee (2007), Darolles, Fan, Florens and Renault (2011).
- Models where both Y and Z are functions.
  See Bosq (2000), Ramsay and Silverman (2005), Cuevas, Febrero, and Fraiman (2002), Yao, Müller, and Wang (2005), Kargin and Onatski (2008), Hovath and Kokoszka (2012), Park and Qian (2012), Crambes and Mas (2013), Aue, Norinho, and Hörmann (2015). Approaches based on functional principal components.

The model is

$$Y = \Pi Z + U \tag{1}$$

where $Y \in \mathcal{E}$, $Z \in \mathcal{F}$, U is a zero mean random element of $\mathcal{E}$ and Π is a nonrandom Hilbert-Schmidt operator from $\mathcal{F}$ to $\mathcal{E}$. Moreover, Z is exogenous so that $cov\left(Z, U\right) = 0$.

For example

$$\mathcal{E} = \left\{ g : \int_{\mathcal{S}} g\left(t\right)^2 dt < \infty \right\}, \ \mathcal{F} = \left\{ f : \int_{\mathcal{T}} f\left(t\right)^2 dt < \infty \right\}$$

where $\mathcal{S}$ and $\mathcal{T}$ are some intervals of $\mathbb{R}$. Then, Π can be represented as an integral operator such that

$$\left(\Pi\varphi\right)\left(s\right) = \int_{\mathcal{T}} \pi\left(s, t\right) \varphi\left(t\right) dt$$

for any $\varphi \in \mathcal{F}$. π is referred to as the kernel of the operator Π.

Estimator

$$\begin{aligned} V_Z &: \mathcal{F} \to \mathcal{F} \\ \varphi &\to V_Z\varphi = E\left[Z\left\langle Z, \varphi \right\rangle\right] \end{aligned}$$

$$\begin{aligned} C_{YZ} &: \mathcal{F} \to \mathcal{E} \\ \varphi &\to C_{YZ}\varphi = E\left[Y\left\langle Z, \varphi \right\rangle\right] \\ &= E\left[\left(\Pi Z + U\right)\left\langle Z, \varphi \right\rangle\right] \\ &= \Pi E\left[Z\left\langle Z, \varphi \right\rangle\right] + E\left[U\left\langle Z, \varphi \right\rangle\right]. \end{aligned}$$

It follows from the exogeneity of Z :

$$C_{YZ} = \Pi V_Z.$$

Let Π^* be the adjoint of Π, then

$$C_{ZY} = V_Z \Pi^*.$$

The solution is **unique** if V_Z is injective.
The unknown operators V_Z and C_{ZY} are replaced by their sample counterparts

$$\begin{aligned} \hat{V}_Z \varphi &= \frac{1}{n} \sum_{i=1}^{n} z_i \langle z_i, \varphi \rangle, \\ \hat{C}_{ZY} \psi &= \frac{1}{n} \sum_{i=1}^{n} z_i \langle y_i, \psi \rangle. \end{aligned}$$

An estimator of Π^* cannot be obtained directly from

$$\hat{C}_{ZY} = \widehat{V}_Z \Pi^*$$

because the initial equation $C_{ZY} = V_Z \Pi^*$ is an ill-posed problem.

We adopt Tikhonov regularization (see Kress, 1999 and Carrasco, Florens, and Renault, 2007).
The estimator of Π^* is defined as

$$\hat{\Pi}_\alpha^* = \left(\alpha I + \hat{V}_Z\right)^{-1} \hat{C}_{ZY} \tag{2}$$

and that of Π is defined by

$$\hat{\Pi}_\alpha = \hat{C}_{YZ} \left(\alpha I + \hat{V}_Z\right)^{-1} \tag{3}$$

where α is some positive regularization parameter which will be allowed to converge to zero as n goes to infinity.
The estimator $\hat{\Pi}_\alpha$ is also a penalized least-squares estimator:

$$\begin{aligned} \hat{\Pi}_\alpha &= \arg\min_{\Pi} \left\| y - \Pi z \right\|^2 + \alpha \left\| \Pi \right\|_{HS}^2 \\ &= \arg\min_{\Pi} \sum_{i=1}^{n} \left\| y_i - \Pi z_i \right\|^2 + \alpha \sum \tilde{\mu}_j^2 \end{aligned}$$

where $\tilde{\mu}_j$ are the singular values of the operator Π.

Implementation

$$
\begin{aligned}
\left(\alpha I+\hat{V}_Z\right)^{-1}\hat{C}_{ZY}\psi &= \hat{\Pi}_\alpha^*\psi \Leftrightarrow \\
\hat{C}_{ZY}\psi &= \left(\alpha I+\hat{V}_Z\right)\hat{\Pi}_\alpha^*\psi \Leftrightarrow \\
\frac{1}{n}\sum_{i=1}^n z_i\left\langle y_i,\psi\right\rangle &= \alpha\hat{\Pi}_\alpha^*\psi+\frac{1}{n}\sum_{i=1}^n z_i\left\langle z_i,\hat{\Pi}_\alpha^*\psi\right\rangle .
\end{aligned}
$$

Then, we take the inner product with z_l, $l=1,2,...,n$ on the left and right hand sides:

$$
\frac{1}{n}\sum_{i=1}^n \left\langle z_l,z_i\right\rangle\left\langle y_i,\psi\right\rangle=\alpha\left\langle z_l,\hat{\Pi}_\alpha^*\psi\right\rangle+\frac{1}{n}\sum_{i=1}^n\left\langle z_l,z_i\right\rangle\left\langle z_i,\hat{\Pi}_\alpha^*\psi\right\rangle ,
$$

l equations with n unknowns $\left\langle z_i,\hat{\Pi}_\alpha^*\psi\right\rangle$, $i=1,2,...,n$.

Let M be the $n \times n$ matrix with (l, i) element $\langle z_l, z_i \rangle / n$, v the $n-$vector of $\left\langle z_i, \hat{\Pi}_\alpha^* \psi \right\rangle$ and w the $n-$vector of $\langle y_i, \psi \rangle$. (**??**) is equivalent to

$$Mw = (\alpha I + M) v.$$

And $v = (\alpha I + M)^{-1} Mw = M (\alpha I + M)^{-1} w$. For a given ψ, we can compute:

$$\begin{aligned}
\hat{\Pi}_\alpha^* \psi &= \frac{1}{\alpha n} \sum_{i=1}^{n} z_i \left(\langle y_i, \psi \rangle - \left\langle z_i, \hat{\Pi}_\alpha^* \psi \right\rangle \right) \\
&= \frac{1}{\alpha n} \underline{z}' \left(I - M (\alpha I + M)^{-1} \right) w \\
&= \frac{1}{n} \underline{z}' (\alpha I + M)^{-1} w
\end{aligned}$$

where $\underline{z}$ is the $n-$vector of z_i.

Mean square error (MSE)

Assumption 1. (U_i, Y_i, Z_i) iid with $E(U_i) = 0$, $cov(U_i, Z_i) = 0$, $cov(U_i, U_j | Z_1, Z_2, ..., Z_n) = 0$ for all $i \neq j$ and $= V_U$ for $i = j$ where V_U is a trace-class operator.

Assumption 2. Π belongs to $\mathcal{H}(\mathcal{F}, \mathcal{E})$ the space of Hilbert-Schmidt operators.

Assumption 3. V_Z is a trace-class operator and $\left\| \hat{V}_Z - V_Z \right\|_{HS}^2 = O_p(1/n)$.

Assumption 4. There is a Hilbert-Schmidt operator R from $\mathcal{E}$ to $\mathcal{F}$ and a constant $\beta > 0$ such that $\Pi^* = V_Z^{\beta/2} R$.

An operator K is trace-class if $\sum_j \left\langle K\phi_j, \phi_j \right\rangle < \infty$ for any basis $\left(\phi_j\right)$. An operator K is Hilbert-Schmidt (noted HS) if $\|K\|_{HS}^2 \equiv \sum_j \left\langle K\phi_j, K\phi_j \right\rangle < \infty$ for any basis $\left(\phi_j\right)$.

The MSE is defined by

$$E\left(\left\|\hat{\Pi}_\alpha - \Pi\right\|_{HS}^2 |Z_1, .., Z_n\right).$$

Replacing y_i by $\Pi z_i + u_i$ in the expression of $\hat{C}_{ZY}$, we obtain

$$\hat{C}_{ZY} = \frac{1}{n}\sum_i z_i \langle y_i, .\rangle = \hat{C}_{ZU} + \hat{V}_Z \Pi^*.$$

We decompose $\hat{\Pi}_\alpha^* - \Pi^*$ in the following manner:

$$\begin{aligned}
\hat{\Pi}_\alpha^* - \Pi^* &= \left(\alpha I + \hat{V}_Z\right)^{-1} \hat{C}_{ZY} - \Pi^* \\
&= \left(\alpha I + \hat{V}_Z\right)^{-1} \hat{C}_{ZU} && (4)\\
&\quad + \left(\alpha I + \hat{V}_Z\right)^{-1} \hat{V}_Z \Pi^* - \left(\alpha I + V_Z\right)^{-1} V_Z \Pi^* && (5)\\
&\quad + \left(\alpha I + V_Z\right)^{-1} V_Z \Pi^* - \Pi^*. && (6)
\end{aligned}$$

To study the rate of convergence of the MSE, we will study the rates of the three terms (4), (5), and (6).

Proposition

Assume Assumptions 1 to 4 hold.
If $\beta > 1$, *then* $MSE = O_p\left(\frac{1}{n\alpha} + \alpha^{\beta \wedge 2}\right)$.
If $\beta < 1$, *then* $MSE = O_p\left(\frac{\alpha^{\beta}}{n\alpha^2} + \alpha^{\beta}\right)$.

Remark: Our method allows for multiple eigenvalues.

Tests

We want to test the null hypothesis:

$$H_0 : \Pi = \Pi_0$$

where Π_0 is known. A simple way to test this hypothesis is to look at $\hat{C}_{ZY} - \hat{V}_Z \Pi_0^*$.
Under H_0, this operator equals $\hat{C}_{ZU}$ and should be close to zero.
Moreover, under H_0,

$$\sqrt{n} \left(\hat{C}_{ZY} - \hat{V}_Z \Pi_0^* \right) \xrightarrow{d} \mathcal{N} \left(0, K_{ZU} \right)$$

where

$$K_{ZU} = E \left[\left(u \otimes Z \right) \widetilde{\otimes} \left(u \otimes Z \right) \right]$$

and $\left(x \otimes y \right) \left(f \right) = \langle x, f \rangle \, y$ and $\left(\Pi_1 \widetilde{\otimes} \Pi_2 \right) T = \langle T, \Pi_1^* \rangle_{\mathcal{H}} \Pi_2$ (see Dauxois, Pousse, and Romain, 1982).

Let $\left\{\phi_j, \psi_j : j = 1, 2, ..., q\right\}$ be a set of test functions, then

$$\begin{bmatrix} \sqrt{n}\left\langle\left(\hat{C}_{ZY} - \hat{V}_Z \Pi_0^*\right)\phi_1, \psi_1\right\rangle \\ \vdots \\ \sqrt{n}\left\langle\left(\hat{C}_{ZY} - \hat{V}_Z \Pi_0^*\right)\phi_q, \psi_q\right\rangle \end{bmatrix}$$

converges to a multivariate normal distribution with mean 0_q and covariance matrix the $q \times q$ matrix Σ with (j, l) element:

$$\begin{aligned} \Sigma_{jl} &= E\left[\left\langle\sqrt{n}\hat{C}_{ZU}\phi_j, \psi_j\right\rangle\left\langle\sqrt{n}\hat{C}_{ZU}\phi_l, \psi_l\right\rangle\right] \\ &= \left\langle\phi_j, V_Z\psi_l\right\rangle\left\langle\phi_j, V_U\psi_l\right\rangle. \end{aligned}$$

The appropriately rescaled quadratic form converges to a chi-square distribution with q degrees of freedom which can be used to test H_0. The test functions could be normal densities with same small variance but centered at different means.

Data-driven selection of the regularization parameter

α can be selected by leave-one-out cross-validation

$$\min_{\alpha} \frac{1}{n} \sum_{j} \left\| y_i - \hat{\Pi}_{\alpha}^{(-i)} z_i \right\|^2$$

where $\hat{\Pi}_{\alpha}^{(-i)}$ has been computed using all observations except for the ith one.

- Centorrino (2015) studies this criterion for nonparametric IV regression and shows that it is rate optimal in mean squared error.
- An alternative approach would be to use a penalized minimum contrast criterion as in Goldenshluger and Lepski (2011). This could lead to a minimax-optimal estimator (Comte and Johannes, 2012).

Discrete observations

Suppose that the data (y_i, z_i) are observed at discrete times. We use some smoothing to construct pairs of curves (y_i^m, z_i^m), $i = 1, 2, ..., n$ such that $y_i^m \in \mathcal{E}$ and $z_i^m \in \mathcal{F}$. This smoothing can be obtained by approximating the curves by step functions or kernel smoothing for instance. The subscript m corresponds to the smallest number of discrete observations across $i = 1, 2, ..., n$. m grows with the sample size n.

Using the smoothed observations, we compute the corresponding estimators of V_Z and C_{ZY} denoted $\hat{V}_Z^m$, $\hat{C}_{ZY}^m$ and the estimator of Π^* denoted $\hat{\Pi}_\alpha^{m*}$:

$$\hat{\Pi}_\alpha^{m*} = \left(\alpha I + \hat{V}_Z^m\right)^{-1} \hat{C}_{ZY}^m.$$

We assume that the discretization error is negligible with respect to the estimation error:

Assumption 6. $\|z_i^m - z_i\| = O_p(f(m))$ and $\|y_i^m - y_i\| = O_p(f(m))$.

Assumption 7.

$$\frac{f(m)}{\alpha n} = o\left(\alpha^{\beta \wedge 2}\right).$$

Proposition

Under Assumptions 1 to 4, 6, and 7, the MSE of $\hat{\Pi}_\alpha^{m*} - \Pi^*$ *has the same rate of convergence as that of the MSE of* $\hat{\Pi}_\alpha^* - \Pi^*$ *in Proposition 1.*

Case where the regressor is endogenous

Potential application: Evaluating the price elasticity of the demand for electricity. Use the wind speed as instrument.

$$
\begin{aligned}
Y &= \text{demand for electricity,} \\
Z &= \text{price of electricity,} \\
W &= \text{wind speed.}
\end{aligned}
$$

Assume Z is endogenous but we observe instrumental variables W such that $\text{cov}(U, W) = 0$. Hence, $E\left(\left(Y - \Pi Z\right)\langle W, .\rangle\right) = 0$. It follows that

$$C_{YW} = \Pi C_{ZW} \tag{7}$$

where $C_{YW} = E\left(Y\langle W, .\rangle\right)$ and $C_{ZW} = E\left(Z\langle W, .\rangle\right)$. Similarly, we have

$$C_{WY} = C_{WZ}\Pi^* \tag{8}$$

where $C_{WZ} = E\left(W\langle Z, .\rangle\right)$

Π is **identified** provided C_{WZ} is injective.

To construct an estimator of Π^*, we first apply the operator C_{ZW} on the l.h.s and r.h.s of Equation (8) to obtain

$$C_{ZW} C_{WY} = C_{ZW} C_{WZ} \Pi^*.$$

Note that $C_{ZW} = C_{WZ}^*$ and therefore the operator $C_{ZW} C_{WZ}$ is self-adjoint. The operators C_{ZW}, C_{WZ}, and C_{WY} can be estimated by their sample counterparts. The estimator of Π^* is defined by

$$\hat{\Pi}_\alpha^* = \left(\alpha I + \hat{C}_{ZW} \hat{C}_{WZ}\right)^{-1} \hat{C}_{ZW} \hat{C}_{WY}. \qquad (9)$$

Similarly, the estimator of Π is given by

$$\hat{\Pi}_\alpha = \hat{C}_{YW} \hat{C}_{WZ} \left(\alpha I + \hat{C}_{ZW} \hat{C}_{WZ}\right)^{-1}.$$

Simulations

Let $\mathcal{E} = \mathcal{F} = L^2[0,1]$ and $\mathcal{S} = \mathcal{T} = [0,1]$. Π is an integral operator from to $L^2[0,1]$ to $L^2[0,1]$ with kernel

$$\pi(s,t) = 1 - |s-t|^2.$$

$$dU(s) = -U(s)ds + \sigma_u dG_u(s), \text{ for } s \in [0,1]$$

where G_u is a Brownian motion. Note that this error function is stationary.

We study the model

$$Y_i = \Pi Z_i + U_i, \qquad i = 1, ..., n$$

with exogenous regressors

$$Z_i(t) = \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i) + \Gamma(\beta_i)} t^{\alpha_i - 1}(1-t)^{\beta_i - 1} + \eta_i$$

for $t \in [0,1]$, with $\alpha_i, \beta_i \sim iidU[2,5]$ and $\eta_i \sim iidN(0,1)$, for all $i = 1, ..., n$.

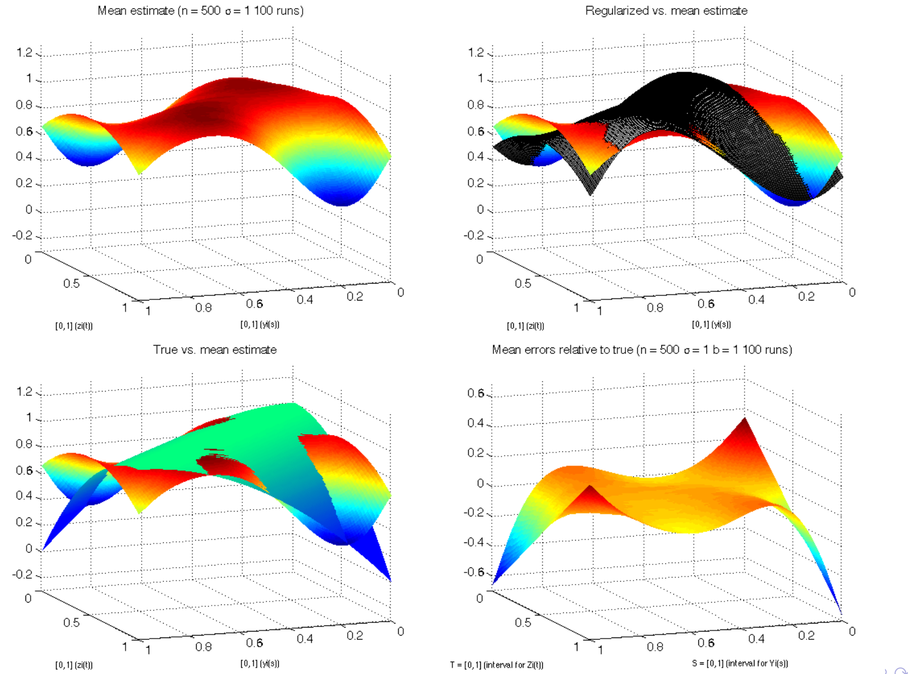
Simulation design

1. Construct both a pseudo-continuous interval of $[0,1]$, $\mathcal{T}$, consisting of 1000 equally-spaced discrete steps, and a discretized interval of $[0,1]$, $\tilde{\mathcal{T}}$, consisting of only 100.
2. Generate n functions $z_i(t)$, $u_i(s)$, where $t, s \in \mathcal{T}$ so as to obtain pseudo-continuous functions.
3. Generate the n response functions $y_i(s)$, $s \in \mathcal{T}$.
4. Generate the sample of n discretized pairs of functions $(\tilde{z}_i, \tilde{y}_i)$ by extracting the corresponding values of the pairs (z_i, y_i) for all $t, s \in \tilde{\mathcal{T}}$.
5. Estimate Π using the regularization method on the sample of n pairs $(\tilde{z}_i, \tilde{y}_i)$ and a fixed smoothing parameter $\alpha = .01$.
6. Repeat steps 2-5 100 times and calculate the *MSE* by averaging the quantities $\|\hat{\Pi}_\alpha - \Pi\|^2_{HS} = \int\limits_{\tilde{\mathcal{T}}}\int\limits_{\tilde{\mathcal{T}}} (\hat{\pi}_\alpha(s,t) - \pi(s,t))^2 dtds$ over all repetitions.

Table: Simulation results: Mean-Square Errors over 100 replications

Std errors	Empirical MSE				Squared bias
	$n = 50$	$n = 100$	$n = 500$	$n = 1000$	$\|\Pi - \Pi_\alpha\|^2_{HS}$
$\sigma_u = 0.1$	.0154	.0135	.0126	.0124	.0095
	(.0027)	(.0017)	(.0008)	(.0005)	
$\sigma_u = 0.25$	.0291	.0205	.0138	.0130	.0095
	(.0098)	(.0063)	(.0022)	(.0013)	
$\sigma_u = 0.5$	.0773	.0438	.0194	.0156	.0095
	(.0363)	(.0193)	(.0057)	(.0028)	
$\sigma_u = 1$	.2909	.1354	.0371	.0257	.0095
	(.1789)	(.0659)	(.0161)	(.0089)	
$\sigma_u = 2$	.9128	.4755	.1245	.0668	.0095
	(.5495)	(.2607)	(.0660)	(.0378)	

Note: Standard errors are reported in parentheses.


Mean estimate (n = 500 σ = 1 100 runs)
[0, 1] (zi(t))
[0, 1] (yi(s))
Regularized vs. mean estimate
[0, 1] (zi(t))
[0, 1] (yi(s))
True vs. mean estimate
[0, 1] (zi(t))
[0, 1] (yi(s))
Mean errors relative to true (n = 500 σ = 1 b = 1 100 runs)
T = [0, 1] (interval for Zi(t))
S = [0, 1] (interval for Yi(s))

Application

- $Y_i(t)$ = total electricity consumption for Ontario at time t for day i. It includes residential $(1/3)$, commercial $(1/3)$, and industrial $(1/3)$ consumptions.
- $Z_i(t)$ = average temperature at time t for day i.
- Data: hourly observations from January 1, 2010 to September 30, 2014.

The temperature is constructed in three steps.

- First, we match a set of 41 Ontarian cities (of above 10,000 inhabitants) to their three nearest weather stations.
- Second, we compute a weighted average using a distance metric.
- Finally, we obtain the province-wide temperature $Z(t)$ as:

$$Z(t) = \sum_{c} \gamma_c \{ \sum_{w(c)} \rho_{w(c)} Z_{w(c)}(t) \}, \qquad \forall i, h$$

where $\gamma_c = \frac{Pop_c}{(\sum_j Pop_j)}$ is city c's weight,

$\rho_{w(c)} = \frac{((lat_c - lat_{w(c)})^6 + (lon_c - lon_{w(c)})^6)^{-1}}{\sum_{l(c)} ((lat_c - lat_{l(c)})^6 + (lon_c - lon_{l(c)})^6)^{-1}}$ is station w's weight for city c's temperature average and $Z_w(t)$ is temperature measurement at station w in hour t.

- Finally, we use robust locally weighted polynomial regression on the constructed temperature series in order to smooth implausible jumps, which maybe due to measurement errors.

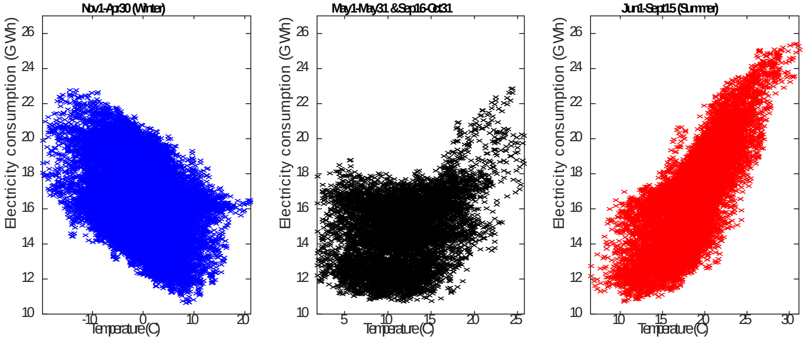


Figure: Electricity Consumption (GWh) vs. Temperature (celsius)

- We focus on summer period: June 1st to September 15.
- Holidays and weekends are dropped.
- Dummy variables for years, months, weekdays and hours of the day.
- We transform the temperature variables into cooling degrees defined by:

$$Z_c(t) = Z(t) - \min_{Z \in \text{Summer}} (Z(t)), \tag{10}$$

Model

Let $\mathcal{E} = L^2[0, 24]$ and $\mathcal{F} = L^2[-24, 48]$ with $\mathcal{S} = [0, 24]$ and $\mathcal{T} = [-24, 48]$.

$$Y_i(s) = \pi_0(s) + \int_{[-24,48]} \pi_1(s,t) Z_i(t) dt + \sum_{j \in J} \beta_j d_{ij} + U_i(s),$$

where $Y_i(s)$ is total electricity consumption in hour t, $\pi_0(s)$ is a constant function, $Z_i(t)$ is temperature in hour $t \in [-24, 48]$, and $U_i(s)$ is a zero-mean error term.
The object of interest is the kernel π_1, which characterizes the dynamic relation between electricity consumption for AC needs and temperature patterns.

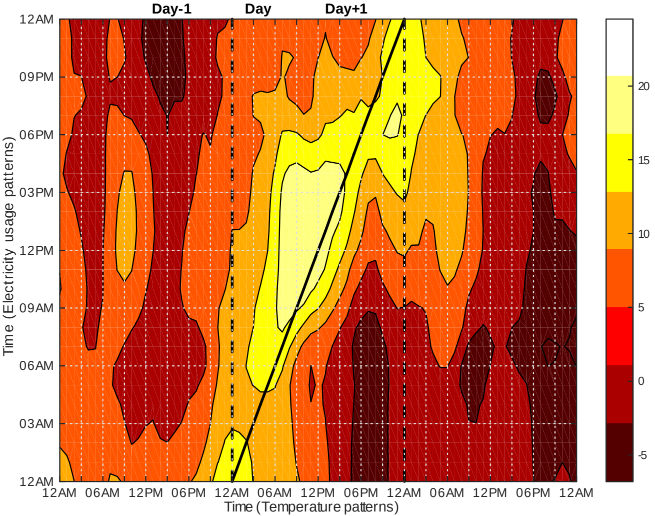
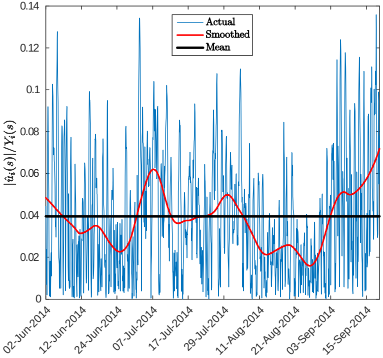
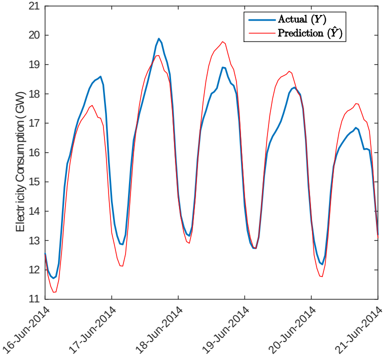


Figure: Contour plots for summer months with $\alpha = 0.7$


Actual
Smoothed
Mean
$|\hat{u}_i(s)|/Y_i(s)$
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
02-Jun-2014
12-Jun-2014
24-Jun-2014
07-Jul-2014
17-Jul-2014
29-Jul-2014
11-Aug-2014
21-Aug-2014
03-Sep-2014
15-Sep-2014


Actual (Y)
Prediction ($\hat{Y}$)
Electricity Consumption (GW)
21
20
19
18
17
16
15
14
13
12
11
16-Jun-2014
17-Jun-2014
18-Jun-2014
19-Jun-2014
20-Jun-2014
21-Jun-2014

Conclusion

- We propose a new estimator based on Tikhonov regularization.
- The model is specified for general spaces and can be applied to various data, including spatial data.
- We consider the case with exogenous regressor and the case with endogenous regressor.
- We derive the rate of convergence of the MSE and study the asymptotic normality.
- So far we focused on iid error, the extension to weakly dependent variables would be interesting.