Functional regression for spatio-temporal data

Y. Matsuda and X. Yuan
Tohoku University

overview

1. Introduction to Spectral analysis of time series
2. Extension to spatial data
3. Applications to functional regression for spatial data
    - Functional regression models for spatial data
    - Estimation on the frequency domain
    - Asymptotic properties
    - Applications to NTT Docomo human mobility survey

Spectral analysis of stationary time series

Time domain

$$\gamma_k = Cov(X_t, X_{t-k}), k = 0,1,\cdots$$

$$X_t, t = 1,\cdots,T$$

$$\log L(\theta) = \log\|\Sigma(\theta)\| + X'\Sigma(\theta)^{-1}X$$

Frequency domain

$$f(\omega) = (2\pi)^{-1}\sum_{k=-\infty}^{\infty}\gamma_k e^{ik\omega}, -\pi < \omega \leq \pi$$

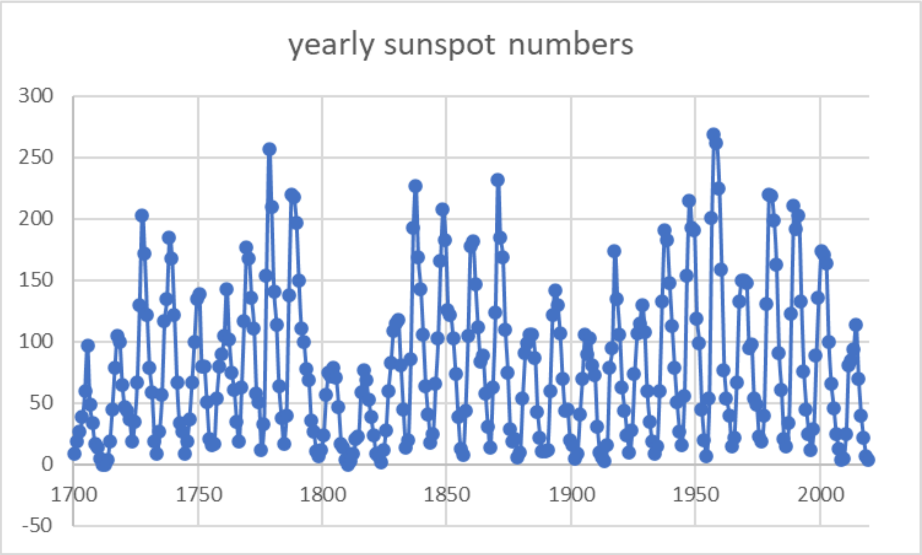
$$\tilde{X}(\omega_k) = \frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T} X_t e^{-i\omega_k t}, \omega_k = \frac{2\pi k}{T}$$

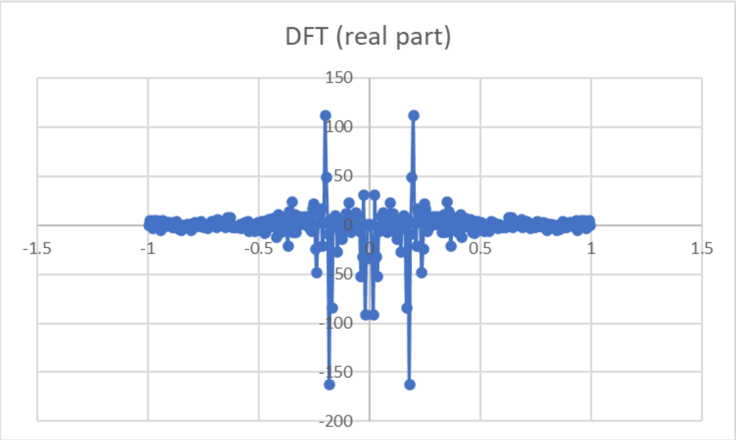
$$\log L(\theta) = \sum_k \log f(\omega_k;\theta) + \frac{\left|\tilde{X}(\omega_k)\right|^2}{f(\omega_k;\theta)}$$

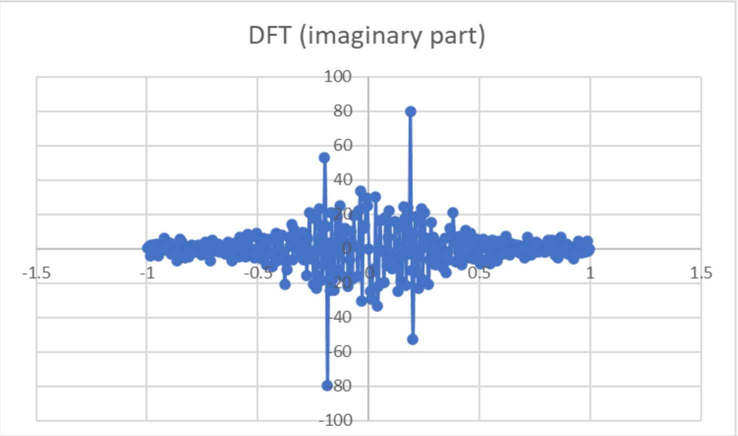
Example: sunspot numbers

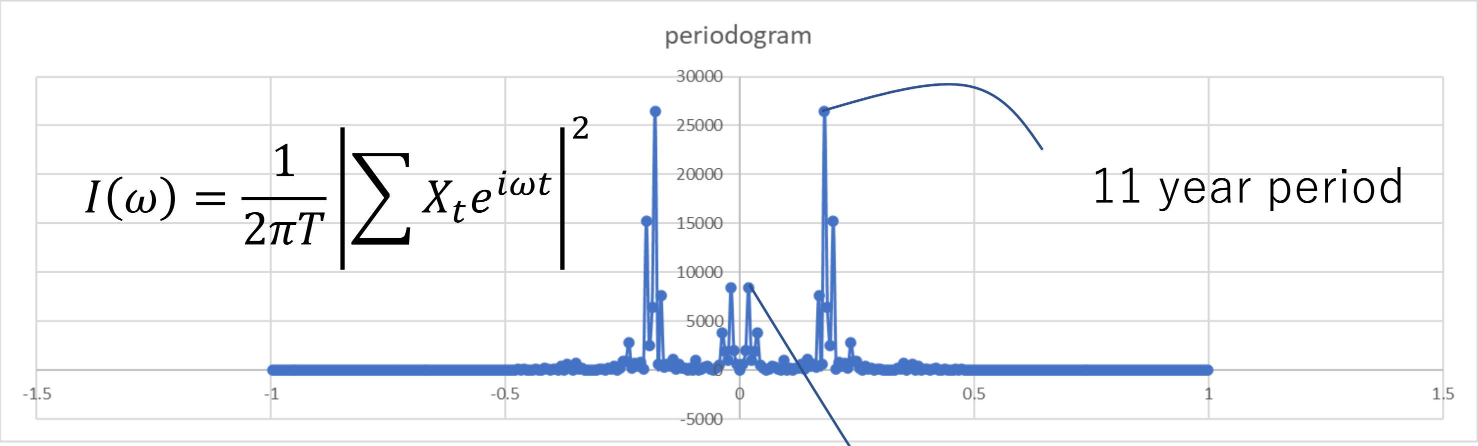
$$\frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T} X_t e^{i\omega t} = \frac{1}{\sqrt{2\pi T}}\sum X_t \cos(\omega t) + \frac{i}{\sqrt{2\pi T}}\sum X_t \sin(\omega t)$$

X_t







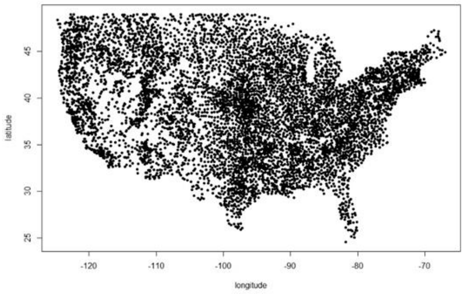


$$I(\omega) = \frac{1}{2\pi T}\left|\sum X_t e^{i\omega t}\right|^2$$

11 year period

107 year period

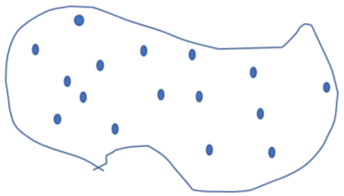
Extension to spatial point data



Spatial domain

$$\gamma(u) = Cov\big(X(s), X(s-u)\big), u \in R^2$$

$$X(s), s \in D$$



$$X(s_j), s_j \in D$$

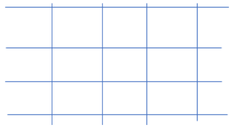
$$\log L(\theta) = \log\|\Sigma(\theta)\| + X'\Sigma(\theta)^{-1}X$$

Frequency domain

$$f(\omega) = (2\pi)^{-2}\int_{R^2} \gamma(u)\, e^{iu'\omega} du, \omega \in R^2$$

$$\tilde{X}(\omega) = \int_D X(s)\, e^{-i\omega' s} ds$$

$$\tilde{X}_N(\omega_k) = \frac{1}{N}\sum_{j=1}^{N} X(s_j) e^{-i\omega_k' s_j}$$



$$\log L(\theta) = \sum_k \log f(\omega_k;\theta) + \frac{\left|\tilde{X}_N(\omega_k)\right|^2}{f(\omega_k;\theta)}$$

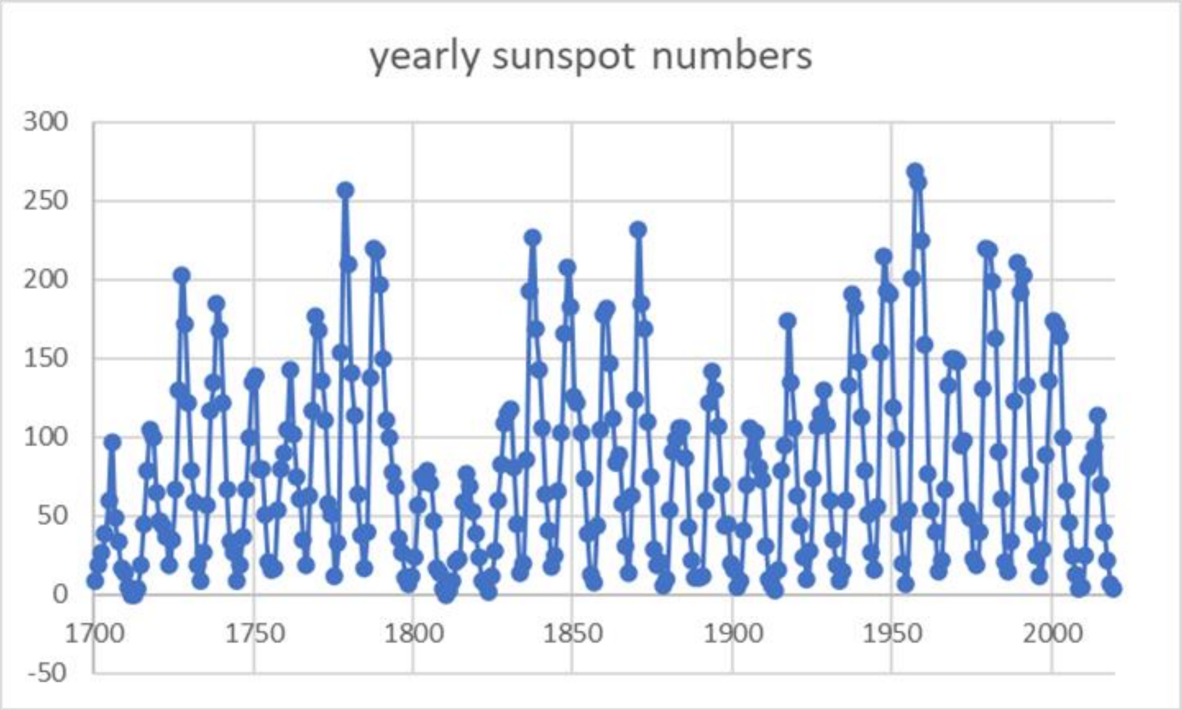
Regression models

lwage	educ	exper	tenure
6.645091	12	11	2
6.694562	18	11	16
6.715384	14	11	9
6.476973	12	13	7
6.331502	11	14	5
7.244227	16	14	2
6.39693	10	13	0
6.985642	18	8	14
7.05099	15	13	1
6.907755	12	16	16
6.835185	18	8	13
6.82546	14	9	11
6.802395	15	4	3
7.183871	16	7	2
7.491087	16	9	9
6.864848	10	17	2
7.21524	15	6	9
6.745236	11	19	10
6.721426	14	4	7
6.154858	12	13	7
7.150702	14	9	1

$$\widehat{\log(wage)} = \underset{(0.11)}{5.49} + \underset{(0.0065)}{0.075}educ + \underset{(0.0034)}{0.015}exper + \underset{(0.0026)}{0.013}tenure$$

$$n = 935, R^2 = 0.155$$

Autoregression



$$X_t = 24.3 + 1.38X_{t-1} - 0.69X_{t-2} + u_t$$
$$(2.41) \quad (0.041) \quad (0.041)$$
$$n = 319, R^2 = 0.828$$

Functional regression

For $y, x \in L^2(S)$, functional data,

$$y_t = \beta_1(x_{t1}) + \cdots + \beta_q(x_{tp}) + \varepsilon_t,$$

where $\beta_j(\cdot), \varphi_j(\cdot)$ are bounded linear operators

Examples

$$S = [a, b] \subset R, \text{curve time series}$$

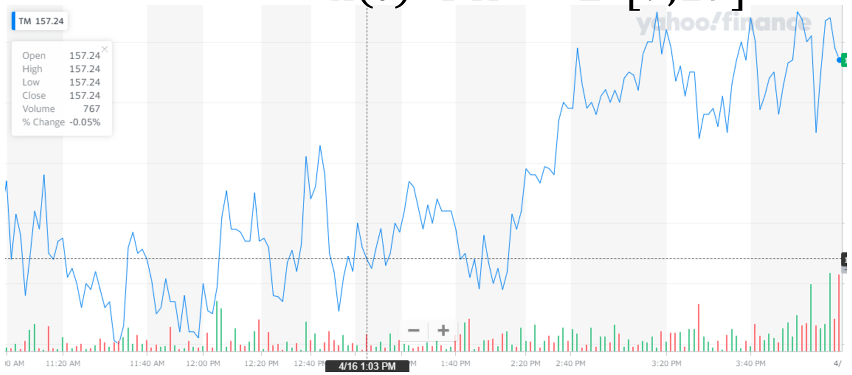
$$S = D \subset R^2, \text{ spatio-temporal data}$$

Stock price of TSLA and TOYOTA

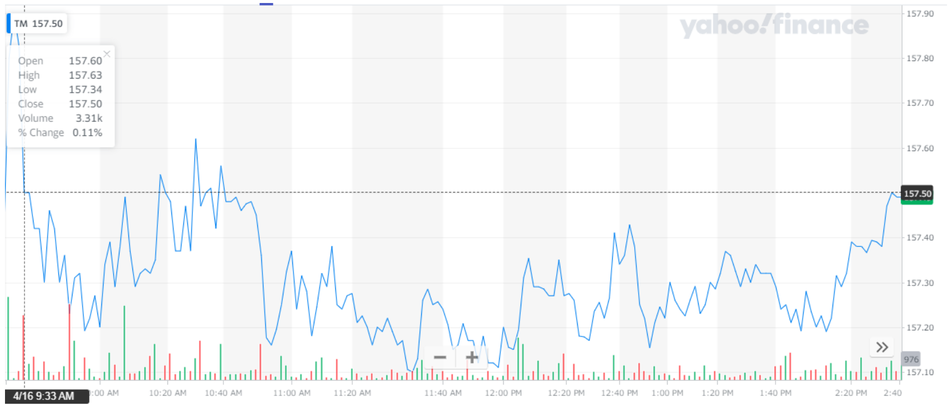
$$y(t) = \beta(x(t)) + u(t)$$

TSLA $y(t) \in H := L^2[9,17]$ TOYOTA $x(t) \in H := L^2[9,17]$

Apr. 17th

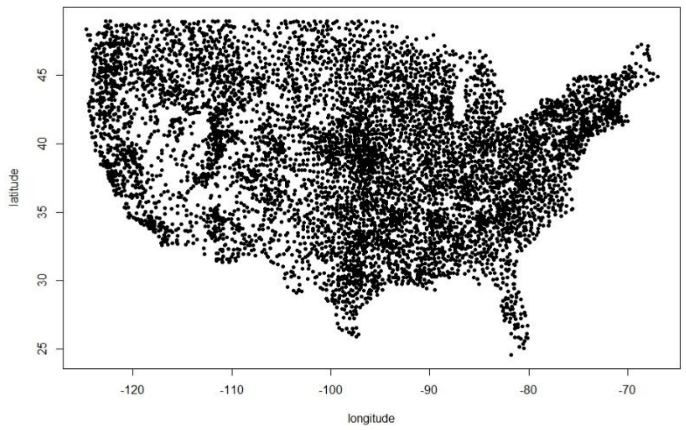


Apr. 18th



US ppt data

Func. regression: $y(s) = \beta_1(x_1(s)) + \beta_2(x_2(s)) + u(s),$



US precipitation data

$y(s) \in L^2(D)$

lon	lat	Jan./1986	Feb	March	April
-85.25	31.57	8.9	11.4	10.2	2.5
-87.18	34.22	2.6	6.4	8.1	1.1
-87.32	34.42	NA	NA	NA	NA
-87.42	32.23	4.5	8.1	9.5	0.9
-86.22	34.25	NA	NA	NA	NA
-85.95	32.95	2.6	7.5	15.9	0.9
-85.87	32.98	NA	NA	NA	NA
-88.13	33.13	2.6	7.3	9.5	NA
-88.28	33.23	4.2	12.2	13.6	1.5
-86.5	31.32	16.3	9.3	19.3	1.8
-85.85	33.58	2.4	5.1	6.2	0.4
-87.22	34.07	NA	NA	NA	NA
-85.83	33.27	1.2	8.3	NA	1.5
-86.27	33.83	NA	NA	NA	NA
-86.98	34.8	NA	NA	NA	NA
-87.48	31.17	18.2	9.8	12.7	1.7
-87.52	31.02	NA	NA	NA	NA
-85.5	32.6	NA	NA	NA	NA
-86.68	32.47	NA	8.8	18.3	1.1
-87.35	33.45	4.9	6.5	6.2	2.1
-87.78	30.88	NA	NA	NA	NA

Temperature (max)

$x_1(s) \in L^2(D)$

Jan	Feb	Mar	Apr
NA	NA	NA	NA
14.4	18.7	21.2	26.7
NA	NA	NA	NA
13	16.9	20.8	NA
NA	NA	21.8	27
12.6	16.5	20.3	24.9
13.6	17.9	NA	25.9
13	14.3	17.8	24.4

Temperature (min)

$x_2(s) \in L^2(D)$

Jan	Feb	Mar	Apr
NA	NA	NA	NA
-2.8	2.4	4.1	8.2
NA	NA	NA	NA
-2.9	3.5	3.6	NA
-0.3	4.1	4.8	7
-1.2	3.3	5.1	7.8
-3.9	1.2	NA	6.2
-2.4	0.9	2.6	6.4

NTT Docomo Mobile Statistics

モバイル空間統計®

Covid-19 cases in Japan

Weekly new cases of covid-19

y

B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
ity	pop	pid	lat	lon	2020/1/1	2020/1/8	2020/1/15	2020/1/22	2020/1/29	2020/2/5	2020/2/12	2020/2/19	2020/2/26	2020/3/4	2020/3/11	2020/3/18	2020/3/25	2020/4/1
1100	1975065	1	43.06209	141.3544	0	0	0	0	0	0	2.470126008	6.58700269	10.703879	21.40776	22.2311341	6.587002688	4.11687668	13.99738
1202	251271	1	41.76864	140.7291	0	0	0	0	0	0	0.660495547	2.64198219	1.3209911	0	0	0	0	0
1203	111422	1	43.19071	140.9945	0	0	0	0	0	0	0	0	0	0	0.55974359	1.119487187	1.119487188	0.559744
1204	329513	1	43.77083	142.365	0	0	0	0	0	0	0	6.1487905	4.7823926	1.366398	0	0.683198945	2.049596835	1.366398
1205	82457	1	42.31521	140.9737	0	0	0	0	0	0	0	0.86189433	0.2154736	0	0.21547358	0	0.215473584	0
1206	165230	1	42.98492	144.3817	0	0	0	0	0	0	0	1.48322696	2.9664539	0	0.74161348	0.741613479	0	2.966454
1207	166690	1	42.92406	143.1962	0	0	0	0	0	0	0	0	0.5006939	0	0	0	0	0
1208	115608	1	43.80393	143.8958	0	0	0	0	0	0	0	1.2674926	4.2249753	1.68999	2.11248767	0	0.422497533	0.422498

Weekly human mobility by NTT Docomo

x_1

city	pop	pid	lat	lon	unit	20190102	20190109	20190116	20190123	20190130	20190206	20190213	20190220	20190227	20190306	20190313	20190320	20190327	20190403
1101	248840	1	43.05539	141.341	current	932848.1	1121272	1214639	1227891	1221009	1158744	1222635	1229758	1218756	1228485	1238896	1133306	1226419	1214379
1102	289667	1	43.09079	141.3409	current	362017	405420.5	427929.3	435397.2	427561.1	401596.8	429156.8	421940.7	428360	432645.3	435902.1	407456.5	429846.3	439032.4
1103	265536	1	43.07611	141.3636	current	289970.5	332413.9	353061.9	352963	359292.2	342211.6	357336.6	353914.7	357924.6	353912.3	356422	339307.9	357335.5	363932.7
1104	212001	1	43.04757	141.4052	current	241103.4	290852.6	307118.2	306447.1	309620.1	288871	309883.4	311787.6	317584.7	314400.1	316731.9	296877.2	317157.9	320179.5
1105	225482	1	43.03134	141.3801	current	266946.1	306153.6	317095.4	317607.8	314209.5	305139.4	309655.7	313906.9	324692.7	321360.3	319073.8	312497.3	342387.8	329129.9
1106	135966	1	42.99001	141.3534	current	103619.8	110347.2	115114.8	114605.2	111689.6	106473.7	112263.9	109858.1	108856.9	106649.6	109683.4	106070.4	112435.3	114284.9
1107	217230	1	43.07445	141.3009	current	214298.3	232732.3	243781.7	246244.4	244693.2	236917.5	251352.4	251011.3	253798.9	251537.2	252065.4	240433.6	250888.1	250633.6
1108	125182	1	43.03639	141.4748	current	147477	164746.2	167432.2	167796.4	166300	157367.3	167311.8	168259.2	172370.5	171334.3	171669.3	165373.4	173209.8	176200.8
1109	142712	1	43.12187	141.2458	current	117637.3	124718.4	127052.9	126215.4	124658.4	121647.3	124925.9	125432	124246.9	123771.6	124262.2	120983.3	126059.3	132111
1110	112449	1	42.99952	141.4438	current	137093.2	134197.3	132671.5	132747.9	131682.7	132759.4	133704.2	135297.8	136711.5	138120	138740	139601.4	138843.9	142811.8

Weekly rainfall

x_2

cid	lat	lon	20190102	20190109	20190116	20190123	20190130	20190206	20190213	20190220	20190227	20190306	20190313	20190320	20190327
11001	45.52	141.935	0.571429	0.285714	2.285714	0.142857	1.785714	2.071429	0.214286	0.5	0.5	0.142857	0.142857	1.642857	0.357143
11016	45.415	141.6783	0.857143	1.785714	2.714286	1.714286	3.071429	3.285714	1.071429	0.142857	0.5	0.714286	0.785714	2.857143	0.214286
11046	45.305	141.045	1.5	0.571429	2.214286	1.785714	1.428571	1.571429	0.785714	0	0.785714	0.5	0.571429	5.071429	0.357143
11061	45.40333	141.8017	0	0.142857	0	0.142857	0.857143	0.714286	0.142857	0.071429	0.285714	0.071429	0	1.285714	0.071429
11076	45.335	142.17	0.428571	0.714286	3.428571	0.642857	2.857143	2.142857	0.428571	0.428571	0.214286	0.142857	0.142857	0.785714	0.928571

Literature review

Functional regression

- Ramsay and Silverman. 2005. Functional Data Analysis. New York: Springer-Verlag. 2nd ed
- Benatia, et al. 2017. Functional linear regression with functional response. JoE

Functional autoregression

- Bosq, D., 2000. Linear Processes in Function Spaces: Theory and Applications. Springer-Verlag, New York.
- Liu, et al. 2016. Convolutional autoregressive models for functional time series. JoE.

Bounded linear operator

For $y \in L^2(D)$ and $x_i \in L^2(D), i = 1, \cdots, p,$

$$y = \beta_1(x_1) + \cdots + \beta_p(x_p) + u,$$

Hilbert Schmidt:

$$\beta(x) = \int \beta(s, u)x(u)du$$

Convolution:

$$\beta(x) = \int \beta(s - u)x(u)du\,,$$

Convolutional regression by CAR kernel

For $y, x_j \in L^2(D)$,

$$y = \beta_1(x_1) + \cdots + \beta_p(x_p) + \varepsilon,$$

where $\beta_j: L^2(D) \to L^2(D)$ such that

$$\beta_j(x_j)(s) = \int_D \beta_j(s-u)x_j(u)du\,,$$

For $y_t \in L^2(D)$,

$$y_t = \beta_1(y_{t-1}) + \cdots + \beta_p(y_{t-p}) + \varepsilon_t,$$

where $\beta_j: L^2(D) \to L^2(D)$ such that

$$\beta_j(y_{t-j})(s) = \int_D \beta_j(s-u)y_{t-j}(u)du$$

Ex. CAR kernel

$$\beta_j(s) = \frac{b_j}{2\pi\delta_j^2} e^{-\frac{\|s\|}{\delta_j}}, j = 1, \cdots, p.$$

Stationary condition for convolutional AR

$$y_t = \varphi_1(y_{t-1}) + \cdots + \varphi_p(y_{t-p}) + \varepsilon_t,$$

causal if

$$1 - \sum_{j=1}^{p} \|\varphi_j\| z^j \neq 0 \text{ for } |z| \leq 1.$$

the stationary condition is

$$1 - \sum_{j=1}^{p} |b_j| z^j \neq 0 \text{ for } |z| \leq 1,$$

for CAR(1) kernel

$$\varphi_j(s) = \frac{b_j}{2\pi\delta_j^2} e^{-\frac{\|s\|}{\delta_j}}.$$

Operator norm of CAR(1) kernel

For a Car(1) kernel:

$$\beta(s) = \frac{b}{2\pi\delta^2} e^{-\frac{\|s\|}{\delta}},$$

the operator norm satisfies:

$$\|\beta\| \leq |b|,$$

since, for $x(u) \in L^2(R^2)$,

$$\left\| \int_{R^2} \beta(s-u)x(u)du \right\|^2 = \left\|\tilde{\beta}(\omega)\tilde{x}(\omega)\right\|^2$$

$$\leq \left| \int_{R^2} |\beta(s)|ds \right|^2 \|x(s)\|^2.$$

$$= |b|^2$$

Summary

$$s \in D \subset R^2, t = 1,2,\cdots,$$

$$y_t(s) = \sum_{j=1}^{p} \int \varphi_j(s-u) y_{t-j}(u) du + \sum_{j=1}^{q} \int \beta_j(s-u) x_{tj}(u) du + \varepsilon_t(s),$$

We shall consider a method to estimate parametric convolutional kernels:

$$\varphi_j(s) = \varphi(\theta_j; s), \beta_j(s) = \beta(\theta_j; s).$$

Fourier transform of convolution

For

$$f(s) \in L^2(R^2),$$

Fourier transform is

$$\tilde{f}(\omega) = \int_{R^2} e^{-i(s,\omega)} f(s) ds\,, \omega \in R^2.$$

For

$$g(s) = \int f_1(s-u) f_2(u) du\,, s \in R^2,$$

$$\tilde{g}(\omega) = \int_{R^2} e^{-i(s,\omega)} g(s) ds = \int_{R^2} f_1(s-u) e^{-i(s-u,\omega)} ds \int_{R^2} f_2(u) e^{-i(u,\omega)} du = \tilde{f}_1(\omega) \tilde{f}_2(\omega)$$

Fourier transform of convolutional regression

$$y(s) = \int_D \beta_1(s-u)x_1(u)du + \cdots + \int_D \beta_p(s-u)x_p(u)du + \varepsilon(s)$$

Fourier transform:

$$\tilde{y}(\omega) = \tilde{\beta}_1(\omega)\tilde{x}_1(\omega) + \cdots + \tilde{\beta}_p(\omega)\tilde{x}_p(\omega) + \tilde{\varepsilon}(\omega)$$

$$y_t(s) = \int_D \beta_1(s-u)y_{t-1}(u)du + \cdots + \int_D \beta_p(s-u)y_{t-p}(u)du + \varepsilon_t(s)$$

Fourier transform:

$$\tilde{y}_t(\omega) = \tilde{\beta}_1(\omega)\tilde{y}_{t-1}(\omega) + \cdots + \tilde{\beta}_p(\omega)\tilde{y}_{t-p}(\omega) + \tilde{\varepsilon}_t(\omega)$$

Estimation

For a convolutional regression,

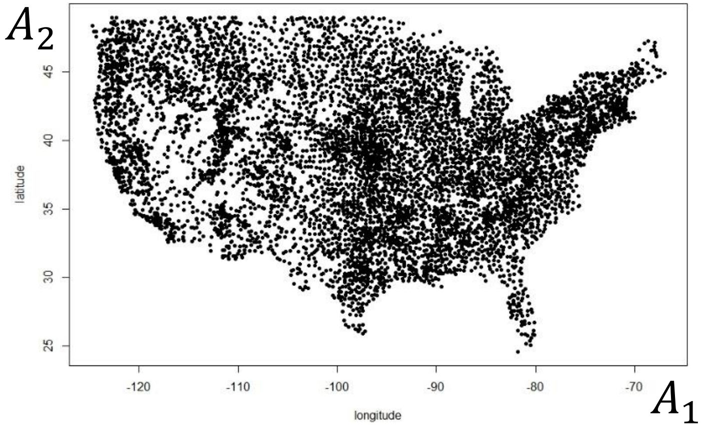
or $\tilde{y}_t(\omega) = \tilde{\beta}(\theta;\ \omega)\tilde{x}_t(\omega) + \tilde{\varepsilon}_t(\omega)$

$$y_t(s) = \int_D \beta(\theta; s-u)x_t(u)du + \varepsilon_t(s)\,,$$

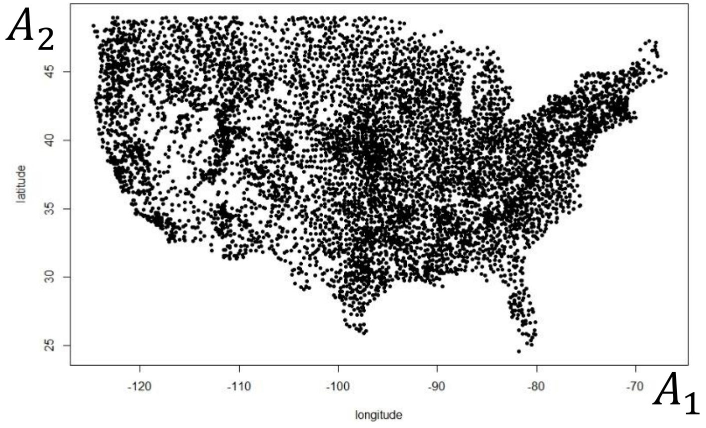
Observations:

$$\{y_t(v_{tj}), x_t(w_{tk}), s_{tj,} w_{tk} \subset D, j = 1, \cdots, n_{t1}, k = 1, \cdots, n_{t2}, t = 1, \cdots, T\}$$

$y_t(s)$



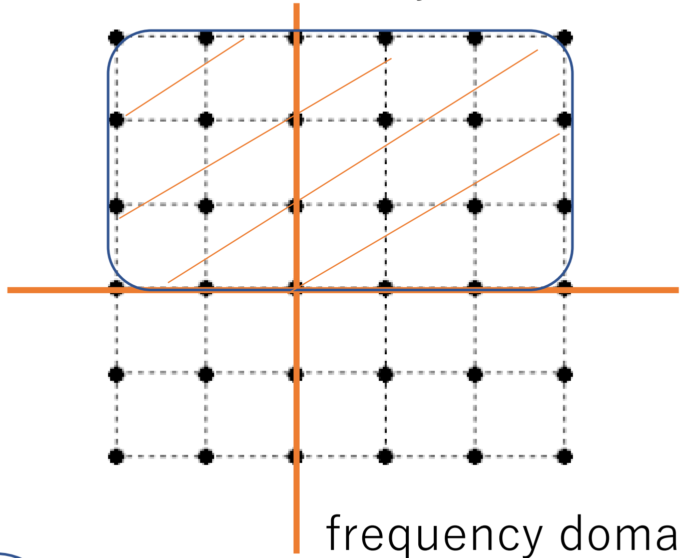
$x_t(s)$



Discrete Fourier Transform

$$\omega_p = \left(\frac{2\pi p_1}{A_1}, \frac{2\pi p_2}{A_2}\right), p = 1.\cdots, N/2$$

$$N = min(n_{jt})$$



frequency domain

For $\omega_p = (\omega_{p1}, \omega_{p2}), p = 1, \cdots, N/2$, let

$$\hat{y}_t(\omega_p) = \frac{1}{n_{0t}} \sum_{j=1}^{n_{0t}} y_t(v_{tj})\, e^{-i\omega_p' v_{tj}},$$

$$\hat{x}_t(\omega_p) = \frac{1}{n_{1t}} \sum_{j=1}^{n_{1t}} x_t(w_{tj})\, e^{-i\omega_p' w_{tj}}.$$

LSE on frequency domain:

$$\hat{\theta} = argmin\, Q(\theta),$$

where

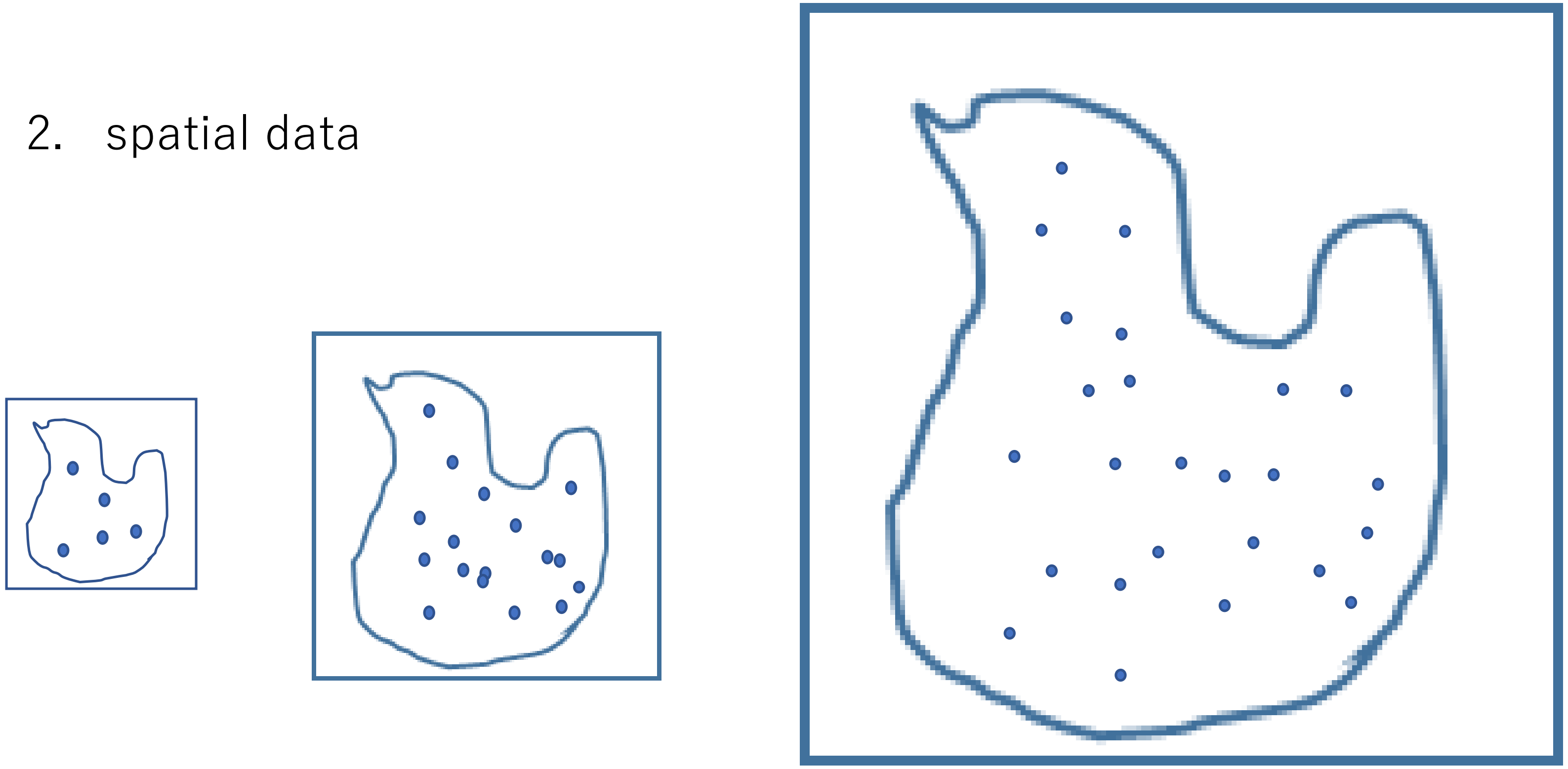
$$Q(\theta) = \sum_{t=1}^{T} \sum_{p=1}^{N/2} \left\{\hat{y}_t(\omega_p) - \tilde{\beta}(\theta; \omega_p)\hat{x}_t(\omega_p)\right\}^2,$$

$$\tilde{y}_t(\omega) = \tilde{\beta}(\theta;\ \omega)\tilde{x}_t(\omega) + \tilde{\varepsilon}_t(\omega)$$

Mixed asymptotics for spatial statistics

1. time series
   $X_1, X_2, \cdots, X_T,\ T \to \infty,$

2. spatial data



Asymptotic scheme

$y_t(s_{tj}), t = 1, \cdots, T, j = 1, \cdots, N$

finite, infinity

C1. The sample sizes n_{at} and the sampling region $A = [0, A_1] \times [0, A_2]$ diverge jointly such that $A_1 \to \infty, A_2 \to \infty, A_1/A_2 = O(1)$ and $|A|/n_{at} \to 0, a = 0, 1, \ldots, p, t = 1, \ldots, T$ for the area $|A| = A_1 \times A_2$. We shall employ a suffix k such as $A = A_k, n_{at} = n_{atk}$ when we indicate explicitly that they diverge as k tends to infinity.

finite

C2. Let S_{at} be the set of sampling points in $A = [0, A_1] \times [0, A_2]$. We assume that elements in S_{at} are written as, for $a = 0, 1, \ldots, p, t = 1, \ldots, T$,

$$s_{atj} = (A_1 u_{1,atj}, A_2 u_{2,atj}), j = 1, \ldots, n_{at},$$

where $u_{atj} = (u_{1,atj}, u_{2,atj})$ is a sequence of independent and identically distributed random vectors with a probability density function $g(x)$ supported on $[0, 1]^2$ which has continuous first derivatives.

conditions

$$E\{Z_t(s)Z_t(u)\} = \gamma_t(s-u)$$

$$f_{t,zz}(\omega) = \begin{pmatrix} f_{t,yy}(\omega) & f_{t,yx}(\omega) \\ f_{t,xy}(\omega) & f_{t,xx}(\omega) \end{pmatrix}$$

C3. $z_t(s) = (y_t(s), x_{t1}(s), \ldots, x_{tp}(s))$ is a zero-mean stationary random field with respect to $s \in \mathbb{R}^2$ for $t = 1, \ldots, T$ with finite moments of all orders and with a positive definite spectral density matrix of $f_{t,zz}(\omega)$. $x_{ta}(s), a = 1, \ldots, p$ is exogenous in the sense that $\varepsilon_t(s)$ is independent of $x_{ta}(u)$ for any $s, u \in \mathbb{R}^2$ and $t = 1, \ldots, T$.

C4. Let Θ be a compact parameter space in $\mathbb{R}^q$ for $\theta = (\theta_1, \ldots, \theta_q)$, and D be a fixed symmetric compact region on $\mathbb{R}^2$. $\tilde{\phi}(\omega; \theta)$ has a continuous first derivative with respect to θ on $\Theta \times D$. $\theta_1 \neq \theta_2$ implies that $\tilde{\phi}(\omega; \theta_1) \neq \tilde{\phi}(\omega; \theta_2)$ on a subset of D with positive Lebesgue measure.

consistency

Theorem 1. *Under Assumptions C2-C4,*

$$\hat{\theta} \rightarrow_p \theta_0$$

in the asymptotic scheme defined by C1.

Asymptotic normality

Theorem 2. *Under Assumptions C2-C4, if* $|A|^{3/2}/n_{at} = o(1), a = 0, 1, \ldots, p,$

$$\sqrt{|A|}(\hat{\theta} - \theta_0) \to N\left(0, b_g \Omega^{-1} \Sigma \Omega^{-1}\right)$$

in the asymptotic scheme defined by C1, where

$$b_g = (2\pi)^2 \left\{ \int_{[0,1]^2} |g(u)|^4 du \right\} \left\{ \int_{[0,1]^2} |g(u)|^2 du \right\}^{-2},$$

$$\Omega = \int_D \dot{\Phi}(\omega; \theta_0) \left\{ \sum_{t=1}^{T} f_{t,xx}(\omega) \right\} \dot{\Phi}'(\omega; \theta_0) d\omega,$$

$$\Sigma = \int_D \dot{\Phi}(\omega; \theta_0) \left\{ \sum_{t=1}^{T} f_{t,\varepsilon\varepsilon}(\omega) f_{t,xx}(\omega) \right\} \dot{\Phi}'(\omega; \theta_0) d\omega,$$

where $\dot{\Phi}(\omega; \theta)$ *is the* q *by* p *matrix whose* (i, j)*th element is*

$$\frac{\partial \tilde{\phi}_j(\omega; \theta)}{\partial \theta_i}, i = 1, \ldots, q, j = 1, \ldots, p.$$

Consistent estimation for asympt. variance

Then the asymptotic variance matrix is consistently estimated by

$$\hat{c}_g \hat{\Omega}^{-1}(\hat{\theta})\hat{\Sigma}(\hat{\theta})\hat{\Omega}(\hat{\theta})^{-1},$$

where

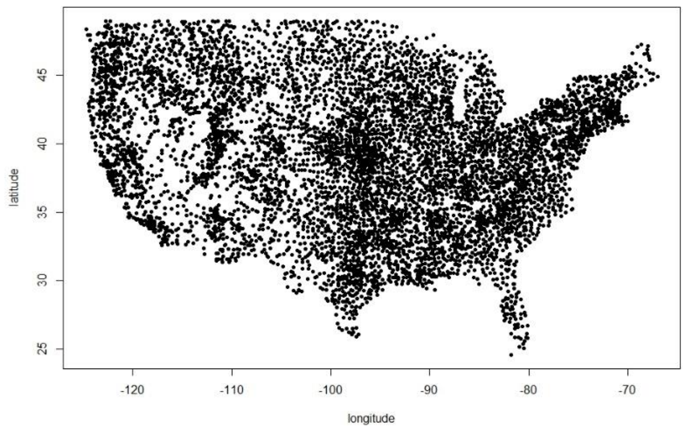
$$\hat{\Omega}(\theta) = \sum_{\omega_f \in D} \dot{\Phi}(\omega_f;\theta) \left\{ \sum_{t=1}^{T} \hat{x}_t(\omega_f)\overline{\hat{x}_t(\omega_f)}' \right\} \dot{\Phi}(\omega_f;\theta)',$$

$$\hat{\Sigma}(\theta) = \sum_{\omega_f \in D} \dot{\Phi}(\omega_f;\theta) \left\{ \sum_{t=1}^{T} |\hat{\varepsilon}_t(\omega_f;\theta)|^2 \, \hat{x}_t(\omega_f)\overline{\hat{x}_t(\omega_f)}' \right\} \dot{\Phi}(\omega_f;\theta)',$$

$$\hat{\varepsilon}_t(\omega_f;\theta) = \hat{y}_t(\omega_f) - \sum^{p} \tilde{\phi}_a(\omega_f;\theta)\hat{x}_{ta}(\omega_f).$$

US ppt data monthly from Jan. 1986 till Dec. 1991

Func. regression: $y(s) = \beta_1(x_1(s)) + \beta_2(x_2(s)) + u(s),$



US precipitation data

$y(s) \in L^2(D)$

lon	lat	Jan./1986	Feb	March	April
-85.25	31.57	8.9	11.4	10.2	2.5
-87.18	34.22	2.6	6.4	8.1	1.1
-87.32	34.42	NA	NA	NA	NA
-87.42	32.23	4.5	8.1	9.5	0.9
-86.22	34.25	NA	NA	NA	NA
-85.95	32.95	2.6	7.5	15.9	0.9
-85.87	32.98	NA	NA	NA	NA
-88.13	33.13	2.6	7.3	9.5	NA
-88.28	33.23	4.2	12.2	13.6	1.5
-86.5	31.32	16.3	9.3	19.3	1.8
-85.85	33.58	2.4	5.1	6.2	0.4
-87.22	34.07	NA	NA	NA	NA
-85.83	33.27	1.2	8.3	NA	1.5
-86.27	33.83	NA	NA	NA	NA
-86.98	34.8	NA	NA	NA	NA
-87.48	31.17	18.2	9.8	12.7	1.7
-87.52	31.02	NA	NA	NA	NA
-85.5	32.6	NA	NA	NA	NA
-86.68	32.47	NA	8.8	18.3	1.1
-87.35	33.45	4.9	6.5	6.2	2.1
-87.78	30.88	NA	NA	NA	NA

Temperature (max)

$x_1(s) \in L^2(D)$

Jan	Feb	Mar	Apr
NA	NA	NA	NA
14.4	18.7	21.2	26.7
NA	NA	NA	NA
13	16.9	20.8	NA
NA	NA	21.8	27
12.6	16.5	20.3	24.9
13.6	17.9	NA	25.9
13	14.3	17.8	24.4

Temperature (min)

$x_2(s) \in L^2(D)$

Jan	Feb	Mar	Apr
NA	NA	NA	NA
-2.8	2.4	4.1	8.2
NA	NA	NA	NA
-2.9	3.5	3.6	NA
-0.3	4.1	4.8	7
-1.2	3.3	5.1	7.8
-3.9	1.2	NA	6.2
-2.4	0.9	2.6	6.4

Application to US precipitation data

Estimation by 60 months from Jan. 1986 till Dec. 1990.

$$ppt(s) = \int \underset{(0.198)}{1.87}\varphi(\underset{(0.051)}{0.502};\ s-u)tmin(u)du - \int \underset{(0.155)}{1.63}\varphi(\underset{(0.027)}{0.310};\ s-u)tmax(u)du$$

$$\varphi(\delta;\ s) = \frac{1}{2\pi\delta^2}e^{-\frac{\|s\|}{\delta}}$$

In-sample: Jan. 1986 - Dec. 1990. out-of-sample: Jan. –Dec.1991

	convolution by CAR kernel		Nonp. Hilbert Schmidt
	freq.	spatal	func. PCA
in-sample MSE	16.0	15.5	12.2
out-of-sample MSE	26.2	25.3	32.0

NTT Docomo Mobile Statistics

モバイル空間統計®

(weekly from Jan. 2020 till Sep. 2021)

Analysis of Covid-19 pandemic in Japan

Weekly new cases of covid-19

B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
ity	pop	pid	lat	lon	2020/1/1	2020/1/8	2020/1/15	2020/1/22	2020/1/29	2020/2/5	2020/2/12	2020/2/19	2020/2/26	2020/3/4	2020/3/11	2020/3/18	2020/3/25	2020/4/1
1100	1975065	1	43.06209	141.3544	0	0	0	0	0	0	2.470126008	6.58700269	10.703879	21.40776	22.2311341	6.587002688	4.11687668	13.99738
1202	251271	1	41.76864	140.7291	0	0	0	0	0	0	0.660495547	2.64198219	1.3209911	0	0	0	0	0
1203	111422	1	43.19071	140.9945	0	0	0	0	0	0	0	0	0	0	0.55974359	1.119487187	1.119487188	0.559744
1204	329513	1	43.77083	142.365	0	0	0	0	0	0	0	6.14879051	4.7823926	1.366398	0	0.683198945	2.049596835	1.366398
1205	82457	1	42.31521	140.9737	0	0	0	0	0	0	0	0.86189433	0.2154736	0	0.21547358	0	0.215473584	0
1206	165230	1	42.98492	144.3817	0	0	0	0	0	0	0	1.48322696	2.9664539	0	0.74161348	0.741613479	0	2.966454
1207	166690	1	42.92406	143.1962	0	0	0	0	0	0	0	0	0.5006939	0	0	0	0	0
1208	115608	1	43.80393	143.8958	0	0	0	0	0	0	0	1.2674926	4.2249753	1.68999	2.11248767	0	0.422497533	0.422498

Weekly human mobility by NTT Docomo

city	pop	pid	lat	lon	unit	20190102	20190109	20190116	20190123	20190130	20190206	20190213	20190220	20190227	20190306	20190313	20190320	20190327	20190403
1101	248840	1	43.05539	141.341	current	932848.1	1121272	1214639	1227891	1221009	1158744	1222635	1229758	1218756	1228485	1238896	1133306	1226419	1214379
1102	289667	1	43.09079	141.3409	current	362017	405420.5	427929.3	435397.2	427561.1	401596.8	429156.8	421940.7	428360	432645.3	435902.1	407456.5	429846.3	439032.4
1103	265536	1	43.07611	141.3636	current	289970.5	332413.9	353061.9	352963	359292.2	342211.6	357336.6	353914.7	357924.6	353912.3	356422	339307.9	357335.5	363932.7
1104	212001	1	43.04757	141.4052	current	241103.4	290852.6	307118.2	306447.1	309620.1	288871	309883.4	311787.6	317584.7	314400.1	316731.9	296877.2	317157.9	320179.5
1105	225482	1	43.03134	141.3801	current	266946.1	306153.6	317095.4	317607.8	314209.5	305139.4	309655.7	313906.9	324692.7	321360.3	319073.8	312497.3	342387.8	329129.9
1106	135966	1	42.99001	141.3534	current	103619.8	110347.2	115114.8	114605.2	111689.6	106473.7	112263.9	109858.1	108856.9	106649.6	109683.4	106070.4	112435.3	114284.9
1107	217230	1	43.07445	141.3009	current	214298.3	232732.3	243781.7	246244.4	244693.2	236917.5	251352.4	251011.3	253798.9	251537.2	252065.4	240433.6	250888.1	250633.6
1108	125182	1	43.03639	141.4748	current	147477	164746.2	167432.2	167796.4	166300	157367.3	167311.8	168259.2	172370.5	171334.3	171669.3	165373.4	173209.8	176200.8
1109	142712	1	43.12187	141.2458	current	117637.3	124718.4	127052.9	126215.4	124658.4	121647.3	124925.9	125432	124246.9	123771.6	124262.2	120983.3	126059.3	132111
1110	112449	1	42.99952	141.4438	current	137093.2	134197.3	132671.5	132747.9	131682.7	132759.4	133704.2	135297.8	136711.5	138120	138740	139601.4	138843.9	142811.8

Weekly rainfall

cid	lat	lon	20190102	20190109	20190116	20190123	20190130	20190206	20190213	20190220	20190227	20190306	20190313	20190320	20190327
11001	45.52	141.935	0.571429	0.285714	2.285714	0.142857	1.785714	2.071429	0.214286	0.5	0.5	0.142857	0.142857	1.642857	0.357143
11016	45.415	141.6783	0.857143	1.785714	2.714286	1.714286	3.071429	3.285714	1.071429	0.142857	0.5	0.714286	0.785714	2.857143	0.214286
11046	45.305	141.045	1.5	0.571429	2.214286	1.785714	1.428571	1.571429	0.785714	0	0.785714	0.5	0.571429	5.071429	0.357143
11061	45.40333	141.8017	0	0.142857	0	0.142857	0.857143	0.714286	0.142857	0.071429	0.285714	0.071429	0	1.285714	0.071429
11076	45.335	142.17	0.428571	0.714286	3.428571	0.642857	2.857143	2.142857	0.428571	0.428571	0.214286	0.142857	0.142857	0.785714	0.928571

Identified model

$$y(s) = \sum_{j=1}^{p} b_j \int \varphi\left(\delta_j; s-u\right) x_j(u) du + \varepsilon(s)$$

$$\varphi(\delta;\ s) = \frac{1}{2\pi\delta^2} e^{-\frac{\|s\|}{\delta}}$$

$covid19_t$

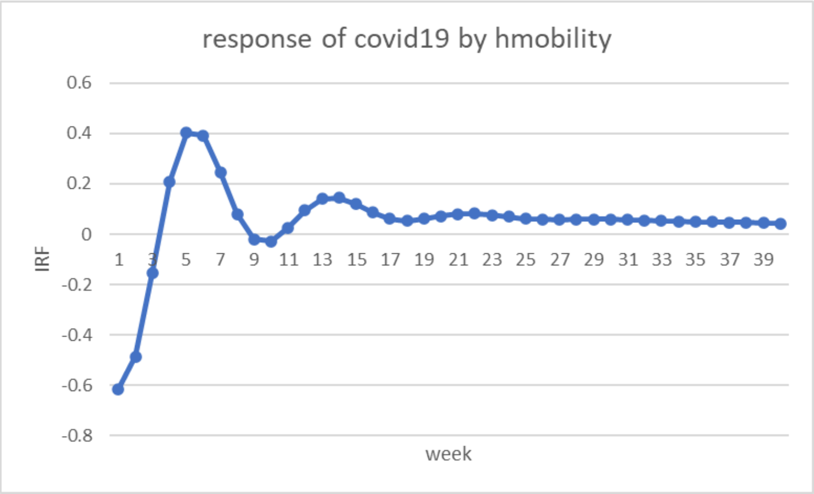
	b	se	delta	se
covid_{t-1}	1.18	0.015	0.13	0.0051
covid_{t-2}	-0.66	0.023	0.33	0.013
hmobil_{t-1}	-0.63	0.068	0.17	0.034
hmobil_{t-2}	0.69	0.074		
rainfall_{t-2}	-0.0071	0.0038	1.11	0.90

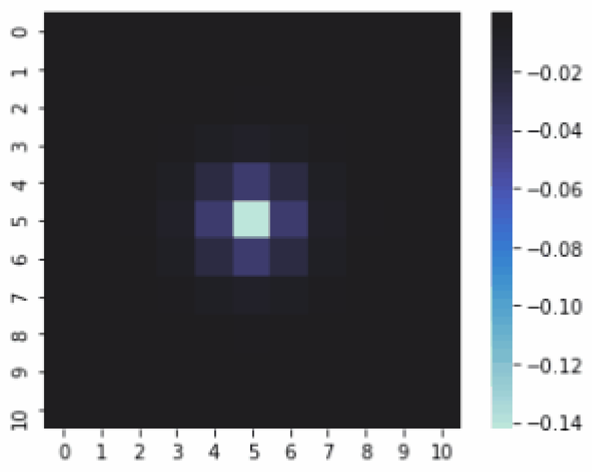
$hmobil_t$

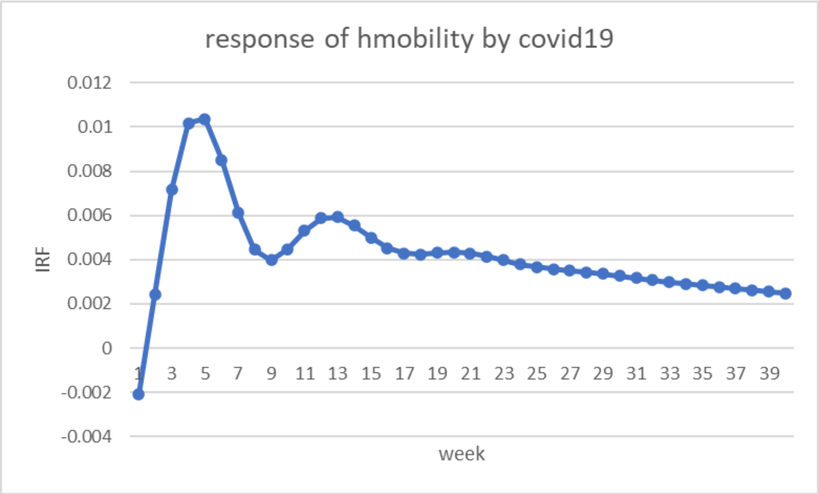
	b	se	delta	se
hmobil_{t-1}	0.71	0.0075	0	
hmobil_{t-2}	0.26	0.0076	0	
covid_{t-1}	-0.0021	0.0005	0.050	0.091
covid_{t-2}	0.0065	0.00073		
rainfall_t	-0.0012	0.00013	0.28	0.074

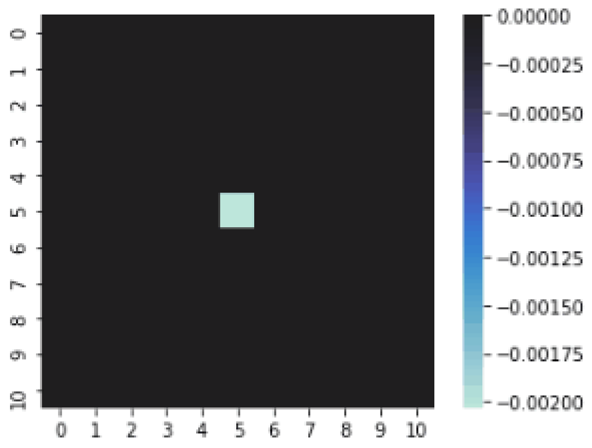
$$\int b_j \varphi\left(\delta_j;\ s-u\right) x_j(u) du \to b_j x_j(s)\ as\ \delta_j \to 0.$$

Impulse response function









Conclusion

- Intercept term
- Functional regression for spatial data

$$y(s) = \int_D \beta(s-u)x(u)du \text{, convolution kernel}$$

$$y(s) = \int_D \beta(s,u)x(u)du \text{, Hilbert Schmidt kernel}$$

Usual approach of nonparametric (spline or PCA) Hilbert Schmidt kernels have overfitting estimation.