GENEVA SCHOOL OF ECONOMICS AND MANAGEMENT

UNIVERSITÉ DE GENÈVE

Statistical analysis of network data: learning from small samples (with a couple of detours)

Davide La Vecchia

University of Geneva

June 30th, 2022

Three-fold aim of the talk

For **small panel data**, I:

Three-fold aim of the talk

For **small panel data**, I:

- illustrate that first-order asymptotic theory suffers from finite sample distortions

Three-fold aim of the talk

For **small panel data**, I:

- illustrate that first-order asymptotic theory suffers from finite sample distortions

- develop saddlepoint techniques which perform well in small samples and feature higher-order accuracy

Three-fold aim of the talk

For **small panel data**, I:

- illustrate that first-order asymptotic theory suffers from finite sample distortions
- develop saddlepoint techniques which perform well in small samples and feature higher-order accuracy
- compare the saddlepoint density approximation to the Edgeworth expansion and asymptotic approximation

Three-fold aim of the talk

For **small panel data**, I:

- illustrate that first-order asymptotic theory suffers from finite sample distortions

- develop saddlepoint techniques which perform well in small samples and feature higher-order accuracy

- compare the saddlepoint density approximation to the Edgeworth expansion and asymptotic approximation

This talk is based on

- *Saddlepoint approximations for spatial panel data models,* **JASA, 2022**, with C.Jiang, E.Ronchetti and O.Scaillet
- *On some connections between Esscher's titlting, saddlepoint approximations, and optimal transportation: a statistical perspective*, **Stat. Science, 2022** with E.Ronchetti and A.Ilievski

Three-fold aim of the talk

For **small network data**, I:

- illustrate that first-order asymptotic theory suffers from finite sample distortions
- develop saddlepoint techniques which perform well in small samples and feature higher-order accuracy
- compare the saddlepoint density approximation to the Edgeworth expansion and asymptotic approximation

This talk is based on

- *Saddlepoint approximations for spatial panel data models*, **JASA, 2022**, with C.Jiang, E.Ronchetti and O.Scaillet
- *On some connections between Esscher's titlting, saddlepoint approximations, and optimal transportation: a statistical perspective*, **Stat. Science, 2022** with E.Ronchetti and A.Ilievski

The problem in spatial econometrics terms

Feldstein and Horioka (1980) investigated the following points:

The problem in spatial econometrics terms

Feldstein and Horioka (1980) investigated the following points:

- Does capital flow among industrial countries to equalise the yield to investors?

The problem in spatial econometrics terms

Feldstein and Horioka (1980) investigated the following points:

- Does capital flow among industrial countries to equalise the yield to investors?
- Alternatively, does the saving that originates in a country remain to be invested there?

The problem in spatial econometrics terms

Feldstein and Horioka (1980) investigated the following points:

- Does capital flow among industrial countries to equalise the yield to investors?
- Alternatively, does the saving that originates in a country remain to be invested there?
- Or does the truth lie somewhere between these two extremes?

The problem in spatial econometrics terms

Feldstein and Horioka (1980) investigated the following points:

- Does capital flow among industrial countries to equalise the yield to investors?
- Alternatively, does the saving that originates in a country remain to be invested there?
- Or does the truth lie somewhere between these two extremes?

Eventually, they found a large, positive (cor)relation of domestic saving rates on domestic investment rates for Organisation for Economic Co-operation and Development (OECD) countries. They interpret this finding as evidence of a high degree of frictions reducing capital flows between countries.

The problem in spatial econometrics terms

Feldstein and Horioka (1980) investigated the following points:

- Does capital flow among industrial countries to equalise the yield to investors?
- Alternatively, does the saving that originates in a country remain to be invested there?
- Or does the truth lie somewhere between these two extremes?

Eventually, they found a large, positive (cor)relation of domestic saving rates on domestic investment rates for Organisation for Economic Co-operation and Development (OECD) countries. They interpret this finding as evidence of a high degree of frictions reducing capital flows between countries. Re-considering FH study, Debarsy & Ertur (2010) analyzed Investment and Saving rates for 24 OECD countries between 1960 and 2000 (41 yrs):

$$\{\mathrm{Inv}_{it}\} \qquad \text{and} \qquad \{\mathrm{Sav}_{it}\}$$

for $i = 1, .., n$ (cross-sectional dimension, $n = 24$) and $t = 1, ..., T$ (time series dimension, $T = 41$).

The problem in spatial econometrics terms

Feldstein and Horioka (1980) investigated the following points:

- Does capital flow among industrial countries to equalise the yield to investors?
- Alternatively, does the saving that originates in a country remain to be invested there?
- Or does the truth lie somewhere between these two extremes?

Eventually, they found a large, positive (cor)relation of domestic saving rates on domestic investment rates for Organisation for Economic Co-operation and Development (OECD) countries. They interpret this finding as evidence of a high degree of frictions reducing capital flows between countries. Re-considering FH study, Debarsy & Ertur (2010) analyzed Investment and Saving rates for 24 OECD countries between 1960 and 2000 (41 yrs):

$$\boxed{\{\mathrm{Inv}_{it}\} \quad \text{and} \quad \{\mathrm{Sav}_{it}\}}$$

for $i = 1, .., n$ (cross-sectional dimension, $n = 24$) and $t = 1, ..., T$ (time series dimension, $T = 41$).

The problem in spatial econometrics terms

Feldstein and Horioka (1980) investigated the following points:

- Does capital flow among industrial countries to equalise the yield to investors?
- Alternatively, does the saving that originates in a country remain to be invested there?
- Or does the truth lie somewhere between these two extremes?

Eventually, they found a large, positive (cor)relation of domestic saving rates on domestic investment rates for Organisation for Economic Co-operation and Development (OECD) countries. They interpret this finding as evidence of a high degree of frictions reducing capital flows between countries. Re-considering FH study, Debarsy & Ertur (2010) analyzed Investment and Saving rates for 24 OECD countries between 1960 and 2000 (41 yrs):

$$\boxed{\underbrace{\{\mathrm{Inv}_{it}\} \quad \overset{\forall i}{\Longleftrightarrow} \quad \{\mathrm{Sav}_{it}\}}_{\text{as in FH}}}$$

for $i = 1, .., n$ (cross-sectional dimension, $n = 24$) and $t = 1, \ldots T$ (time series dimension, $T = 41$).

The problem in spatial econometrics terms

Feldstein and Horioka (1980) investigated the following points:

- Does capital flow among industrial countries to equalise the yield to investors?
- Alternatively, does the saving that originates in a country remain to be invested there?
- Or does the truth lie somewhere between these two extremes?

Eventually, they found a large, positive (cor)relation of domestic saving rates on domestic investment rates for Organisation for Economic Co-operation and Development (OECD) countries. They interpret this finding as evidence of a high degree of frictions reducing capital flows between countries. Re-considering FH study, Debarsy & Ertur (2010) analyzed Investment and Saving rates for 24 OECD countries between 1960 and 2000 (41 yrs):

$$\boxed{\{\mathrm{Inv}_{it}\} \overset{?}{\Longleftrightarrow} \{\mathrm{Inv}_{kt}\} \overset{?}{\Longleftrightarrow} \{\mathrm{Inv}_{jt}\} \overset{?}{\Longleftrightarrow} \underbrace{\{\mathrm{Inv}_{it}\} \Longleftrightarrow \{\mathrm{Sav}_{it}\}}_{\text{as in FH}}}$$

for $i = 1, .., n$ (cross-sectional dimension, $n = 24$) and $t = 1, ..., T$ (time series dimension, $T = 41$).

The problem in spatial econometrics terms

Feldstein and Horioka (1980) investigated the following points:

- Does capital flow among industrial countries to equalise the yield to investors?
- Alternatively, does the saving that originates in a country remain to be invested there?
- Or does the truth lie somewhere between these two extremes?

Eventually, they found a large, positive (cor)relation of domestic saving rates on domestic investment rates for Organisation for Economic Co-operation and Development (OECD) countries. They interpret this finding as evidence of a high degree of frictions reducing capital flows between countries. Re-considering FH study, Debarsy & Ertur (2010) analyzed Investment and Saving rates for 24 OECD countries between 1960 and 2000 (41 yrs):

$$\boxed{\underbrace{\{\mathrm{Inv}_{it}\} \overset{?}{\Longleftrightarrow} \{\mathrm{Inv}_{kt}\} \overset{?}{\Longleftrightarrow} \{\mathrm{Inv}_{jt}\} \overset{?}{\Longleftrightarrow} \{\mathrm{Inv}_{it}\}}_{\text{spillover}} \Longleftrightarrow \{\mathrm{Sav}_{it}\}}$$

for $i = 1, .., n$ (cross-sectional dimension, $n = 24$) and $t = 1, ..., T$ (time series dimension, $T = 41$).

The problem in spatial econometrics terms

Inference goal

Test for the presence of spillover (spatial autocorrelation) between country i and country j, $i \neq j$, in the investment-saving relationship, e.g. using p-value, as in the finance/economics literature.

The problem in spatial econometrics terms

Inference goal

Test for the presence of spillover (spatial autocorrelation) between country i and country j, $i \neq j$, in the investment-saving relationship, e.g. using p-value, as in the finance/economics literature.

In Debarsy & Ertur econometric words: test if a "change in the saving rate in one country, i say, affects the investment rate of that country, which in turn affects the investment rates of other countries (j, k say), which then feed back to the investment rate of country i."

The problem in spatial econometrics terms

The Dataset at a glance. As in Debarsy & Ertur (2010), countries considered are: Australia, Austria, Belgium, Canada, Switzerland, Denmark, Spain, Finland, France, United Kingdom, Greece, Ireland, Iceland, Italy, Japan, Mexico, Netherlands, Norway, New Zealand, Portugal, Republic of Korea, Sweden, Turkey and United States.

Ratios of Investment come from the Penn World Table and are defined as investment share of real GDP per capita. Ratios of Savings are defined as the percentage share of current savings to GDP. They are actually computed as subtracting consumption share and government share of real gross domestic product per capita, from 100.

The Table reports some descriptive statistics for the two variables under consideration:

Descriptive statistics for ratio of investment and saving.

	Obs.	Mean	Std deviation	Min	Max
Investment	984	25.54793	5.987409	9.798207	44.85361
Savings	984	24.84542	6.990893	0.1976267	41.84594

The problem in spatial econometrics terms

In the econometric literature (see Debarsy & Ertur (2010)), the problem is tackled using spatial panel data models; see e.g. Lee (2004), Lee and Yu (2010).

The problem in spatial econometrics terms

In the econometric literature (see Debarsy & Ertur (2010)), the problem is tackled using spatial panel data models; see e.g. Lee (2004), Lee and Yu (2010).

We consider a Gaussian random field described by the SARAR(1,1) (Kelejian and Prucha (1998))

$$
\begin{aligned}
\mathrm{Inv}_{nt} &= \lambda_0 W_n \,\mathrm{Inv}_{nt} + \mathrm{Sav}_{nt}\beta_0 + c_{n0} + E_{nt}, \\
E_{nt} &= \rho_0 M_n E_{nt} + V_{nt}, \qquad\qquad t = 1, 2, \cdots, T
\end{aligned}
$$

The problem in spatial econometrics terms

In the econometric literature (see Debarsy & Ertur (2010)), the problem is tackled using spatial panel data models; see e.g. Lee (2004), Lee and Yu (2010).

We consider a Gaussian random field described by the SARAR(1,1) (Kelejian and Prucha (1998))

$$\begin{aligned} \mathrm{Inv}_{nt} &= \lambda_0 W_n \,\mathrm{Inv}_{nt} + \mathrm{Sav}_{nt}\beta_0 + c_{n0} + E_{nt}, \\ E_{nt} &= \rho_0 M_n E_{nt} + V_{nt}, \qquad t = 1, 2, \cdots, T \end{aligned} \tag{1}$$

- Inv_{nt} is the $n \times 1$ vector of investment rates for all countries
- Sav_{nt} is the $n \times 1$ vector of saving rates (time-varying)
- V_{nt} is an $n \times 1$ vector and each element v_{it} in it is i.i.d across i and t, having Gaussian distribution with zero mean and variance σ_0^2.
- c_{n0} is an $n \times 1$ vector of fixed effects
- **the scalars λ_0 and ρ_0 control for spatial autocorrelation**
- we assume $W_n = M_n$ and these matrices control for the spatial relationships (e.g. Rook, Queen, k-nearest neighbours)...

The problem in spatial econometrics terms

The matrix W_n (M_n) is an $n \times n$ spatial weight matrix (also called contiguity/connectivity matrix) that generates the spatial dependence among cross sectional units.

The problem in spatial econometrics terms

The matrix W_n (M_n) is an $n \times n$ spatial weight matrix (also called contiguity/connectivity matrix) that generates the spatial dependence among cross sectional units.

To gain some understanding, let us consider the SAR(1) model

$$\mathrm{Inv}_{nt} = \lambda_0 W_n \, \mathrm{Inv}_{nt} + V_{nt}$$

The problem in spatial econometrics terms

The matrix W_n (M_n) is an $n \times n$ spatial weight matrix (also called contiguity/connectivity matrix) that generates the spatial dependence among cross sectional units.

To gain some understanding, let us consider the SAR(1) model

$$\begin{aligned} \mathrm{Inv}_{nt} &= \lambda_0 W_n \,\mathrm{Inv}_{nt} + V_{nt} \\ &= \mathcal{L}(W_n)\,\mathrm{Inv}_{nt} + V_{nt}, \end{aligned}$$

The problem in spatial econometrics terms

The matrix W_n (M_n) is an $n \times n$ spatial weight matrix (also called contiguity/connectivity matrix) that generates the spatial dependence among cross sectional units.

To gain some understanding, let us consider the SAR(1) model

$$\begin{aligned} \text{Inv}_{nt} &= \lambda_0 W_n \,\text{Inv}_{nt} + V_{nt} \\ &= \mathcal{L}(W_n)\,\text{Inv}_{nt} + V_{nt}, \end{aligned}$$

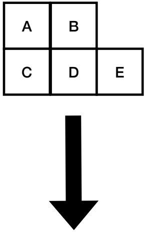
where I emphasize an analogy to time series:

time series AR(1)	spatial AR(1)
$\phi(L)X_t = \phi_1 L X_t$	$\mathcal{L}(W_m)\text{Inv}_{nt} = \lambda_0 W_n \text{Inv}_{nt}.$

The problem in spatial econometrics terms

For instance, on a plane, one may define a Rook-type contiguity matrix W_n as

Underlying Geometry
n = 5



$w_{ij} = 1$ **if** $i \sim j$ **and** 0 **otherwise**

Rook-type W_n

Original matrix

	A	B	C	D	E
A	**0**	1	1	0	0
B	1	**0**	0	1	0
C	1	0	**0**	1	0
D	0	1	1	**0**	1
E	0	0	0	1	**0**

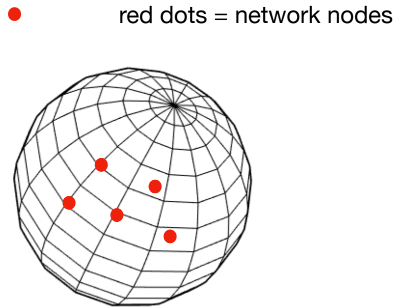
Row normalisation

	A	B	C	D	E
A	**0**	1/2	1/2	0	0
B	1/2	**0**	0	1/2	0
C	1/2	0	**0**	1/2	0
D	0	1/3	1/3	**0**	1/3
E	0	0	0	1	**0**

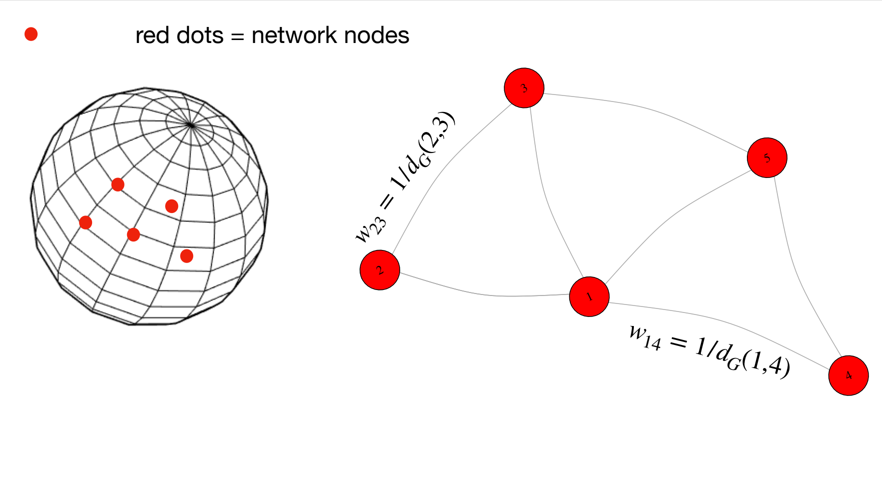
Remark: zero diagonal elements $\rightarrow$ no self-influence.

... or we may consider points scattered over a sphere (more generally a Riemaniann manifold), where w_{ij} is determined using the geodesic distance

... or we may consider points scattered over a sphere (more generally a Riemaniann manifold), where w_{ij} is determined using the geodesic distance



... or we may consider points scattered over a sphere (more generally a Riemaniann manifold), where w_{ij} is determined using the geodesic distance (d_G)



The problem in spatial econometrics terms

SARAR(1,1) model as in Debarsy & Ertur (2010):

- Because of macroeconomic reasons (deregulating financial markets), the period can be divided into shorter sub-periods

The problem in spatial econometrics terms

SARAR(1,1) model as in Debarsy & Ertur (2010):

- Because of macroeconomic reasons (deregulating financial markets), the period can be divided into shorter sub-periods
- W_n specified as either inverse distance or 7 nearest neighbours

The problem in spatial econometrics terms

SARAR(1,1) model as in Debarsy & Ertur (2010):

- Because of macroeconomic reasons (deregulating financial markets), the period can be divided into shorter sub-periods
- W_n specified as either inverse distance or 7 nearest neighbours
- Gaussian MLE of the parameters β_0, λ_0, ρ_0

The problem in spatial econometrics terms

SARAR(1,1) model as in Debarsy & Ertur (2010):

- Because of macroeconomic reasons (deregulating financial markets), the period can be divided into shorter sub-periods
- W_n specified as either inverse distance or 7 nearest neighbours
- Gaussian MLE of the parameters β_0, λ_0, ρ_0

	Inverse distance			7 nearest neighbours		
	1960-1970	1971-1985	1986-2000	1960-1970	1971-1985	1986-2000
β_0	0.935	0.638	0.356	0.932	0.633	0.368
λ_0	0.004	0.381	0.430	-0.016	0.340	0.437
ρ_0	-0.305	0.334	0.222	-0.219	0.258	0.025

The problem in spatial econometrics terms

The goal is to test for spatial dependency in the investments, thus:

$$\underbrace{\mathcal{H}_0 : \; \lambda_0 = 0}_{\text{no spatial dep.}} \quad \text{vs} \quad \underbrace{\mathcal{H}_1 : \; \lambda_0 > 0}_{\text{pos. spatial dep.}}.$$

The problem in spatial econometrics terms

The goal is to test for spatial dependency in the investments, thus:

$$\underbrace{\mathcal{H}_0 : \ \lambda_0 = 0}_{\text{no spatial dep.}} \quad \text{vs} \quad \underbrace{\mathcal{H}_1 : \ \lambda_0 > 0}_{\text{pos. spatial dep.}}.$$

Other interesting hypotheses refer to $\rho_0 = 0$ and $\lambda_0 = \rho_0 = 0$.

The problem in spatial econometrics terms

The goal is to test for spatial dependency in the investments, thus:

$$\underbrace{\mathcal{H}_0 : \ \lambda_0 = 0}_{\text{no spatial dep.}} \quad \text{vs} \quad \underbrace{\mathcal{H}_1 : \ \lambda_0 > 0}_{\text{pos. spatial dep.}}.$$

Other interesting hypotheses refer to $\rho_0 = 0$ and $\lambda_0 = \rho_0 = 0$. Standard 1st order Gaussian asymptotic theory (Gaussian MLE, $n \to \infty$ and T fix) is applied and p-values are computed.

The problem in spatial econometrics terms

The goal is to test for spatial dependency in the investments, thus:

$$\underbrace{\mathcal{H}_0 :\ \lambda_0 = 0}_{\text{no spatial dep.}} \quad \text{vs} \quad \underbrace{\mathcal{H}_1 :\ \lambda_0 > 0}_{\text{pos. spatial dep.}}.$$

Other interesting hypotheses refer to $\rho_0 = 0$ and $\lambda_0 = \rho_0 = 0$. Standard 1st order Gaussian asymptotic theory (Gaussian MLE, $n \to \infty$ and T fix) is applied and p-values are computed.

		Inverse distance			7 nearest neighbours		
		1960-1970	1971-1985	1986-2000	1960-1970	1971-1985	1986-2000
$\lambda_0 = 0$	ASY	1.0000	0.0116	<u>0.5679</u>	0.9987	0.0130	<u>0.1123</u>
$\rho_0 = 0$	ASY	0.5890	0.2261	0.9578	0.7101	0.3898	0.9998
$\lambda_0 = \rho_0 = 0$	ASY	0.4615	0.0000	0.0000	0.5042	0.0000	0.0000

The problem in spatial econometrics terms

Issue

- *Is the use of asymptotics legitimate? Is the asymptotics suffering from size distortion due to the small samples, in both the cross section and in time?*

The problem in spatial econometrics terms

Issue

- *Is the use of asymptotics legitimate? Is the asymptotics suffering from size distortion due to the small samples, in both the cross section and in time?*
- *If this is the case, can we rely on small sample techniques?*

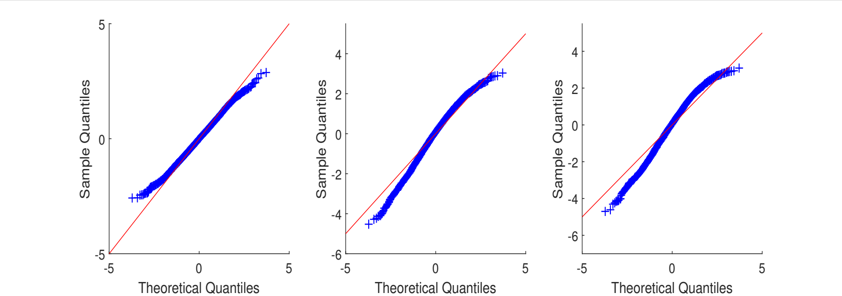
The problem in spatial econometrics terms

Issue

- *Is the use of asymptotics legitimate? Is the asymptotics suffering from size distortion due to the small samples, in both the cross section and in time?*
- *If this is the case, can we rely on small sample techniques?*

We check (MC simulation) for $\hat{\lambda}$, in a SARAR with $n = 24$ and different W_n:

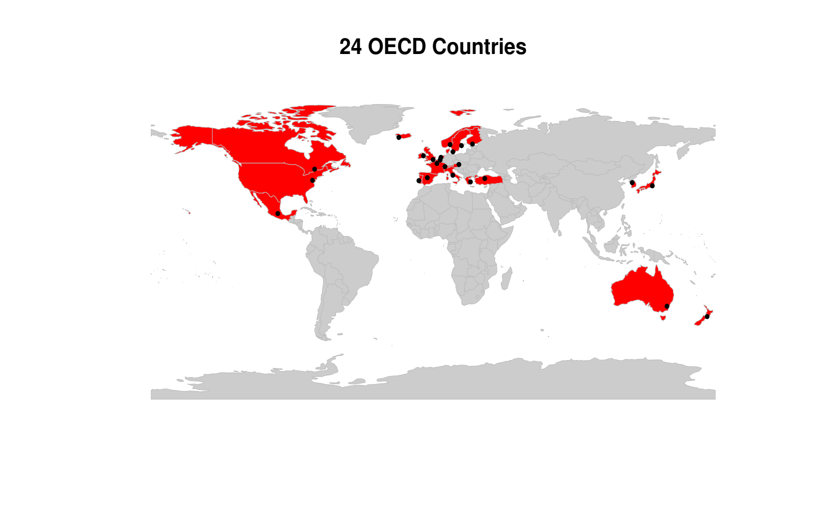
Rook Queen Queen torus



The problem in spatial econometrics terms

The problem can be reformulated in terms of network analysis and use graph theory jargon to describe it...

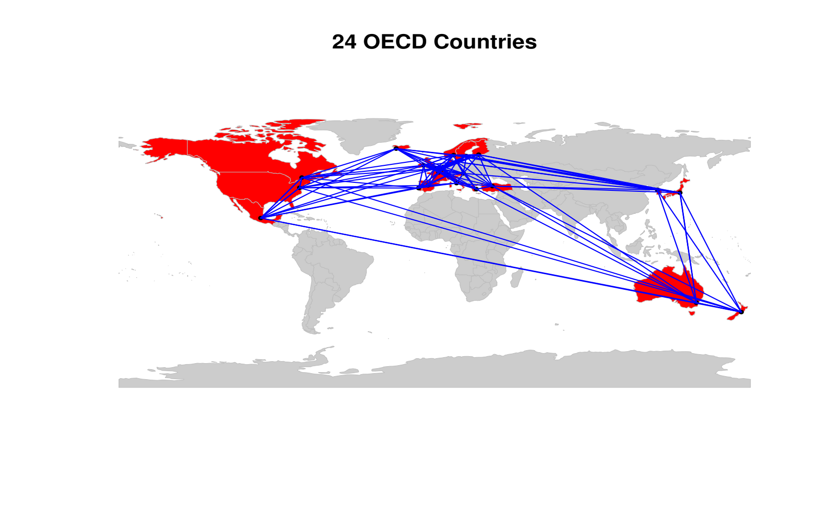
The problem in network analysis terms



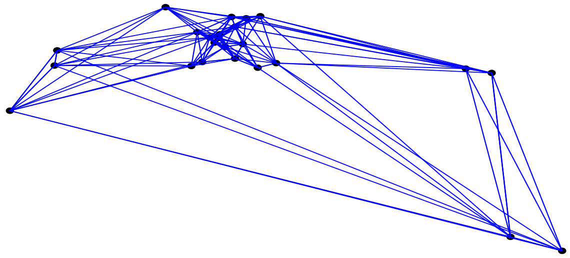
The problem in network analysis terms

On this map we define an order structure by specifying a “neighborhood set” for each observation. This is achieved by W_n, which specifies neighbours (the connectivity, indeed, according to some notion of distance and some rule, e.g. 7NN or inverse geodesic distance) for each capital city.

The problem in network analysis terms



The problem in network analysis terms



The problem in network analysis terms

- W_n is a weighted adjacency matrix. We have a graph $\mathcal{G} = (V, E)$, where V is the set of vertices (capital cities) and E is the set of edges (connections between capital cities). We have $|V| = n$ and W_n is such that if there is an edge $e = (i, j)$ connecting vertices i and j, the entry w_{ij} of W_n represents the weight of the edge—we assume no self-edges $w_{ii} = 0$.

The problem in network analysis terms

- W_n is a weighted adjacency matrix. We have a graph $\mathcal{G} = (V, E)$, where V is the set of vertices (capital cities) and E is the set of edges (connections between capital cities). We have $|V| = n$ and W_n is such that if there is an edge $e = (i, j)$ connecting vertices i and j, the entry w_{ij} of W_n represents the weight of the edge—we assume no self-edges $w_{ii} = 0$.
- we encode the graph structure into the Laplacian: letting D be the degree matrix (a diagonal matrix whose i-th diagonal element d_{ii} is equal to the sum of the weights of all the edges incident to vertex i), one has

$$L_n = D_n - W_n.$$

The problem in network analysis terms

- W_n is a weighted adjacency matrix. We have a graph $\mathcal{G} = (V, E)$, where V is the set of vertices (capital cities) and E is the set of edges (connections between capital cities). We have $|V| = n$ and W_n is such that if there is an edge $e = (i, j)$ connecting vertices i and j, the entry w_{ij} of W_n represents the weight of the edge—we assume no self-edges $w_{ii} = 0$.
- we encode the graph structure into the Laplacian: letting D be the degree matrix (a diagonal matrix whose i-th diagonal element d_{ii} is equal to the sum of the weights of all the edges incident to vertex i), one has

$$L_n = D_n - W_n.$$

- we model a Markovian (in the space) Gaussian random field $\{\mathrm{Inv}_{it}\}$ taking values on a graph, randomness is induced by a Gaussian random field for the disturbances. We add some vertex specific covariates (saving rate of i-th country)

The problem in network analysis terms

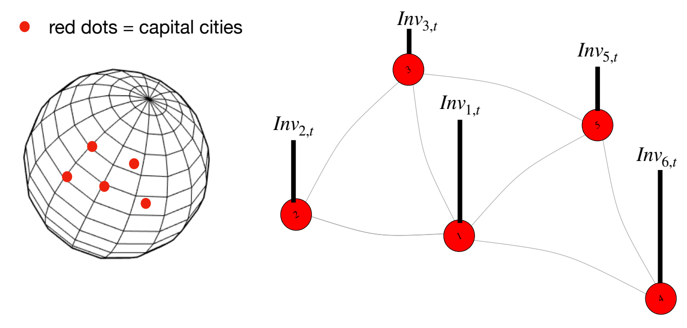
- W_n is a weighted adjacency matrix. We have a graph $\mathcal{G} = (V, E)$, where V is the set of vertices (capital cities) and E is the set of edges (connections between capital cities). We have $|V| = n$ and W_n is such that if there is an edge $e = (i, j)$ connecting vertices i and j, the entry w_{ij} of W_n represents the weight of the edge—we assume no self-edges $w_{ii} = 0$.
- we encode the graph structure into the Laplacian: letting D be the degree matrix (a diagonal matrix whose i-th diagonal element d_{ii} is equal to the sum of the weights of all the edges incident to vertex i), one has

$$L_n = D_n - W_n.$$

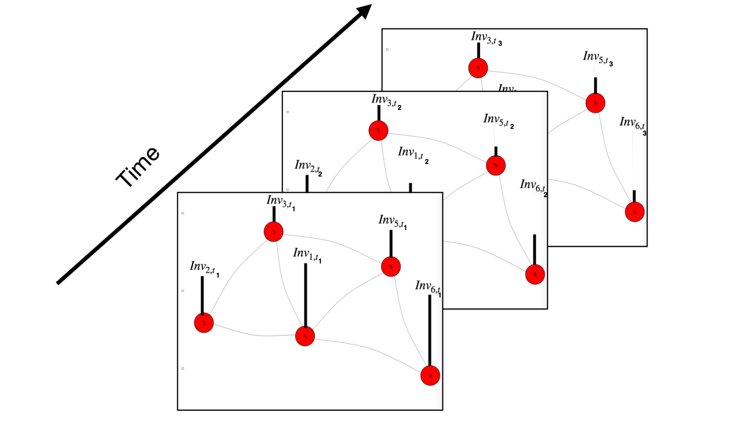
- we model a Markovian (in the space) Gaussian random field $\{\mathrm{Inv}_{it}\}$ taking values on a graph, randomness is induced by a Gaussian random field for the disturbances. We add some vertex specific covariates (saving rate of i-th country)
- we have repeated observations (one for each time point t) of the random field on the graph.

The problem in network analysis terms

- W_n is a weighted adjacency matrix. We have a graph $\mathcal{G} = (V, E)$, where V is the set of vertices (capital cities) and E is the set of edges (connections between capital cities). We have $|V| = n$ and W_n is such that if there is an edge $e = (i, j)$ connecting vertices i and j, the entry w_{ij} of W_n represents the weight of the edge—we assume no self-edges $w_{ii} = 0$.
- we encode the graph structure into the Laplacian: letting D be the degree matrix (a diagonal matrix whose i-th diagonal element d_{ii} is equal to the sum of the weights of all the edges incident to vertex i), one has

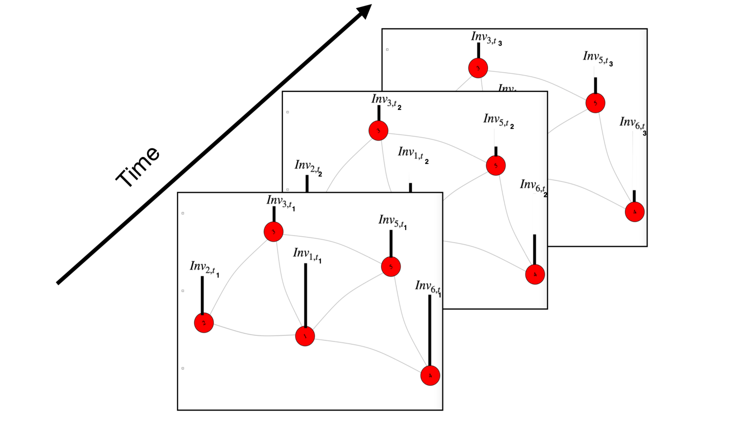


The problem in network analysis terms



Graph Signal Processes (GSP) theory at the intersection between stochastic processes, network science and machine learning: see e.g. Shuman et al, 2013, Dong et al 2020.

The problem in network analysis terms



Graph Signal Processes (GSP) theory at the intersection between stochastic processes, network science and machine learning: see e.g. Shuman et al, 2013, Dong et al 2020.

Testing in small samples

General problem: For a given statistics $\hat{\theta}_{n,T}$, tail probabilities

$$P[\hat{\theta}_{n,T} > x]$$

are needed to carry out statistical inference (essentially, tests and confidence intervals).

Unless the (test) statistic $\hat{\theta}_{n,T}$ has a simple form (e.g. linear) and/or the underlying distribution of data has a particular form (e.g. normal), tail probabilities (more generally the whole distribution) cannot be computed exactly.

Testing in small samples

General problem: For a given statistics $\hat{\theta}_{n,T}$, tail probabilities

$$P[\hat{\theta}_{n,T} > x]$$

are needed to carry out statistical inference (essentially, tests and confidence intervals).

Unless the (test) statistic $\hat{\theta}_{n,T}$ has a simple form (e.g. linear) and/or the underlying distribution of data has a particular form (e.g. normal), tail probabilities (more generally the whole distribution) cannot be computed exactly.

$\Rightarrow$ we have to rely on approximations

Asymptotic theory versus finite sample techniques

We can approximate tail probabilities via

- *Asymptotic theory*: use of Central Limit Theorem to get a Gaussian approximation in large samples,

Asymptotic theory versus finite sample techniques

We can approximate tail probabilities via

- *Asymptotic theory*: use of Central Limit Theorem to get a Gaussian approximation in large samples,
- *Resampling techniques*: use of resampling (bootstrap, subsampling) to get an approximation in small samples,
- *Analytical techniques*: use of expansions (Edgeworth, saddlepoint) to get an approximation in small samples.

Edgeworth expansion

In the iid setting, with sample size n, Edgeworth expansions are approx to the density of a statistic: expansions in powers of $n^{-1/2}$, where the leading term is the normal density—Maclaurin series for the cumulant generating function. Typically, the expansion is computed up to order $O(n^{-1})$ and it entails absolute error.

Edgeworth expansion

In the iid setting, with sample size n, Edgeworth expansions are approx to the density of a statistic: expansions in powers of $n^{-1/2}$, where the leading term is the normal density—Maclaurin series for the cumulant generating function. Typically, the expansion is computed up to order $O(n^{-1})$ and it entails absolute error.

By construction, Edgeworth expansions provide in general a **good approximation at the mean of the density**, **BUT**

Edgeworth expansion

In the iid setting, with sample size n, Edgeworth expansions are approx to the density of a statistic: expansions in powers of $n^{-1/2}$, where the leading term is the normal density—Maclaurin series for the cumulant generating function. Typically, the expansion is computed up to order $O(n^{-1})$ and it entails absolute error.

By construction, Edgeworth expansions provide in general a **good approximation at the mean of the density**, **BUT**

- they can be inaccurate in the tails

Edgeworth expansion

In the iid setting, with sample size n, Edgeworth expansions are approx to the density of a statistic: expansions in powers of $n^{-1/2}$, where the leading term is the normal density—Maclaurin series for the cumulant generating function. Typically, the expansion is computed up to order $O(n^{-1})$ and it entails absolute error.

By construction, Edgeworth expansions provide in general a **good approximation at the mean of the density**, **BUT**

- they can be inaccurate in the tails
- they can even become negative in the tails.

Edgeworth expansion

In the iid setting, with sample size n, Edgeworth expansions are approx to the density of a statistic: expansions in powers of $n^{-1/2}$, where the leading term is the normal density—Maclaurin series for the cumulant generating function. Typically, the expansion is computed up to order $O(n^{-1})$ and it entails absolute error.

By construction, Edgeworth expansions provide in general a **good approximation at the mean of the density**, **BUT**

- they can be inaccurate in the tails
- they can even become negative in the tails.

$\Rightarrow$ To solve these problems, saddlepoint approximations have been introduced.

Saddlepoint density approx for the mean

As in Daniels (1954), let $X_1, \ldots, X_n$ be **iid** copies of $X \sim F$ with $M(\eta) = E[e^{\eta X}]$ being the moment generating function such that

$$\mathcal{K}(\eta) = \log M(\eta)$$

is the cumulant generating function. Assume w.l.o.g. that $E[X] = 0$.

Inferential problem

Approximate, for every n, the density f_n (say) of the

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

which is an M-estimator of location.

Saddlepoint techniques for the mean (cont.)

By standard Fourier inversion, the density f_n, at a point α, is obtained as

$$f_n(\alpha) = \frac{n}{2\pi i} \int_{-i\infty}^{i\infty} \exp\{n(\mathcal{K}(\eta) - \eta\alpha)\} d\eta.$$

Saddlepoint techniques for the mean (cont.)

By standard Fourier inversion, the density f_n, at a point α, is obtained as

$$f_n(\alpha) = \frac{n}{2\pi i} \int_{-i\infty}^{i\infty} \exp\{n(\mathcal{K}(\eta) - \eta\alpha)\} d\eta.$$

The integral is typically not available in a closed form and an approximation is obtained by the method of the steepest descent, which yields the saddlepoint approx $p_n(\alpha)$:

$$f_n(\alpha) = p_n(\alpha) \left\{1 + O\left(n^{-1}\right)\right\},$$

$$p_n(\alpha) = \left[\frac{n}{2\pi \mathcal{K}^{('')}\{\eta(\alpha)\}}\right]^{1/2} \exp\left(n\left[\mathcal{K}\{\eta(\alpha)\} - \eta(\alpha)\alpha\right]\right),$$

and $\eta(\alpha)$ (saddlepoint) is the solution to $\boxed{\mathcal{K}^{(')}(\eta) - \alpha = 0}$.

Remark

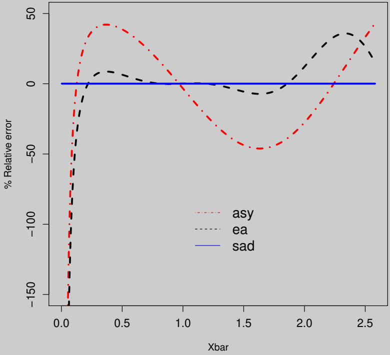
The approximation p_n features relative error of order $O(n^{-1})$ over the whole $\mathbb{R}$

Example (Mean of exp(1) rvs)

Density approx for $\bar{X}$ when $n = 5$, with $X_i \sim \exp(1)$?

Example (Mean of exp(1) rvs)

Density approx for $\bar{X}$ when $n = 5$, with $X_i \sim \exp(1)$? We compare the relative error implied by the approx obtained by: asymptotic theory (asy); **Edg** (ea); Saddlepoint (sad), as derived by tilted Edg.



Spatial Autoregressive models with covariates (SARAR)

Q. How about MLE for the SARAR model?

Spatial Autoregressive models with covariates (SARAR)

Q. How about MLE for the SARAR model?

For user-specified spatial matrices W_n and M_n, let's re-consider the model (1) in more general terms:

$$
\begin{aligned}
Y_{nt} &= \lambda_0 W_n Y_{nt} + X_{nt}\beta_0 + c_{n0} + E_{nt}, \\
E_{nt} &= \rho_0 M_n E_{nt} + V_{nt}, \quad t = 1, ..., T.
\end{aligned}
$$

where $Y_{nt} = (y_{1t}, y_{2t}, ..., y_{nt})$, X_{nt} is an matrix of non stochastic time-varying regressors, c_{n0} is an $n \times 1$ vector of fixed effects and $V_{nt} = (v_{1t}, v_{2t}, .., v_{nt})'$ are $n \times 1$ vectors and $v_{it} \sim \mathcal{N}(0, \sigma_0^2)$, i.i.d. across i and t (Lee and Yu (JoE, 2010)).

The model parameter is

$$\theta_0 := (\lambda_0, \beta_0, \rho_0, \sigma_0^2)$$

and we estimate it using the Gaussian MLE ($\hat{\theta}_{n,T}$).

SARAR (first-order asymptotics)

Let's set $m = n(T-1)$.

SARAR (first-order asymptotics)

Let's set $m = n(T-1)$. Moreover, let us define $S_n(\lambda) = I_n - \lambda W_n$, and analogously $R_n(\rho) = I_n - \rho M_n$. Then, we introduce

Assumption A. (Lee & Yu)

(i) *The elements ω_{ij} of W_n and the elements m_{ij} of M_n are at most of order $\tilde{h}_n^{-1}$, denoted by $O(1/\tilde{h}_n)$, uniformly in all i,j, where the rate sequence $\{\tilde{h}_n\}$ is bounded, and $\tilde{h}_n$ is bounded away from zero for all n. As a normalization, we have $\omega_{ii} = m_{ii} = 0$, for all i.*

(ii) *n diverges, while $T \geq 2$ and it is finite.*

(iii) *Denote $C_n = \ddot{G}_n - n^{-1}tr(\ddot{G}_n)I_n$ and $D_n = H_n - n^{-1}tr(H_n)I_n$ where*

$$\ddot{G}_n = R_n G_n R_n^{-1} \text{ and } H_n = M_n R_n^{-1}.$$

Then $C_n^s = C_n + C_n'$ and $D_n^s = D_n + D_n'$. The limit of $n^{-2}\left[tr(C_n^s C_n^s)tr(D_n^s D_n^s) - tr^2(C_n^s D_n^s)\right]$ is strictly positive as $n \to \infty$.

SARAR (first-order asymptotics)

Let $\hat{\theta}_{n,T}$ be the Gaussian MLE. Under Assumption A(i)-A(iv), Theorem 1 part(ii) in Lee & Yu shows that

$$\lim_{n\to\infty} \hat{\theta}_{n,T} = \theta_0.$$

SARAR (first-order asymptotics)

Let $\hat{\theta}_{n,T}$ be the Gaussian MLE. Under Assumption A(i)-A(iv), Theorem 1 part(ii) in Lee & Yu shows that

$$\lim_{n\to\infty} \hat{\theta}_{n,T} = \theta_0.$$

Furthermore, as $n \to \infty$, the MLE $\hat{\theta}_{n,T}$ satisfies

$$\sqrt{m}(\hat{\theta}_{n,T} - \theta_0) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \Sigma_{0,T}^{-1}\right),$$

and $\Sigma_{0,T} = \text{plim}_{n\to\infty}\Sigma_{0,n,T}$. The expression of $\Sigma_{0,n,T}$ is available in Lee & Yu.

Our contribution: Main theoretical result

We derive the saddlepoint density approximation for $q(\hat{\theta}_{n,T})$ at point α, where

$$q : \Theta \to \mathbb{R}$$

and $\theta \in \Theta \subset \mathbb{R}^p, p \geq 1$:

$$p_{n,T}(\alpha) = \left[\frac{n}{2\pi \boxed{\tilde{\mathcal{K}}^{('')}_{n,T}}(\eta)} \right]^{1/2} \exp\left\{ n \left[\boxed{\tilde{\mathcal{K}}_{n,T}}(\nu) - \eta\alpha \right] \right\},$$

with, under additional assumptions, relative error of order $O(m^{-1})$, where $\eta := \eta_\alpha$ is the saddlepoint defined by

$$\boxed{\boxed{\tilde{\mathcal{K}}^{(')}_{n,T}}(\eta) - \alpha = 0.}$$

See paper for the explicit expression of $\boxed{\tilde{\mathcal{K}}_{n,T}}$ (approximate c.g.f. of $q(\hat{\theta}_{n,T})$).

The construction at a glance

Idea

Our construction relies on five steps:

The construction at a glance

Idea

Our construction relies on five steps:

(i) *eliminating the fixed effect without creating dependence—to this end, we adopt the transformation approach of Lee & Yu and derive the Gaussian MLE* Jump to MLE

The construction at a glance

Idea

Our construction relies on five steps:

(i) *eliminating the fixed effect without creating dependence—to this end, we adopt the transformation approach of Lee & Yu and derive the Gaussian MLE* Jump to MLE

(ii) *defining a second-order asymptotic (von Mises) expansion for the MLE of θ and using it to define an asymptotic expansion for $q(\hat{\theta}_{n,T})$*

The construction at a glance

Idea

Our construction relies on five steps:

(i) *eliminating the fixed effect without creating dependence—to this end, we adopt the transformation approach of Lee & Yu and derive the Gaussian MLE* Jump to MLE

(ii) *defining a second-order asymptotic (von Mises) expansion for the MLE of θ and using it to define an asymptotic expansion for $q(\hat{\theta}_{n,T})$*

(iii) *identifying the corresponding U-statistic, whose cumulants define* $\tilde{\mathcal{K}}_{n,T}$

The construction at a glance

Idea

Our construction relies on five steps:

(i) *eliminating the fixed effect without creating dependence—to this end, we adopt the transformation approach of Lee & Yu and derive the Gaussian MLE* Jump to MLE

(ii) *defining a second-order asymptotic (von Mises) expansion for the MLE of θ and using it to define an asymptotic expansion for $q(\hat{\theta}_{n,T})$*

(iii) *identifying the corresponding U-statistic, whose cumulants define* $\tilde{\mathcal{K}}_{n,T}$

(iv) *using (iii) to derive an Edgeworth expansion (Bickel et al, 1986)* Jump to HOA

The construction at a glance

Idea

Our construction relies on five steps:

(i) *eliminating the fixed effect without creating dependence—to this end, we adopt the transformation approach of Lee & Yu and derive the Gaussian MLE* Jump to MLE

(ii) *defining a second-order asymptotic (von Mises) expansion for the MLE of θ and using it to define an asymptotic expansion for $q(\hat{\theta}_{n,T})$*

(iii) *identifying the corresponding U-statistic, whose cumulants define* $\tilde{\mathcal{K}}_{n,T}$

(iv) *using (iii) to derive an Edgeworth expansion (Bickel et al, 1986)* Jump to HOA

(v) *defining the tilted version (tilted-Edgeworth) we obtain the saddlepoint density (Gatto & Ronchetti, 1996). The key tool is the* Esscher's tilting.

The construction at a glance

Idea

Our construction relies on five steps:

(i) *eliminating the fixed effect without creating dependence—to this end, we adopt the transformation approach of Lee & Yu and derive the Gaussian MLE* Jump to MLE

(ii) *defining a second-order asymptotic (von Mises) expansion for the MLE of θ and using it to define an asymptotic expansion for $q(\hat{\theta}_{n,T})$*

(iii) *identifying the corresponding U-statistic, whose cumulants define* $\tilde{\mathcal{K}}_{n,T}$

(iv) *using (iii) to derive an Edgeworth expansion (Bickel et al, 1986)* Jump to HOA

(v)


Home ▶ All Journals ▶ Journal of the American Statistical Association ▶ List of Issues ▶ Latest Articles ▶ Saddlepoint Approximations for Spatial P ....

Journal of the American Statistical Association ›
Latest Articles
Submit an article | Journal homepage

Enter keywords, authors, DOI, ORCID etc | This Journal
Advanced search

312 Views
0 CrossRef citations to date
0 Altmetric

Listen

Full access

Theory and Methods

Saddlepoint Approximations for Spatial Panel Data Models

Chaonan Jiang, Davide La Vecchia, Elvezio Ronchetti & Olivier Scaillet

The construction at a glance

Idea

Intuition for tilted-Edgeworth (BNS & Cox, 89, Ch 4): let us consider again $\bar{X}$ in the iid setting, with $E[X] = 0$ and X real-valued.

The construction at a glance

Idea

Intuition for tilted-Edgeworth (BNS & Cox, 89, Ch 4): let us consider again $\bar{X}$ in the iid setting, with $E[X] = 0$ and X real-valued. The Edgeworth expansion for the pdf at $\alpha \in \mathbb{R}$ is

$$p^{Edg}(\alpha) = \phi(\alpha)\left\{1 + \frac{1}{\sqrt{n}}\kappa_3(\alpha^3 - 3\alpha) + O(n^{-1})\right\},$$

The construction at a glance

Idea

Intuition for tilted-Edgeworth (BNS & Cox, 89, Ch 4): let us consider again $\bar{X}$ in the iid setting, with $E[X]=0$ and X real-valued. The Edgeworth expansion for the pdf at $\alpha \in \mathbb{R}$ is

$$p^{Edg}(\alpha) = \phi(\alpha)\left\{1 + \frac{1}{\sqrt{n}}\kappa_3(\alpha^3 - 3\alpha) + O(n^{-1})\right\},$$

Now, let us compute this approx at the zero, namely $\exists \eta(\alpha)$ such that

$$\alpha = 0 = E[X] = \mathcal{K}^{(\prime)}(\eta(\alpha)) \Rightarrow \boxed{\mathcal{K}^{(\prime)}(\eta(\alpha)) = 0}$$

and notice that the term of order $n^{-1/2}$ disappears: approx with relative error $O(n^{-1})$.

Let's exploit it! For each α, we perform an Esscher's tilting, we move the probability mass of the underling distribution in such a way that each X_i has expectation equal to α under the tilted distribution.

The construction at a glance

Idea

Intuition for tilted-Edgeworth (BNS & Cox, 89, Ch 4): let us consider again $\bar{X}$ in the iid setting, with $E[X]=0$ and X real-valued. The Edgeworth expansion for the pdf at $\alpha \in \mathbb{R}$ is

$$p^{Edg}(\alpha) = \phi(\alpha)\left\{1 + \frac{1}{\sqrt{n}}\kappa_3(\alpha^3 - 3\alpha) + O(n^{-1})\right\},$$

Now, let us compute this approx at the zero, namely $\exists \eta(\alpha)$ such that

$$\alpha = 0 = E[X] = \mathcal{K}^{(\prime)}(\eta(\alpha)) \Rightarrow \boxed{\mathcal{K}^{(\prime)}(\eta(\alpha)) = 0}$$

and notice that the term of order $n^{-1/2}$ disappears: approx with relative error $O(n^{-1})$.

Let's exploit it! For each α, we perform an Esscher's tilting, we move the probability mass of the underling distribution in such a way that each X_i has expectation equal to α under the tilted distribution. In this way, we compute an accurate Edgeworth!

The construction at a glance

Idea

Intuition for tilted-Edgeworth (BNS & Cox, 89, Ch 4): let us consider again $\bar{X}$ in the iid setting, with $E[X]=0$ and X real-valued. The Edgeworth expansion for the pdf at $\alpha \in \mathbb{R}$ is

$$p^{Edg}(\alpha) = \phi(\alpha)\left\{1 + \frac{1}{\sqrt{n}}\kappa_3(\alpha^3 - 3\alpha) + O(n^{-1})\right\},$$

Now, let us compute this approx at the zero, namely $\exists \eta(\alpha)$ such that

$$\alpha = 0 = E[X] = \mathcal{K}^{(\prime)}(\eta(\alpha)) \Rightarrow \boxed{\mathcal{K}^{(\prime)}(\eta(\alpha)) = 0}$$

and notice that the term of order $n^{-1/2}$ disappears: approx with relative error $O(n^{-1})$.

Let's exploit it! For each α, we perform an Esscher's tilting, we move the probability mass of the underling distribution in such a way that each X_i has expectation equal to α under the tilted distribution. In this way, we compute an accurate Edgeworth! The resulting **tilted-Edgeworth** is the **saddlepoint density approx.**

The construction at a glance

Idea

Intuition for tilted-Edgeworth (BNS & Cox, 89, Ch 4): let us consider again $\bar{X}$ in the iid setting, with $E[X] = 0$ and X real-valued. The Edgeworth expansion for the pdf at $\alpha \in \mathbb{R}$ is

$$p^{Edg}(\alpha) = \phi(\alpha)\left\{1 + \frac{1}{\sqrt{n}}\kappa_3(\alpha^3 - 3\alpha) + O(n^{-1})\right\},$$

Now, let us compute this approx at the zero, namely $\exists\eta(\alpha)$ such that

$$\alpha = 0 = E[X] = \mathcal{K}^{(\prime)}(\eta(\alpha)) \Rightarrow \boxed{\mathcal{K}^{(\prime)}(\eta(\alpha)) = 0}$$

and notice that the term of order $n^{-1/2}$ disappears: approx with relative error $O(n^{-1})$.

Let's exploit it! For each α, we perform an Esscher's tilting, we move the probability mass of the underling distribution in such a way that each X_i has expectation equal to α under the tilted distribution. In this way, we compute an accurate Edgeworth! The resulting **tilted-Edgeworth** is the **saddlepoint density approx.**

We apply this idea to approximate the distribution of $q(\hat{\theta}_{n,T})$.

The problem in optimal transportation terms

The Esscher's tilting can be reformulated in terms of optimal transportation theory...

Measure transportation

Looking at the issue of finding the best way to move given piles of sand to fill up given holes of the same total volume, **Gaspard Monge** (1746-1818) formulated a **mathematical problem** that in modern jargon reads as:

Measure transportation

Looking at the issue of finding the best way to move given piles of sand to fill up given holes of the same total volume, **Gaspard Monge** (1746-1818) formulated a **mathematical problem** that in modern jargon reads as:

Let μ *and* ν *denote two probability measures over (for simplicity)* $(\mathbb{R}^d, \mathcal{B}^d)$*, for* $d \geq 1$*. Let* $c : \mathbb{R}^{2d} \to \mathbb{R}$ *be a Borel-measurable cost function such that* $c(\mathrm{x}, \mathrm{y})$ *represents the cost of transporting* x *to* y*. Then, find a measurable transport map* $\mathcal{T} : \mathbb{R}^d \to \mathbb{R}^d$ *that achieves*

$$\inf_{\mathcal{T} \in M} \int_{\mathbb{R}^d} c[\mathrm{x}, \mathcal{T}(\mathrm{x})] \mathrm{d}\mu \tag{2}$$

where

$$M := \{\mathcal{T} : \mathrm{X} \to \mathrm{Y}\},$$

with $\mathrm{X} \sim \mu$*,* $\mathrm{Y} \sim \nu$*. The map* $\mathcal{T} \# \mu = \nu$ *does the push forward of* μ *to* ν*.*

Measure transportation

Looking at the issue of finding the best way to move given piles of sand to fill up given holes of the same total volume, **Gaspard Monge** (1746-1818) formulated a **mathematical problem** that in modern jargon reads as:

Let μ and ν denote two probability measures over (for simplicity) $(\mathbb{R}^d, \mathcal{B}^d)$, for $d \geq 1$. Let $c : \mathbb{R}^{2d} \to \mathbb{R}$ be a Borel-measurable cost function such that $c(\mathrm{x}, \mathrm{y})$ represents the cost of transporting x to y. Then, find a measurable transport map $\mathcal{T} : \mathbb{R}^d \to \mathbb{R}^d$ that achieves

$$\inf_{\mathcal{T} \in M} \int_{\mathbb{R}^d} c[\mathrm{x}, \mathcal{T}(\mathrm{x})] \mathrm{d}\mu \tag{2}$$

where

$$M := \{\mathcal{T} : \mathrm{X} \to \mathrm{Y}\},$$

with $\mathrm{X} \sim \mu$, $\mathrm{Y} \sim \nu$. The map $\mathcal{T} \# \mu = \nu$ does the push forward of μ to ν.

$\Rightarrow$ The map solution to (2) is called the optimal transportation map.

Measure transportation

Monge's problem remained open until the 1940s, when it was revisited by **Leonid Vitaliyevitch Kantorovich** (1912-1986; Nobel Prize in Economics in 1975) for the economic problem of optimal allocation of resources; see e.g. Villani (2008), Santambrogio (2015), Galichon (2016).

Measure transportation

Monge's problem remained open until the 1940s, when it was revisited by **Leonid Vitaliyevitch Kantorovich** (1912-1986; Nobel Prize in Economics in 1975) for the economic problem of optimal allocation of resources; see e.g. Villani (2008), Santambrogio (2015), Galichon (2016).

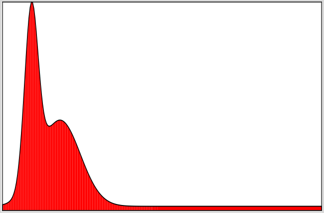
In the Kantorovich primal problem, the objective is to find the optimal transportation plan γ, which solves

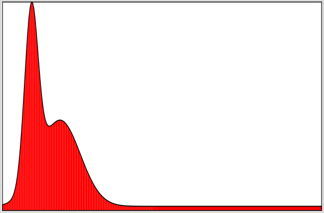
$$\inf_{\gamma \in \Gamma(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(\mathrm{x}, \mathrm{y}) \mathrm{d}\gamma(\mathrm{x}, \mathrm{y}), \tag{3}$$

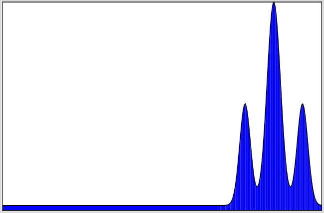
where the infimum is over all coupling $(\mathbf{X}, \mathbf{Y})$ of (μ, ν), belonging to $\Gamma(\mu, \nu)$, the set of probability measures γ on $\mathbb{R}^d \times \mathbb{R}^d$, satisfying

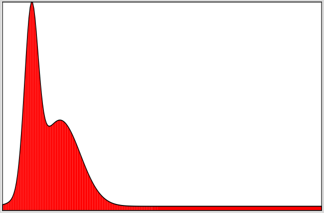
$$\gamma(A \times \mathbb{R}^d) = \mu(A) \text{ and } \gamma(\mathbb{R}^d \times B) = \nu(B),$$

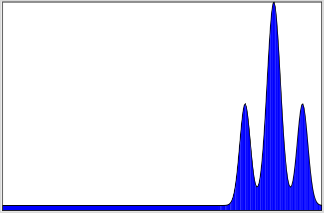
for measurable sets $A, B \subset \mathbb{R}^d$.

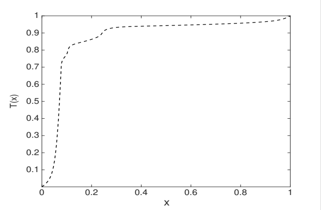


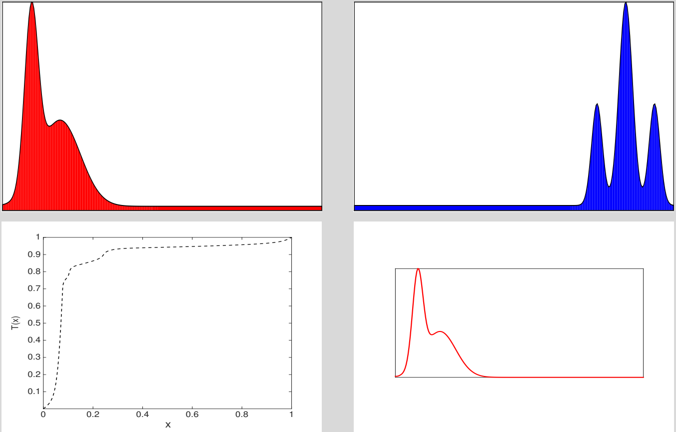


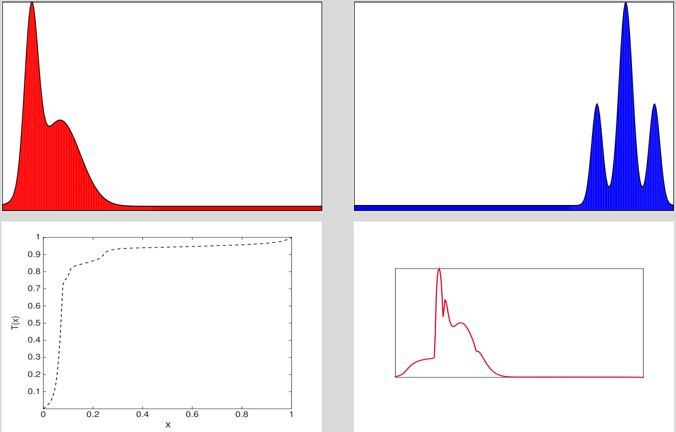


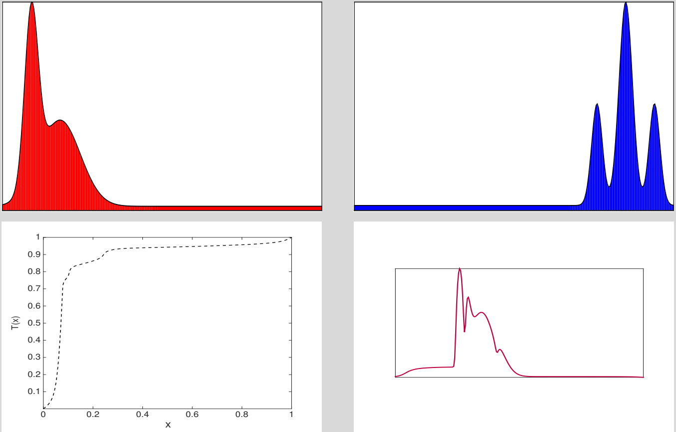


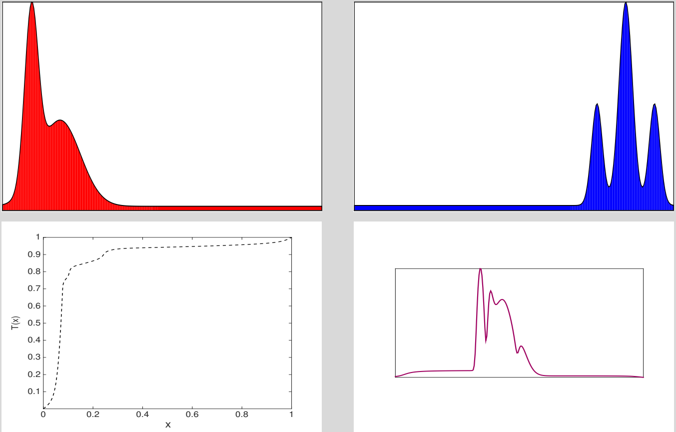


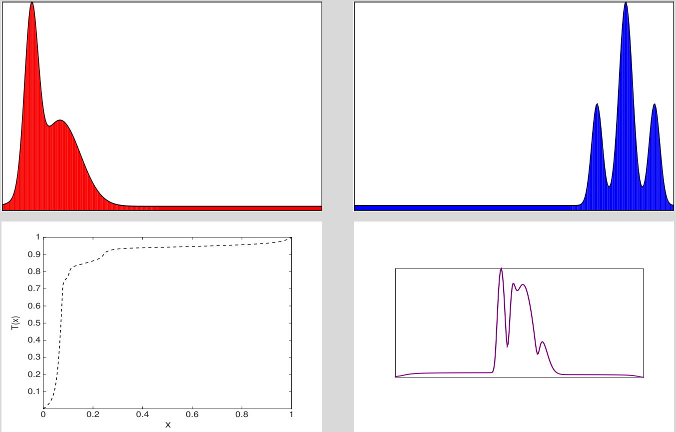

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
T(x)
0
0.2
0.4
0.6
0.8
1
x

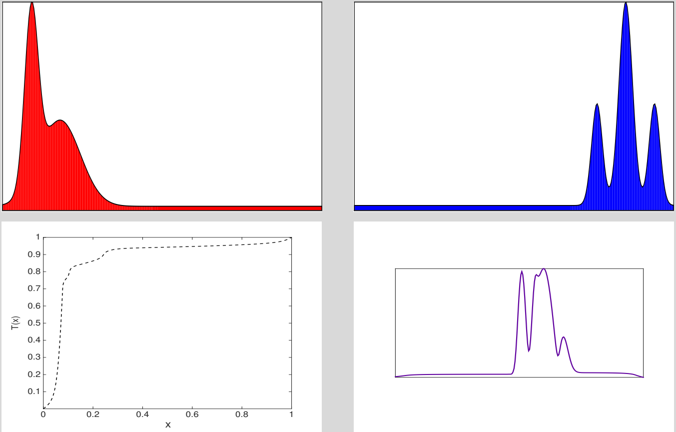

T(x)
x
0
0.2
0.4
0.6
0.8
1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

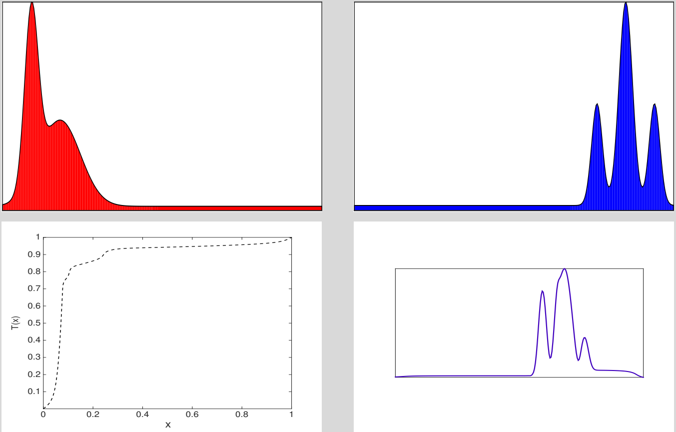

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
T(x)
0
0.2
0.4
0.6
0.8
1
x

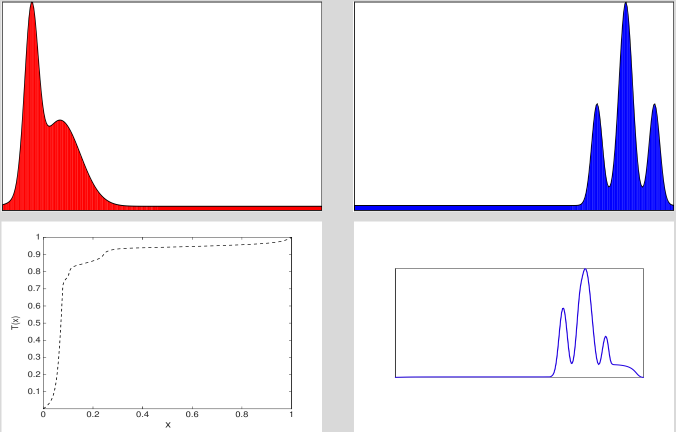

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
T(x)
0
0.2
0.4
0.6
0.8
1
x

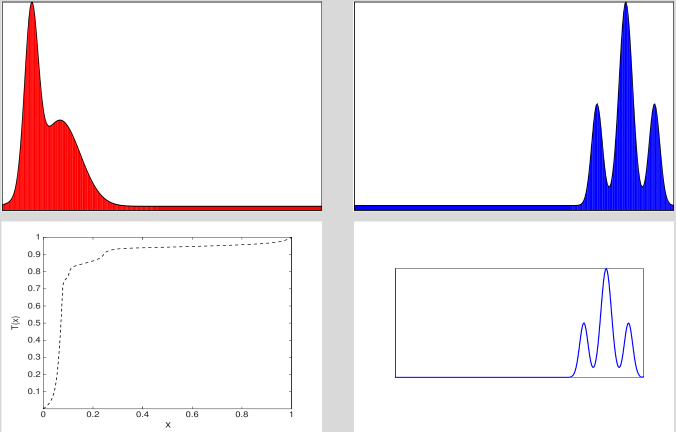

T(x)
x
0
0.2
0.4
0.6
0.8
1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


T(x)
x
0
0.2
0.4
0.6
0.8
1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


T(x)
x
0
0.2
0.4
0.6
0.8
1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


T(x)
x
0
0.2
0.4
0.6
0.8
1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9


T(x)
x
0
0.2
0.4
0.6
0.8
1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9


T(x)
x
0
0.2
0.4
0.6
0.8
1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Let us consider again $\bar{X}$, for $X \in \mathbb{R}$. For $\alpha \in \mathbb{R}$, let us consider two real-valued rv:

$$X \sim \mu(-\infty, x] = F(x) \quad \text{and} \quad Y_\alpha \sim \nu_\alpha(-\infty, x] = H_\alpha(y).$$

Then, the considered transformations are such that

- Esscher's tilting: for every $\alpha \in \mathbb{R}$,

$$\mu \mapsto \nu_\alpha \Rightarrow E_\mu[X] = 0 \quad E_{\nu_\alpha}[Y_\alpha] = \alpha$$

Let us consider again $\bar{X}$, for $X \in \mathbb{R}$. For $\alpha \in \mathbb{R}$, let us consider two real-valued rv:

$$X \sim \mu(-\infty, x] = F(x) \quad \text{and} \quad Y_\alpha \sim \nu_\alpha(-\infty, x] = H_\alpha(y).$$

Then, the considered transformations are such that

- Esscher's tilting: for every $\alpha \in \mathbb{R}$,

$$\mu \mapsto \nu_\alpha \Rightarrow E_\mu[X] = 0 \quad E_{\nu_\alpha}[Y_\alpha] = \alpha$$

- Optimal transportation: for every $\alpha \in \mathbb{R}$

$$X \mapsto Y_\alpha \Rightarrow E_\mu[X] = 0 \quad E_{\nu_\alpha}[Y_\alpha] = \alpha$$

where $\mathcal{T} = H_\alpha^{-1} \circ F$.

Let us consider again $\bar{X}$, for $X \in \mathbb{R}$. For $\alpha \in \mathbb{R}$, let us consider two real-valued rv:

$$X \sim \mu(-\infty, x] = F(x) \quad \text{and} \quad Y_\alpha \sim \nu_\alpha(-\infty, x] = H_\alpha(y).$$

Then, the considered transformations are such that

- Esscher's tilting: for every $\alpha \in \mathbb{R}$,

$$\mu \mapsto \nu_\alpha \Rightarrow E_\mu[X] = 0 \quad E_{\nu_\alpha}[Y_\alpha] = \alpha$$

- Optimal transportation: for every $\alpha \in \mathbb{R}$

$$X \mapsto Y_\alpha \Rightarrow E_\mu[X] = 0 \quad E_{\nu_\alpha}[Y_\alpha] = \alpha$$

where $\mathcal{T} = H_\alpha^{-1} \circ F$.

Connection: **Esscher's tilting** is a form of **mass transportation** from μ to ν_α: This is related to the **optimal transportation**. In the case of $\bar{X}$, the push-forward $\boxed{\mathcal{T}}$ depends on the Legendre transform of $\boxed{\mathcal{K}}$. Details available in Proposition 2.1 in La Vecchia et al. (2022).

Let us consider again $\bar{X}$, for $X \in \mathbb{R}$. For $\alpha \in \mathbb{R}$, let us consider two real-valued rv:

$$X \sim \mu(-\infty, x] = F(x) \quad \text{and} \quad Y_\alpha \sim \nu_\alpha(-\infty, x] = H_\alpha(y).$$

Then, the considered transformations are such that

- Esscher's tilting: for every $\alpha \in \mathbb{R}$,

$$\mu \mapsto \nu_\alpha \Rightarrow E_\mu[X] = 0 \quad E_{\nu_\alpha}[Y_\alpha] = \alpha$$

- Optimal transportation: for every $\alpha \in \mathbb{R}$

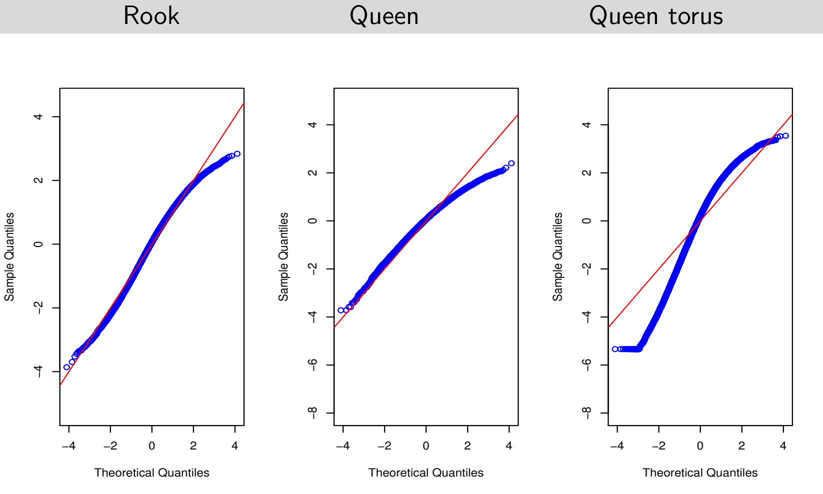
$$X \mapsto Y_\alpha \Rightarrow E_\mu[X] = 0 \quad E_{\nu_\alpha}[Y_\alpha] = \alpha$$

where $\mathcal{T} = H_\alpha^{-1} \circ F$.

Connection: **Esscher's tilting** is a form of **mass transportation** from μ to ν_α: This is related to the **optimal transportation**. In the case of $\bar{X}$, the push-forward $\mathcal{T}$ depends on the Legendre transform of $\mathcal{K}$. Details available in Proposition 2.1 in La Vecchia et al. (2022).

Comparison with asymptotic theory

SAR(1): distribution of $\hat{\lambda}$, for $n = 24$, $T = 2$ and $\lambda_0 = 0.2$



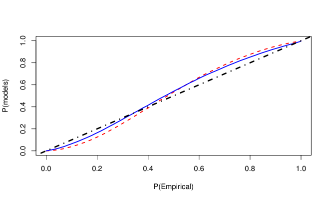
Comparison with asymptotic theory (cont'd)

Is the saddlepoint density approximation doing any better?

Comparison with asymptotic theory (cont'd)

Is the saddlepoint density approximation doing any better?

Rook

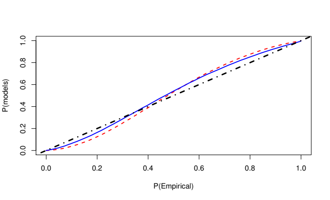


Queen

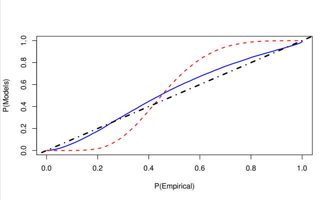
Comparison with asymptotic theory (cont'd)

Is the saddlepoint density approximation doing any better?

Rook



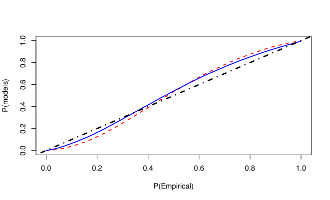
Queen



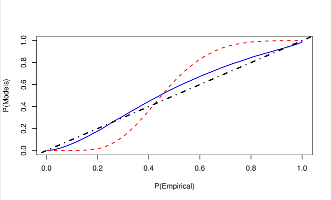
Comparison with asymptotic theory (cont'd)

Is the saddlepoint density approximation doing any better?

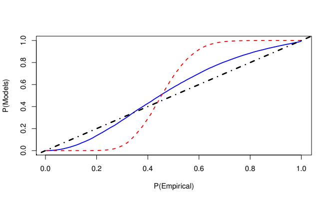
Rook



Queen



Queen torus

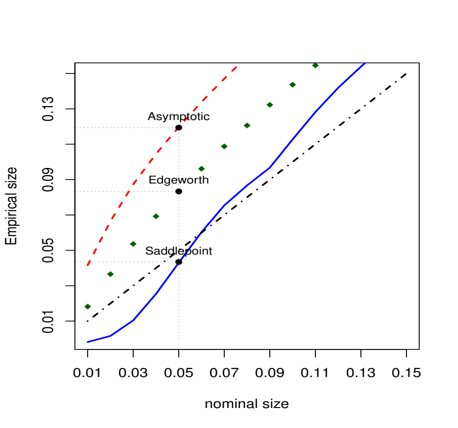


"All together now" ($n = 24$ and W_n is Rook)

We study the level of the test $\mathcal{H}_0 : \lambda_0 = 0$ vs $\mathcal{H}_1 : \lambda_0 > 0$.

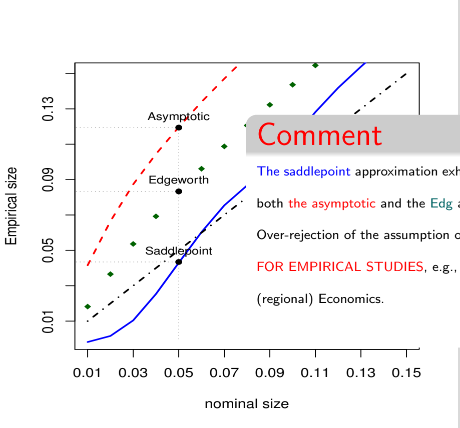
"All together now" ($n = 24$ and W_n is Rook)

We study the level of the test $\mathcal{H}_0 : \lambda_0 = 0$ vs $\mathcal{H}_1 : \lambda_0 > 0$.



"All together now" ($n = 24$ and W_n is Rook)

We study the level of the test $\mathcal{H}_0 : \lambda_0 = 0$ vs $\mathcal{H}_1 : \lambda_0 > 0$.

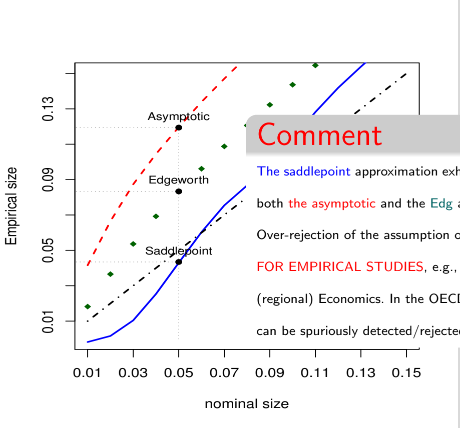


Comment

The saddlepoint approximation exhibits only mild size distortions, whilst both the asymptotic and the Edg approx display **severe size distortion**. Over-rejection of the assumption of no spatial correlation $\Rightarrow$ ALARM FOR EMPIRICAL STUDIES, e.g., in (applied) Econometrics and/or (regional) Economics.

“All together now” ($n = 24$ and W_n is Rook)

We study the level of the test $\mathcal{H}_0 : \lambda_0 = 0$ vs $\mathcal{H}_1 : \lambda_0 > 0$.



Comment

The saddlepoint approximation exhibits only mild size distortions, whilst both the asymptotic and the Edg approx display **severe size distortion**. Over-rejection of the assumption of no spatial correlation $\Rightarrow$ ALARM FOR EMPIRICAL STUDIES, e.g., in (applied) Econometrics and/or (regional) Economics. In the OECD example, spatial autocorrelation can be spuriously detected/rejected by the asymptotic theory!

Test results for the OCDE countries

Report the p-values (testing in presence of **nuisance parameters**, Robinson et al. (2003) and see La Vecchia et al. (2022)) of Saddlepoint (SAD) and first-order asymptotic (ASY) approximation.

		Inverse distance			7 nearest neighbours		
		1960-1970	1971-1985	1986-2000	1960-1970	1971-1985	1986-2000
$\lambda_0 = 0$	*SAD*	1.0000	0.0096	**0.0000**	0.9998	0.2248	**0.0000**
	ASY	1.0000	0.0116	0.5679	0.9987	0.0130	0.1123
$\rho_0 = 0$	*SAD*	0.1134	0.0024	0.1217	0.3232	**0.0403**	0.9993
	ASY	0.5890	0.2261	0.9578	0.7101	0.3898	0.9998
$\lambda_0 = \rho_0 = 0$	*SAD*	0.1414	0.0000	0.0000	0.2603	0.0000	0.0000
	ASY	0.4615	0.0000	0.0000	0.5042	0.0000	0.0000

Test results for the OCDE countries

- In the sub-period 86-00, there are **large differences between *p*-values under the two approximations** when testing $\lambda_0 = 0$. We find that there is no spatial dependence of investing rates across countries for that period, and vice-versa for the asymptotic approximation. This is in line with the overrejection of the ASY that we find in the Monte Carlo experiments for λ.

Test results for the OCDE countries

- In the sub-period 86-00, there are **large differences between p-values under the two approximations** when testing $\lambda_0 = 0$. We find that there is no spatial dependence of investing rates across countries for that period, and vice-versa for the asymptotic approximation. This is in line with the overrejection of the ASY that we find in the Monte Carlo experiments for λ.

- In the sub-period 71-85 under 7NN W_n, the ASY test does not find evidence against $\rho_0 = 0$, while the *SAD* test rejects this composite hypothesis. Thus, the *SAD* test indicates a spillover through the contemporary shocks between countries. *This spillover goes through the innovations, i.e., through the unexpected part of the model dynamics, a finding not documentable when one relies on the first-order asymptotic theory*.

Take home message

- First-order asymptotics and Edgeworth expansions may deliver poor inference

Take home message

- First-order asymptotics and Edgeworth expansions may deliver poor inference
- Saddlepoint techniques are more accurate and faster (no resampling).

Jump to Parametric Bootstrap

Take home message

- First-order asymptotics and Edgeworth expansions may deliver poor inference
- Saddlepoint techniques are more accurate and faster (no resampling).
  Jump to Parametric Bootstrap
- (Interesting) Connections with the statistical analysis of network data and optimal transportation theory deserve further investigation

Take home message

- First-order asymptotics and Edgeworth expansions may deliver poor inference
- Saddlepoint techniques are more accurate and faster (no resampling).
  Jump to Parametric Bootstrap
- (Interesting) Connections with the statistical analysis of network data and optimal transportation theory deserve further investigation

... and next ...

Take home message

- First-order asymptotics and Edgeworth expansions may deliver poor inference
- Saddlepoint techniques are more accurate and faster (no resampling).
  Jump to Parametric Bootstrap
- (Interesting) Connections with the statistical analysis of network data and optimal transportation theory deserve further investigation

... and next ...

Locally stationary random fields over a network deserve further investigation

Jump to tvSAR

Thank you

For questions: `davide.lavecchia@unige.ch`

Gaussian MLE

The log-likelihood reads as:

$$\begin{aligned} \ell_{n,T}(\theta) &= -\frac{n(T-1)}{2}\ln(2\pi\sigma^2) + (T-1)[\ln|S_n(\lambda)| \\ &+ \ln|R_n(\rho)|] - \frac{1}{2\sigma^2}\sum_{t=1}^{T} \tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta), \end{aligned}$$

where the **transformed errors**

$$\tilde{V}_{nt}(\zeta) = R_n(\rho)[S_n(\lambda)\tilde{Y}_{nt} - \tilde{X}_{nt}\beta],$$

Gaussian MLE

The log-likelihood reads as:

$$\begin{aligned}\ell_{n,T}(\theta) &= -\frac{n(T-1)}{2}\ln(2\pi\sigma^2) + (T-1)[\ln|S_n(\lambda)| \\ &+ \ln|R_n(\rho)|] - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta),\end{aligned}$$

where the **transformed errors**

$$\tilde{V}_{nt}(\zeta) = R_n(\rho)[S_n(\lambda)\tilde{Y}_{nt} - \tilde{X}_{nt}\beta],$$

which are iid Gaussian and are obtained using the transformed variables and transformed covariates (essentially, the transformation is based the time-mean operator $J_T = (I_T - \frac{1}{T}l_T l_T')$, where I_T is the $T \times T$ identity matrix, and $l_T = (1, \ldots, 1)'$, namely the $T \times 1$ vector of ones), so

$$\tilde{Y}_{nt} = Y_{nt} - \underbrace{\sum_{t=1}^{T} Y_{nt}/T}_{\text{suff stat for } c_{n0} \text{ under Gauss}} \quad \text{and } \tilde{X}_{nt} = X_{nt} - \sum_{t=1}^{T} X_{nt}/T.$$

Gaussian MLE

The log-likelihood reads as:

$$\ell_{n,T}(\theta) = -\frac{n(T-1)}{2}\ln(2\pi\sigma^2) + (T-1)[\ln|S_n(\lambda)| + \ln|R_n(\rho)|] - \frac{1}{2\sigma^2}\sum_{t=1}^{T} \tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta),$$

where the **transformed errors**

$$\tilde{V}_{nt}(\zeta) = R_n(\rho)[S_n(\lambda)\tilde{Y}_{nt} - \tilde{X}_{nt}\beta],$$

which are iid Gaussian and are obtained using the transformed variables and transformed covariates (essentially, the transformation is based the time-mean operator $J_T = (I_T - \frac{1}{T} l_T l'_T)$, where I_T is the $T \times T$ identity matrix, and $l_T = (1, \ldots, 1)'$, namely the $T \times 1$ vector of ones), so

$$\tilde{Y}_{nt} = Y_{nt} - \underbrace{\sum_{t=1}^{T} Y_{nt}/T}_{\text{suff stat for } c_{n0} \text{ under Gauss}} \quad \text{and } \tilde{X}_{nt} = X_{nt} - \sum_{t=1}^{T} X_{nt}/T.$$

The MLE $\hat{\theta}_{n,T}$ for θ is obtained solving $\hat{\theta}_{n,T} = \arg\max_{\theta\in\Theta} \ell_{n,T}(\theta)$. This is an **M-estimator** related to the log-likelihood score function, which can be written as a quadratic form in $\tilde{V}_{nt}(\zeta)$.

Jump Back

HOA for the Gaussian MLE

Set $d = \dim(\Theta)$, the MLE is the solution to

$$\frac{1}{n}\sum_{t=1}^{T}\left(\sum_{i=1}^{n}(T-1)^{-1}\psi_{i,t,1}(\hat{\theta}_{n,T}), \ldots, \sum_{i=1}^{n}(T-1)^{-1}\psi_{i,t,d}(\hat{\theta}_{n,T})\right)' = 0,$$

with $\psi_{i,t,j}$ represents the j-th component of the likelihood score, at time t in location i. The M-functional ϑ related to the MLE is implicitly defined as the unique functional root of:

$$\mathbb{E}\left\{\sum_{t=1}^{T}(T-1)^{-1}\,\psi_{nt}\left[\vartheta(P_{\theta_0})\right]\right\} = 0, \tag{4}$$

or equivalently via the asymptotic maximization

$$\theta_0 = \arg\max_{\theta\in\Theta}\mathbb{E}[\ell_{n,T}(\theta_0)] = \vartheta(P_{\theta_0})$$

The finite sample version of the M-functional is the M-estimator $\hat{\theta}_{n,T} = \vartheta(P_{n,T})$.

HOA for the Gaussian MLE (cont'd)

Under additional assumptions, the following von Mises expansion holds:

$$\vartheta(P_{n,T}) - \vartheta(P_{\theta_0}) = \underbrace{\frac{1}{n}\sum_{i=1}^{n} IF_{i,T}(\psi, P_{\theta_0})}_{\text{1st order}} + \underbrace{\frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \varphi_{i,j,T}(\psi, P_{\theta_0})}_{\text{2nd order}} + O_P(m^{-3/2}), \quad (5)$$

where, for $M_{i,T}(\psi, P_{\theta_0}) = \mathbb{E}\left[-(T-1)^{-1}\sum_{t=1}^{T} \partial\psi_{i,t}(\theta)/\partial\theta\Big|_{\theta=\theta_0}\right]$, we have

$$IF_{i,T}(\psi, P_{\theta_0}) = M_{i,T}^{-1}(\psi, P_{\theta_0})(T-1)^{-1}\sum_{t=1}^{T}\psi_{i,t}(\theta_0), \quad (6)$$

and (the expression of $\Gamma_{i,j,T}(\psi, P_{\theta_0})$ is provided in the paper)

$$\begin{aligned}\varphi_{i,j,T}(\psi, P_{\theta_0}) \quad = \quad & IF_{i,T}(\psi, P_{\theta_0}) + IF_{j,T}(\psi, P_{\theta_0}) + M_{i,T}^{-1}(\psi, P_{\theta_0})\Gamma_{i,j,T}(\psi, P_{\theta_0}) \\ & + M_{i,T}^{-1}(\psi, P_{\theta_0})\left\{(T-1)^{-1}\sum_{t=1}^{T}\frac{\partial\psi_{j,t}(\theta)}{\partial\theta}\Big|_{\theta_0} IF_{i,T}(\psi, P_{\theta_0})\right. \\ & \left. +(T-1)^{-1}\sum_{t=1}^{T}\frac{\partial\psi_{i,t}(\theta)}{\partial\theta}\Big|_{\theta=\theta_0} IF_{j,T}(\psi, P_{\theta_0})\right\} \end{aligned} \quad (7)$$

HOA for the Gaussian MLE (cont'd)

Let q be a function from $\mathbb{R}^d$ to $\mathbb{R}$, which has continuous and nonzero gradient at $\theta = \theta_0$ and continuous second derivative at $\theta = \theta_0$. Then, the following expansion holds:

$$q[\vartheta(P_{n,T})] - q[\vartheta(P_{\theta_0})] = \underbrace{\frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} h_{i,j,T}\left(\psi, P_{\theta_0}\right)}_{\text{U-statistic of order two}} + O_P(m^{-3/2}),$$

where

$$\begin{aligned} h_{i,j,T}\left(\psi, P_{\theta_0}\right) &= g_{i,T}\left(\psi, P_{\theta_0}\right) + g_{j,T}\left(\psi, P_{\theta_0}\right) + \gamma_{i,j,T}(\psi, P_{\theta_0}) \\ &= \frac{1}{2}\left\{ IF'_{i,T}(\psi, P_{\theta_0}) + IF'_{j,T}(\psi, P_{\theta_0}) + \varphi'_{i,j,T}(\psi, P_{\theta_0}) \right\} \frac{\partial q(\vartheta)}{\partial \vartheta}\Big|_{\theta=\theta_0} \\ &+ \frac{1}{2} IF'_{i,T}(\psi, P_{\theta_0}) \frac{\partial^2 q(\vartheta)}{\partial \vartheta \partial \vartheta'}\Big|_{\theta=\theta_0} IF_{j,T}(\psi, P_{\theta_0}), \end{aligned}$$

with

$$g_{i,T}(\psi, P_{\theta_0}) = \frac{1}{2}\left(IF'_{i,T}(\psi, P_{\theta_0}) \frac{\partial q(\vartheta)}{\partial \vartheta}\Big|_{\theta=\theta_0} \right),$$

$$\gamma_{i,j,T}(\psi, P_{\theta_0}) = \frac{1}{2}\left(\varphi'_{i,j,T}(\psi, P_{\theta_0}) \frac{\partial q(\vartheta)}{\partial \vartheta}\Big|_{\theta=\theta_0} + IF'_{i,T}(\psi, P_{\theta_0}) \frac{\partial^2 q(\vartheta)}{\partial \vartheta \partial \vartheta'}\Big|_{\theta=\theta_0} IF_{j,T}(\psi, P_{\theta_0}) \right).$$

HOA for the Gaussian MLE (cont'd)

We derive the Edgeworth expansion for $q[\vartheta(P_{n,T})] - q[\vartheta(P_{\theta_0})]$, working along the same lines as in Bickel et al., 1986, AoS.

Our derivation of the saddlepoint density approximation to $f_{n,T}(z)$ is based on the tilted-Edgeworth expansion for U-statistics of order two.

HOA for the Gaussian MLE (cont'd)

We derive the Edgeworth expansion for $q[\vartheta(P_{n,T})] - q[\vartheta(P_{\theta_0})]$, working along the same lines as in Bickel et al., 1986, AoS.

Our derivation of the saddlepoint density approximation to $f_{n,T}(z)$ is based on the tilted-Edgeworth expansion for U-statistics of order two.

Now, let $\sigma_{n,T}$ be the standard deviation of $q[\vartheta(P_{n,T})] - q[\vartheta(P_{\theta_0})]$. The Edgeworth expansion $\Lambda_m(z)$ for the c.d.f. F_m of

$$\sigma_{n,T}^{-1}\{q[\vartheta(P_{n,T})] - q[\vartheta(P_{\theta_0})]\}$$

is

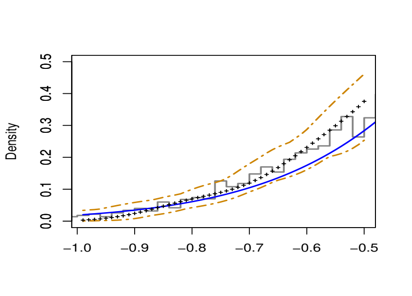
$$\Lambda_m(z) = \Phi(z) - \phi(z)\left\{n^{-1/2}\frac{\kappa_{n,T}^{(3)}}{3!}(z^2-1) + n^{-1}\frac{\kappa_{n,T}^{(4)}}{4!}(z^3-3z) + n^{-1}\frac{\kappa_{n,T}^{(3)}}{72}(z^5-10z^2+15z)\right\} \tag{8}$$

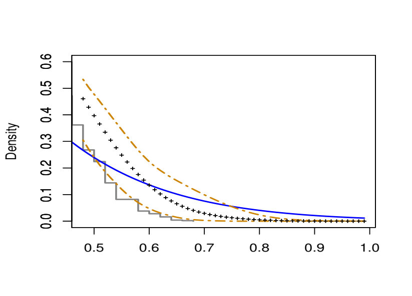
where $z \in \mathcal{A}$, , $\Phi(z)$ and $\phi(z)$ are the c.d.f. and p.d.f of a standard normal r.v. respectively, $\kappa_{n,T}^{(3)}n^{-1/2}$ and $\kappa_{n,T}^{(4)}n^{-1}$ are the approximate third and fourth cumulants of $\sigma_{n,T}^{-1}\{q[\vartheta(P_{n,T})] - q[\vartheta(P_{\theta_0})]\}$, as defined in the paper.

Jump Back

Comparison with Parametric Bootstrap

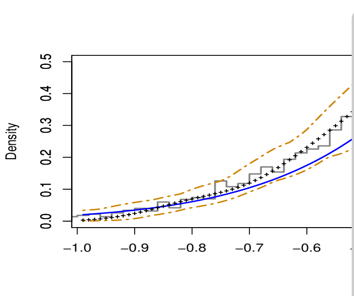
The parametric bootstrap represents a computer-based competitor: we compare our saddlepoint approximation to the one obtained by bootstrap, for $B = 499$ bootstrap repetitions. We display the functional boxplots (as obtained iterating the procedure 100 times) of the bootstrap approximated density, for sample size $n = 24$ and for W_n is Queen.





Comparison with Parametric Bootstrap

The parametric bootstrap represents a computer-based competitor: we compare our saddlepoint approximation to the one obtained by bootstrap, for $B = 499$ bootstrap repetitions. We display the functional boxplots (as obtained iterating the procedure 100 times) of the bootstrap approximated density, for sample size $n = 24$ and for W_n is Queen.



The bootstrap median functional curve is close to the actual density, the range between the quartile curves illustrates that the bootstrap approximation has a variability, which depends on B. However, larger values of B entail bigger computational costs: when $B = 499$, the bootstrap is almost as fast as the saddlepoint density approximation, but for $B = 999$, it is three times slower.

Jump Back

tvSAR [ongoing]

I plan to define a novel class of M-type estimators for local stationarity for random fields, focussing in the first place on estimation for tv SAR models. My ideas:

$$Inv_{nt,T} = \lambda_0(t/T) W_n Inv_{nt,T} + V_{nt} \quad t = 1, 2, \cdots, T \tag{9}$$

Notice that, in line with the time series literature on locally stationary processes, Eq. (9) is written for the triangular array $\{Inv_{nt,T}\}$, observed from $[0, T]$.

tvSAR [ongoing]

I plan to define a novel class of M-type estimators for local stationarity for random fields, focussing in the first place on estimation for tv SAR models. My ideas:

$$Inv_{nt,T} = \lambda_0(t/T) W_n Inv_{nt,T} + V_{nt} \quad t = 1, 2, \cdots, T \tag{9}$$

Notice that, in line with the time series literature on locally stationary processes, Eq. (9) is written for the triangular array $\{Inv_{nt,T}\}$, observed from $[0, T]$.

The spatial relation in (9) is controlled by the curve $\lambda_0(\cdot)$, which depends on the rescaled time t/T. To estimate the time-varying parameter $\lambda_0(\cdot)$, we define the stationary process $\{\widetilde{Inv}_{nt}(u)\}$, indexed by $u \in [0, 1]$ and having dynamics:

$$\widetilde{Inv}_{nt}(u) = \lambda_0(u) W_n \widetilde{Inv}_{nt}(u) + V_{nt}, \quad t = 1, 2, \cdots, T \tag{10}$$

tvSAR [ongoing]

Intuition (twofold)

- *In some neighbourhood of* t*, the non stationary process* $\{Inv_{nt,T}\}$ *can be approximated (in a suitable stochastic sense) by the stationary one* $\{\widetilde{Inv}_{nt}(u)\}$*. Since the spatial parameter of* $\{Inv_{nt,T}\}$ *changes smoothly over* $[0, T]$*, the approximation accuracy should depends both on* T *and on* $|t/T - u|$*. Indeed, a comparison of (9) and (10) suggests that if* $t/T \approx u$*, then* $Inv_{nt,T}$ *and* $\widetilde{Inv}_{nt}(u)$ *should be close with high probability.*

tvSAR [ongoing]

Intuition (twofold)

- *In some neighbourhood of* t*, the non stationary process* $\{Inv_{nt,T}\}$ *can be approximated (in a suitable stochastic sense) by the stationary one* $\{\widetilde{Inv}_{nt}(u)\}$*. Since the spatial parameter of* $\{Inv_{nt,T}\}$ *changes smoothly over* $[0, T]$*, the approximation accuracy should depends both on* T *and on* $|t/T - u|$*. Indeed, a comparison of (9) and (10) suggests that if* $t/T \approx u$*, then* $Inv_{nt,T}$ *and* $\widetilde{Inv}_{nt}(u)$ *should be close with high probability.*

- *Use Gaussian PMLE over segments of time (controlled by a kernel in time which pre-multiply the estimating function)*

tvSAR [ongoing]

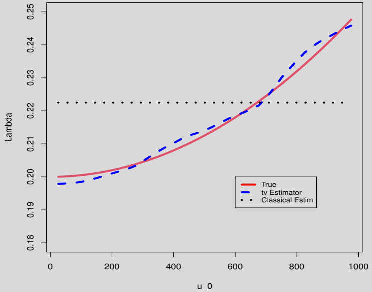
I conducted the following simulation study: (1) I generated 100 samples for $\{Inv_{nt,T}\}$, for $n = 24$ and $T = 1000$, from the model in (9), with $\lambda_0(\cdot)$ being a parabola in the rescaled time;

tvSAR [ongoing]

I conducted the following simulation study: (1) I generated 100 samples for $\{Inv_{nt,T}\}$, for $n = 24$ and $T = 1000$, from the model in (9), with $\lambda_0(\cdot)$ being a parabola in the rescaled time; (2) I computed the Segment-PMLE (which in this case admits a simple closed-form expression) and the extant Gaussian likelihood estimation method (which is applied under the assumption that the spatial parameter is constant).

tvSAR [ongoing]

I conducted the following simulation study: (1) I generated 100 samples for $\{Inv_{nt,T}\}$, for $n = 24$ and $T = 1000$, from the model in (9), with $\lambda_0(\cdot)$ being a parabola in the rescaled time; (2) I computed the Segment-PMLE (which in this case admits a simple closed-form expression) and the extant Gaussian likelihood estimation method (which is applied under the assumption that the spatial parameter is constant).



tvSAR [ongoing]

I conducted the following simulation study: (1) I generated 100 samples for $\{Inv_{nt,T}\}$, for $n = 24$ and $T = 1000$, from the model in (9), with $\lambda_0(\cdot)$ being a parabola in the rescaled time; (2) I computed the Segment-PMLE (which in this case admits a simple closed-form expression) and the extant Gaussian likelihood estimation method (which is applied under the assumption that the spatial parameter is constant).