Spatial Modelling of Mortality Rates with Heteroscedastic Residuals

TIES-GRASPA 2017, Bergamo

Robert Garthoff, **Philipp Otto**, Wolfgang Schmid

European University Viadrina, Frankfurt (Oder), Germany

July 26, 2017

Motivation

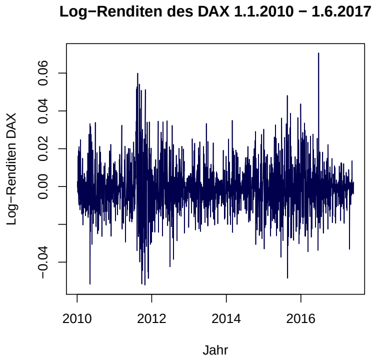

Figure: Daily log-returns of the German index DAX-30 from January 2010 to June 2017.

- Daily, weekly, and monthly log-returns are not normally distributed.
- The log-returns have heavy tails. The largest tails are observed for the daily returns, followed by weekly, monthly and annual returns.
- All log-returns are "slighty" skewed to the left.
- The log-returns seem to be weakly dependent or even uncorrelated, but the squared and the absolute log-returns are correlated at least for daily and the weekly observations.
- The volatility seems to be asymmetric. A negative return shock has more impact than a positive return shock.

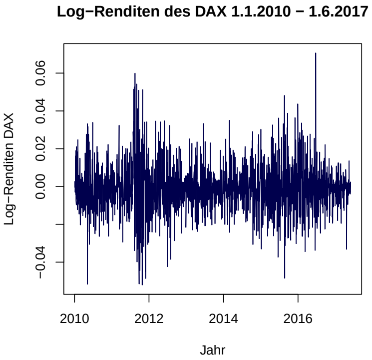

Figure: Daily log-returns of the German index DAX-30 from January 2010 to June 2017.

- Daily, weekly, and monthly log-returns are not normally distributed.
- The log-returns have heavy tails. The largest tails are observed for the daily returns, followed by weekly, monthly and annual returns.
- All log-returns are "slighty" skewed to the left.
- The log-returns seem to be weakly dependent or even uncorrelated, but the squared and the absolute log-returns are correlated at least for daily and the weekly observations.
- The volatility seems to be asymmetric. A negative return shock has more impact than a positive return shock.

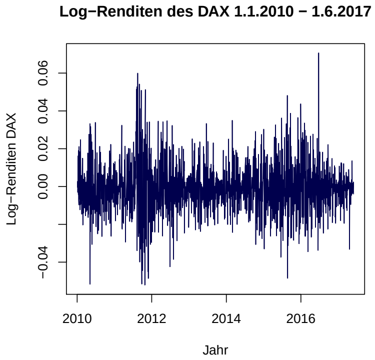

Figure: Daily log-returns of the German index DAX-30 from January 2010 to June 2017.

- Daily, weekly, and monthly log-returns are not normally distributed.
- The log-returns have heavy tails. The largest tails are observed for the daily returns, followed by weekly, monthly and annual returns.
- All log-returns are "slighty" skewed to the left.
- The log-returns seem to be weakly dependent or even uncorrelated, but the squared and the absolute log-returns are correlated at least for daily and the weekly observations.
- The volatility seems to be asymmetric. A negative return shock has more impact than a positive return shock.

Engle (1982)

A process $\{X_t\}$ is called anautoregressive conditional heteroscedastic process of order 1, briefly ARCH(1), if

$$X_t = \varepsilon_t \, \sigma_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 \, X_{t-1}^2$$

with $\alpha_0 > 0$, $\alpha_1 \geq 0$ and

$$E(\varepsilon_t) = 0, Cov(\varepsilon_t, \varepsilon_s) = 0 \text{ for } s \neq t, Var(\varepsilon_t) = 1.$$

- Generalized ARCH process (GARCH, Bollerslev 1986)
- Further generalizations, extensions (NGARCH, EGARCH, TGARCH, ...)

Engle (1982)

A process $\{X_t\}$ is called an autoregressive conditional heteroscedastic process of order 1, briefly ARCH(1), if

$$X_t = \varepsilon_t \, \sigma_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 \, X_{t-1}^2$$

with $\alpha_0 > 0$, $\alpha_1 \geq 0$ and

$$E(\varepsilon_t) = 0, Cov(\varepsilon_t, \varepsilon_s) = 0 \text{ for } s \neq t, Var(\varepsilon_t) = 1.$$

- Generalized ARCH process (GARCH, Bollerslev 1986)
- Further generalizations, extensions (NGARCH, EGARCH, TGARCH, ...)

Empirical Motivation

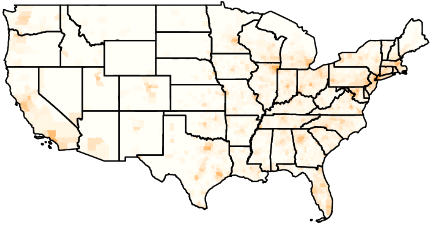

Figure: Population density of the U.S. counties (except Alaska and Hawaii) in 2010 projected on the map. The darker the area is drawn the higher is the observation in the respective location.

Empirical Motivation

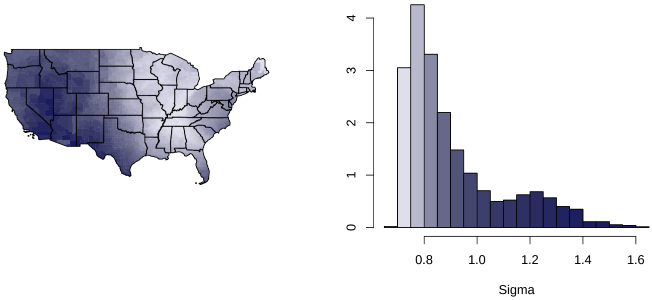

Figure: Sample standard deviation of the estimated residuals within a radius of 500 km (310.686 $miles$) of a spatial autoregressive process with $\hat{\lambda} = 0.8578$ and $\hat{\sigma_\xi}^2 = 0.9104$.

Empirical Motivation

$\rightsquigarrow$ **Spatial ARCH Model**

Simulation and Visualization

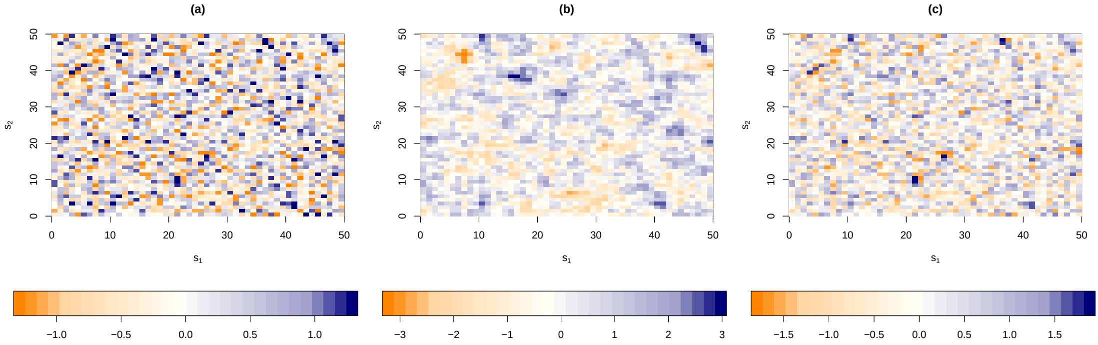

Figure: (a) Spatial White Noise, (b) spatial autoregressive process, and (c) spatial ARCH process on a two-dimensional spatial lattice ($\boldsymbol{s} \in \mathbb{Z}^2$).

1. Motivation
2. Modelling Approach
3. Estimation of the ARCH parameters
4. Modelling of Mortality Rates with Heteroscedastic Residuals
5. Extensions and concluding remarks

The talk is based on the paper:
Spatial and Spatiotemporal Autoregressive Conditional Heteroscedasticity
with Wolfgang Schmid and Robert Garthoff,
`arXiv:1609.00711`

2. Theoretical Framework

Model

- Univariate spatial stochastic process $\{Y(\boldsymbol{s}) \in \mathbb{R} : \boldsymbol{s} \in D_{\boldsymbol{s}}\}$
- $D_{\boldsymbol{s}}$ is a subset of the q-dimensional set of real numbers $\mathbb{R}^q$ or the q-dimensional set of integers $\mathbb{Z}^q$
- $(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n)'$ denote all locations
- $\boldsymbol{Y} = (Y(\boldsymbol{s}_i))_{i=1,\ldots,n}$, $\boldsymbol{\alpha} = (\alpha(\boldsymbol{s}_i))_{i=1,\ldots,n}$
- Matrix $\mathbf{W}$ of spatial weights is assumed to be non-stochastic with zero elements on the diagonal

$$\boldsymbol{Y} = \mathrm{diag}(\boldsymbol{h})^{1/2}\boldsymbol{\varepsilon} \qquad Y(\boldsymbol{s}_i) = h(\boldsymbol{s}_i)^{1/2}\varepsilon(\boldsymbol{s}_i)$$

$$\boldsymbol{h} = \boldsymbol{\alpha} + \mathbf{W}(\boldsymbol{Y} \circ \boldsymbol{Y}) \qquad h(\boldsymbol{s}_i) = \alpha_i + \sum_{v=1}^{n} w_{iv} Y(\boldsymbol{s}_v)^2$$

Model

- Univariate spatial stochastic process $\{Y(\boldsymbol{s}) \in \mathbb{R} : \boldsymbol{s} \in D_{\boldsymbol{s}}\}$
- $D_{\boldsymbol{s}}$ is a subset of the q-dimensional set of real numbers $\mathbb{R}^q$ or the q-dimensional set of integers $\mathbb{Z}^q$
- $(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n)'$ denote all locations
- $\boldsymbol{Y} = (Y(\boldsymbol{s}_i))_{i=1,\ldots,n}$, $\boldsymbol{\alpha} = (\alpha(\boldsymbol{s}_i))_{i=1,\ldots,n}$
- Matrix $\mathbf{W}$ of spatial weights is assumed to be non-stochastic with zero elements on the diagonal

$$\boldsymbol{Y} = \text{diag}(\boldsymbol{h})^{1/2}\boldsymbol{\varepsilon} \qquad Y(\boldsymbol{s}_i) = h(\boldsymbol{s}_i)^{1/2}\varepsilon(\boldsymbol{s}_i)$$

$$\boldsymbol{h} = \boldsymbol{\alpha} + \mathbf{W}(\boldsymbol{Y} \circ \boldsymbol{Y}) \qquad h(\boldsymbol{s}_i) = \alpha_i + \sum_{v=1}^{n} w_{iv} Y(\boldsymbol{s}_v)^2$$

Model

- Univariate spatial stochastic process $\{Y(\boldsymbol{s}) \in \mathbb{R} : \boldsymbol{s} \in D_{\boldsymbol{s}}\}$
- $D_{\boldsymbol{s}}$ is a subset of the q-dimensional set of real numbers $\mathbb{R}^q$ or the q-dimensional set of integers $\mathbb{Z}^q$
- $(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n)'$ denote all locations
- $\boldsymbol{Y} = (Y(\boldsymbol{s}_i))_{i=1,\ldots,n}$, $\boldsymbol{\alpha} = (\alpha(\boldsymbol{s}_i))_{i=1,\ldots,n}$
- Matrix $\mathbf{W}$ of spatial weights is assumed to be non-stochastic with zero elements on the diagonal

$$\begin{aligned} \boldsymbol{Y} &= \text{diag}(\boldsymbol{h})^{1/2}\boldsymbol{\varepsilon} \\ \boldsymbol{h} &= \boldsymbol{\alpha} + \mathbf{W}(\boldsymbol{Y} \circ \boldsymbol{Y}) \end{aligned}$$

$$\begin{aligned} Y(\boldsymbol{s}_i) &= h(\boldsymbol{s}_i)^{1/2}\varepsilon(\boldsymbol{s}_i) \\ h(\boldsymbol{s}_i) &= \alpha_i + \sum_{v=1}^{n} w_{iv} Y(\boldsymbol{s}_v)^2 \end{aligned}$$

Table: Selected models that are covered by the spatial approach.

Model	q	D_s	$(w_{ij})_{i,j=1,\ldots,n}$	triangular
Time series				
ARCH(1) Engle (1982)	1	$\mathbb{Z}$	$\alpha \mathbb{1}_{\{s_i - s_j = 1\}}$	✓
ARCH(p) Engle (1982)	1	$\mathbb{Z}$	$\sum_{k=1}^{p} \alpha_k \mathbb{1}_{\{s_i - s_j = k\}}$	✓
Spatial time series				
spatial ARCH				
Borovkova und Lopuhaa (2012)	1	$\mathbb{Z}$	$(a_{1,i} + a_{2,i} w_{ij}) \mathbb{1}_{\{s_i - s_j = 1\}}$	✓
Spatial models				
SARCH(1) Bera und Simlai (2004)	$2, 3$	$\mathbb{Z}^q$, $\mathbb{R}^q$	$\alpha_1 w_{ij}^2$	
Proposed models				
spARCH(p)	≥ 1	$\mathbb{Z}^q$, $\mathbb{R}^q$	Visualization	
oriented	≥ 1	$\mathbb{Z}^q$, $\mathbb{R}^q$	Visualization	✓

3. Estimation of the ARCH parameters

Theorem

Suppose that $\boldsymbol{\alpha} \geq \mathbf{0}$, $w_{ij} \geq 0$ *for all* $i, j = 1, \ldots, n$, $w_{ii} = 0$ *for all* $i = 1, \ldots, n$ *and that* $\det(\mathbf{I} - \mathbf{A}^2) \neq 0$. *If all elements of the matrix*

$$(\mathbf{I} - \mathbf{A}^2)^{-1} \tag{1}$$

are nonnegative, then all components of $\boldsymbol{Y}^{(2)}$ *are nonnegative; i.e.,* $Y(\boldsymbol{s}_i)^2 \geq 0$ *for* $i = 1, \ldots, n$. *Moreover,* $h(\boldsymbol{s}_i) \geq 0$ *for* $i = 1, \ldots, n$.

Lemma

The elements of (1) *are nonnegative, if*

- $\mathbf{W}$ *is a strictly triangular matrix (directional process).* comparison $\mathbf{W}$
- $||\mathbf{A}^2|| < 1$ *for an arbitrary induced matrix norm.*
- $|\varepsilon(\boldsymbol{s}_i)| \leq a < 1/||\mathbf{W}^2||_1^{1/4}$ *for* $i = 1, \ldots, n$. influence of compact support

properties

Theorem

Suppose that $\boldsymbol{\alpha} \geq \mathbf{0}$, $w_{ij} \geq 0$ *for all* $i, j = 1, \ldots, n$, $w_{ii} = 0$ *for all* $i = 1, \ldots, n$ *and that* $\det(\mathbf{I} - \mathbf{A}^2) \neq 0$. *If all elements of the matrix*

$$(\mathbf{I} - \mathbf{A}^2)^{-1} \tag{1}$$

are nonnegative, then all components of $\boldsymbol{Y}^{(2)}$ *are nonnegative; i.e.,* $Y(\boldsymbol{s}_i)^2 \geq 0$ *for* $i = 1, \ldots, n$. *Moreover,* $h(\boldsymbol{s}_i) \geq 0$ *for* $i = 1, \ldots, n$.

Lemma

The elements of (1) *are nonnegative, if*

- $\mathbf{W}$ *is a strictly triangular matrix (directional process).* comparison $\mathbf{W}$
- $||\mathbf{A}^2|| < 1$ *for an arbitrary induced matrix norm.*
- $|\varepsilon(\boldsymbol{s}_i)| \leq a < 1/||\mathbf{W}^2||_1^{1/4}$ *for* $i = 1, \ldots, n$. influence of compact support

properties

Estimation

Joint distribution

$$
\begin{aligned}
f_{\boldsymbol{Y}}(\boldsymbol{y}) &= f_{(Y(\boldsymbol{s}_1),\ldots,Y(\boldsymbol{s}_n))}(y_1,\ldots,y_n) \\
&= f_{(\varepsilon(\boldsymbol{s}_1),\ldots,\varepsilon(\boldsymbol{s}_n))}\left(\frac{y_1}{\sqrt{h_1}},\ldots,\frac{y_n}{\sqrt{h_n}}\right)\left|\det\left(\left(\frac{\partial y_j/\sqrt{h_j}}{\partial y_i}\right)_{i,j=1,\ldots,n}\right)\right| \qquad (2)
\end{aligned}
$$

with

$$
\left|\det\left(\left(\frac{\partial y_j/\sqrt{h_j}}{\partial y_i}\right)_{i,j=1,\ldots,n}\right)\right| = \prod_{i=1}^{n}\frac{y_i^2}{h_i^{3/2}} \cdot \left|\det\left(\mathrm{diag}\left(\frac{h_1}{y_1^2},\ldots,\frac{h_n}{y_n^2}\right)+\mathbf{W}'\right)\right|.
$$

Estimation

$\mathbf{W}$ is a strictly triangular matrix

$$
\begin{aligned}
f_{\boldsymbol{Y}}(\boldsymbol{y}) &= \prod_{i=1}^{n} \left(f_{\varepsilon}\left(\frac{y_i}{\sqrt{h_i}}\right) \frac{1}{\sqrt{h_i}} \right) = \prod_{i=1}^{n} f_{Y(\boldsymbol{s}_i)|Y(\boldsymbol{s}_{i-1}),\ldots,Y(\boldsymbol{s}_1)}(y_i | y_{i-1}, \ldots, y_1) \\
\log(f_{\boldsymbol{Y}}(\boldsymbol{y})) &= \sum_{i=1}^{n} \left(\log\left(f_{\varepsilon}\left(\frac{y_i}{\sqrt{h_i}}\right)\right) - \frac{1}{2}\log(h_i) \right) .
\end{aligned}
$$

$\mathbf{W}$ is an arbitrary matrix

$$
\log\left|\det\left(\left(\frac{\partial y_j/\sqrt{h_j}}{\partial y_i}\right)_{i,j=1,\ldots,n}\right)\right| = \sum_{i=1}^{n}\left(2\log y_i - \frac{3}{2}\log h_i\right) + \sum_{i=1}^{n}\log|\lambda_i|
$$

and λ_i is the i-th eigenvalue of $\left(\mathrm{diag}\left(\frac{h_1}{y_1^2}, \ldots, \frac{h_n}{y_n^2}\right) + \rho\mathbf{W}'\right)$

Estimation

$\mathbf{W}$ is a strictly triangular matrix

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \prod_{i=1}^{n} \left(f_{\varepsilon}\left(\frac{y_i}{\sqrt{h_i}}\right) \frac{1}{\sqrt{h_i}} \right) = \prod_{i=1}^{n} f_{Y(\boldsymbol{s}_i)|Y(\boldsymbol{s}_{i-1}),\ldots,Y(\boldsymbol{s}_1)}(y_i|y_{i-1},\ldots,y_1)$$

$$\log(f_{\boldsymbol{Y}}(\boldsymbol{y})) = \sum_{i=1}^{n} \left(\log\left(f_{\varepsilon}\left(\frac{y_i}{\sqrt{h_i}}\right)\right) - \frac{1}{2}\log(h_i) \right) .$$

$\mathbf{W}$ is an arbitrary matrix

$$\log\left|\det\left(\left(\frac{\partial y_j/\sqrt{h_j}}{\partial y_i}\right)_{i,j=1,\ldots,n}\right)\right| = \sum_{i=1}^{n} \left(2\log y_i - \frac{3}{2}\log h_i\right) + \sum_{i=1}^{n} \log|\lambda_i|$$

and λ_i is the i-th eigenvalue of $\left(\mathrm{diag}\left(\frac{h_1}{y_1^2},\ldots,\frac{h_n}{y_n^2}\right) + \rho\mathbf{W}'\right)$

Performance for sparse data

$n = 4$

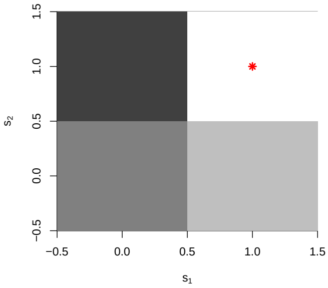

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

Performance for sparse data

$$n = 4,\ \rho = 0.2,\ \alpha = 5$$

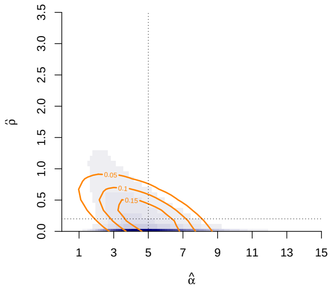

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

Performance for sparse data

$$n = 4,\ \rho = 0.6,\ \alpha = 5$$

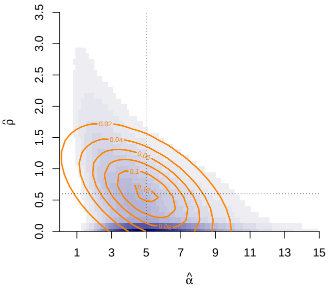

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

Performance for sparse data

$$n = 4,\ \rho = 0.9,\ \alpha = 5$$

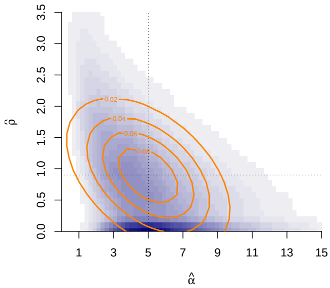

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

Performance for sparse data

$$n = 25$$

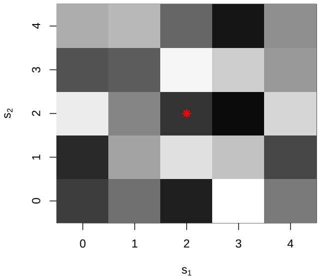

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

Performance for sparse data

$$n = 25,\ \rho = 0.2,\ \alpha = 5$$

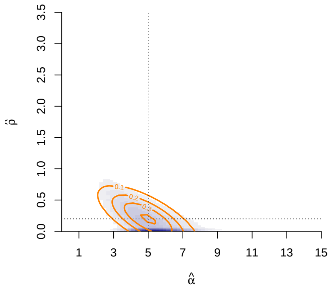

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

Performance for sparse data

$$n = 25,\ \rho = 0.6,\ \alpha = 5$$

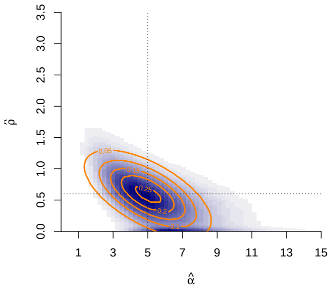

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

Performance for sparse data

$$n = 25,\ \rho = 0.9,\ \alpha = 5$$

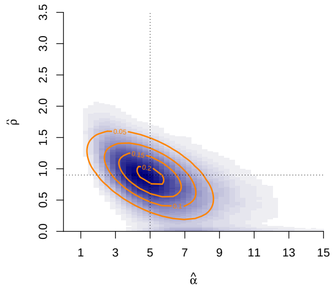

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

Performance for sparse data

$n = 100$

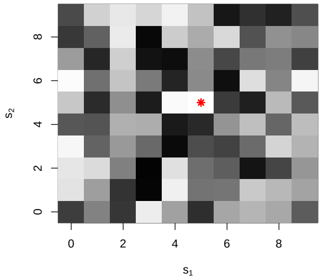

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

Performance for sparse data

$$n = 100,\ \rho = 0.2,\ \alpha = 5$$

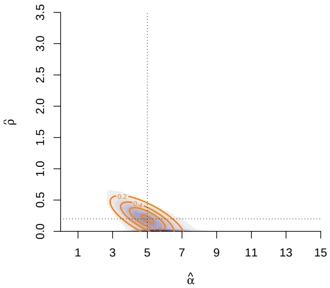

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

Performance for sparse data

$$n = 100,\ \rho = 0.6,\ \alpha = 5$$

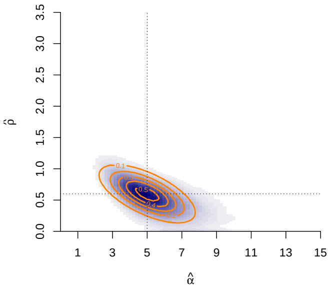

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

Performance for sparse data

$$n = 100, \ \rho = 0.9, \ \alpha = 5$$

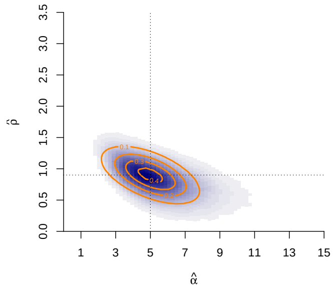

Figure: Performance of the estimators for sparse data and a triangular matrix $\mathbf{W}$. The estimated distributions of the parameter estimates are plotted as two-dimensional histograms, whereas the theoretical distribution of $(\hat{\alpha}, \hat{\rho})$ with the median elements of the covariance matrices of each replication is added as orange contour lines.

4. Modelling of Mortality Rates with Heteroscedastic Residuals

$$\boldsymbol{Y} = \mu\mathbf{1} + \lambda\mathbf{B}\boldsymbol{Y} + \boldsymbol{\xi}\text{, i.e.}\quad \boldsymbol{Y} = (\mathbf{I} - \lambda\mathbf{B})^{-1}(\mu\mathbf{1} + \boldsymbol{\xi}) \qquad \text{with}$$

$$\boldsymbol{\xi} = \text{diag}(\boldsymbol{h})^{1/2}\boldsymbol{\varepsilon} \qquad \text{and}$$

$$\boldsymbol{h} = \boldsymbol{\alpha} + \mathbf{W}\,\text{diag}(\boldsymbol{\xi})\boldsymbol{\xi}\,.$$

- 5-year average mortality (2008–2012) caused by cancer of the lungs or bronchus provided by the Center for Disease Control and Prevention (U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute (2015))
- death rates are age-adjusted to the 2000 U.S. standard population
- 3108 counties; all U.S. counties excluding Alaska and Hawaii
- covariates:
  - particulate matter $PM_{2.5}$, PM_{10}, NO_2, SO_2, CO, O_3
  - percentage of smokers in 2012
  - personal income per capita

$$\boldsymbol{Y} = \mu\mathbf{1} + \lambda\mathbf{B}\boldsymbol{Y} + \boldsymbol{\xi}\,, \text{i.e.} \quad \boldsymbol{Y} = (\mathbf{I} - \lambda\mathbf{B})^{-1}(\mu\mathbf{1} + \boldsymbol{\xi}) \qquad \text{with}$$

$$\begin{aligned} \boldsymbol{\xi} &= \operatorname{diag}(\boldsymbol{h})^{1/2}\boldsymbol{\varepsilon} \qquad \text{and} \\ \boldsymbol{h} &= \boldsymbol{\alpha} + \mathbf{W}\operatorname{diag}(\boldsymbol{\xi})\boldsymbol{\xi}\,. \end{aligned}$$

- 5-year average mortality (2008–2012) caused by cancer of the lungs or bronchus provided by the Center for Disease Control and Prevention (U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute (2015))
- death rates are age-adjusted to the 2000 U.S. standard population
- 3108 counties; all U.S. counties excluding Alaska and Hawaii
- covariates:
  - particulate matter $PM_{2.5}$, PM_{10}, NO_2, SO_2, CO, O_3
  - percentage of smokers in 2012
  - personal income per capita

$$\boldsymbol{Y} = \mu\mathbf{1} + \lambda\mathbf{B}\boldsymbol{Y} + \boldsymbol{\xi}\text{, i.e. } \boldsymbol{Y} = (\mathbf{I} - \lambda\mathbf{B})^{-1}(\mu\mathbf{1} + \boldsymbol{\xi}) \quad \text{with}$$

$$\begin{aligned} \boldsymbol{\xi} &= \operatorname{diag}(\boldsymbol{h})^{1/2}\boldsymbol{\varepsilon} \quad \text{and} \\ \boldsymbol{h} &= \boldsymbol{\alpha} + \mathbf{W}\operatorname{diag}(\boldsymbol{\xi})\boldsymbol{\xi}\,. \end{aligned}$$

- 5-year average mortality (2008–2012) caused by cancer of the lungs or bronchus provided by the Center for Disease Control and Prevention (U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute (2015))
- death rates are age-adjusted to the 2000 U.S. standard population
- 3108 counties; all U.S. counties excluding Alaska and Hawaii
- covariates:
  - particulate matter $PM_{2.5}$, PM_{10}, NO_2, SO_2, CO, O_3
  - percentage of smokers in 2012
  - personal income per capita

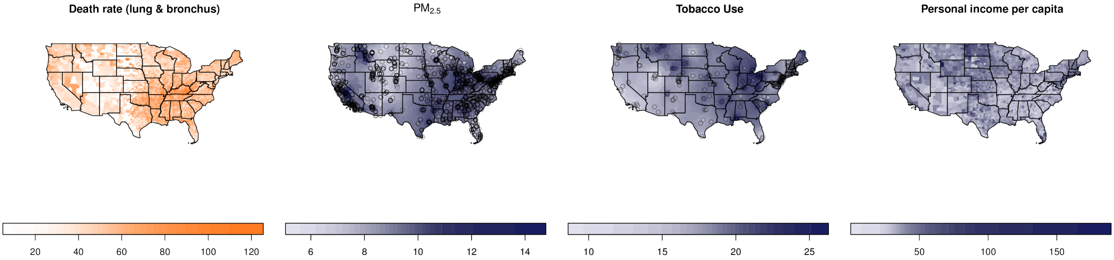

Figure: Mortality caused by cancer of the lungs or bronchus in U.S. counties (most left) and main covariates: annual average of $PM_{2.5}$ in 2012 (middle left), percentage of smokers in 2012 (middle right), and personal income per capita in thousand U.S. dollars (most right). The measurement stations of the covariates $PM_{2.5}$ and the percentage of smokers are indicated on the maps via empty circles.

Results I

	Linear Regression		
	Estimate	Standard Error	p-Value
Intercept	-10.8575	0.9964	0.0000
Environmental Covariates			
PM_{10}	-1.1734	0.0874	0.0000
$PM_{2.5}$	2.1193	0.1427	0.0000
SO_2	0.1210	0.0422	0.0042
NO_2	0.6217	0.0799	0.0000
O_3	-2.4133	0.2489	0.0000
CO	-0.4041	0.1151	0.0005
Behavioral Covariates			
Tobacco Use	1.2859	0.1677	0.0000
Economic Covariates			
Personal Income	-	-	-
Spatial Coefficients			
λ_1			
λ_2			
$\sigma^2_{\boldsymbol{\xi}}$			
α			
ρ			
Summary Statistics			
Moran's I $\boldsymbol{\xi}$	0.2203	0.0106	0.0000
Moran's I $\boldsymbol{\xi}^{(2)}$	0.3331	0.0106	0.0000
AIC		7091.062	

Results II

	SAR Estimate	SAR Standard Error	SAR p-Value
Intercept	-4.3157	0.9347	0.0000
Environmental Covariates			
PM_{10}	-0.3584	0.0844	0.0000
$PM_{2.5}$	0.6162	0.1402	0.0000
SO_2	-	-	-
NO_2	0.2731	0.0732	0.0002
O_3	-0.8082	0.2251	0.0003
CO	-0.1759	0.1026	0.0863
Behavioral Covariates			
Tobacco Use	0.6090	0.1381	0.0000
Economic Covariates			
Personal Income	-	-	-
Spatial Coefficients			
λ_1	0.2449	0.0262	0.0000
λ_2	0.4431	0.0327	0.0000
$\sigma^2_{\boldsymbol{\xi}}$	0.4628	0.0119	0.0000
α			
ρ			
Summary Statistics			
Moran's I $\boldsymbol{\xi}$	-0.0114	0.0106	0.2966
Moran's I $\boldsymbol{\xi}^{(2)}$	0.3212	0.0106	0.0000
AIC		6560.509	

Results III

	SARspARCH		
	Estimate	Standard Error	p-Value
Intercept	-0.0059	0.1629	0.9712
Environmental Covariates			
PM_{10}	-0.2641	0.0397	0.0000
$PM_{2.5}$	0.5365	0.0632	0.0000
SO_2	-	-	-
NO_2	-	-	-
O_3	-	-	-
CO	-	-	-
Behavioral Covariates			
Tobacco Use	0.3188	0.0665	0.0000
Economic Covariates			
Personal Income	-	-	-
Spatial Coefficients			
λ_1	0.2624	0.0278	0.0000
λ_2	0.3888	0.0400	0.0000
$\sigma^2_{\boldsymbol{\xi}}$			
α	0.0601	0.0015	0.0000
ρ	0.6680	0.0161	0.0000
Summary Statistics			
Moran's I $\boldsymbol{\varepsilon}$	0.0075	0.0106	0.4565
Moran's I $\boldsymbol{\varepsilon}^{(2)}$	0.0067	0.0106	0.4852
AIC		2484.686	

4. Extensions and concluding remarks

Further research, open problems

- R-package spGARCH for simulation and estimation of spatial (G)ARCH models
- Influence of conditional spatial heteroscedasticity on the estimated coefficients of a SAR model
- Stationarity, isotropy, and multivariate extension of the model; spatial GARCH model
- Finance: volatility clusters, asset returns depend on many exogenous factors $\rightsquigarrow$ but usually no spatial dependence
  However, one might expect spatial dependence for spatially constraint markets (e.g. real-estate market)
- Influence of the level of spatial aggregation on spatial heteroscedasticity

Further research, open problems

- R-package spGARCH for simulation and estimation of spatial (G)ARCH models
- Influence of conditional spatial heteroscedasticity on the estimated coefficients of a SAR model
- Stationarity, isotropy, and multivariate extension of the model; spatial GARCH model
- Finance: volatility clusters, asset returns depend on many exogenous factors $\rightsquigarrow$ but usually no spatial dependence
  However, one might expect spatial dependence for spatially constraint markets (e.g. real-estate market)
- Influence of the level of spatial aggregation on spatial heteroscedasticity

Further research, open problems

- R-package spGARCH for simulation and estimation of spatial (G)ARCH models
- Influence of conditional spatial heteroscedasticity on the estimated coefficients of a SAR model
- Stationarity, isotropy, and multivariate extension of the model; spatial GARCH model
- Finance: volatility clusters, asset returns depend on many exogenous factors $\rightsquigarrow$ but usually no spatial dependence
  However, one might expect spatial dependence for spatially constraint markets (e.g. real-estate market)
- Influence of the level of spatial aggregation on spatial heteroscedasticity

Further research, open problems

- R-package spGARCH for simulation and estimation of spatial (G)ARCH models
- Influence of conditional spatial heteroscedasticity on the estimated coefficients of a SAR model
- Stationarity, isotropy, and multivariate extension of the model; spatial GARCH model
- Finance: volatility clusters, asset returns depend on many exogenous factors $\rightsquigarrow$ but usually no spatial dependence
  However, one might expect spatial dependence for spatially constraint markets (e.g. real-estate market)
- Influence of the level of spatial aggregation on spatial heteroscedasticity

Thank you for your attention!

ID 508: Philipp Otto "Spatial Modelling of Mortality Rates with Heteroscedastic Residuals"

Spatial and Spatiotemporal Autoregressive Conditional Heteroscedasticity
with Wolfgang Schmid and Robert Garthoff,
arXiv:1609.00711

Thank you for your attention!

ID 508: Philipp Otto "Spatial Modelling of Mortality Rates with Heteroscedastic Residuals"

Spatial and Spatiotemporal Autoregressive Conditional Heteroscedasticity
with Wolfgang Schmid and Robert Garthoff,
arXiv:1609.00711

Literaturverzeichnis I

Bera, Anil K und Simlai, Pradosh. 2004. "Testing for Spatial Dependence and a Formulation of Spatial ARCH (SARCH) Model with Applications." Technical report, Working paper University of Illinois.

Bollerslev, Tim. 1986. "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics*, 31(3): 307–327.

Borovkova, Svetlana und Lopuhaa, Rik. 2012. "Spatial GARCH: A Spatial Approach to Multivariate Volatility Modeling." *Available at SSRN 2176781*.

Caporin, Massimiliano und Paruolo, Paolo. 2006. "GARCH models with spatial structure." *SIS Statistica*: 447–450.

Engle, Robert F. 1982. "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation." *Econometrica: Journal of the Econometric Society*: 987–1007.

Noiboar, Amir und Cohen, Israel. 2005. "Two-dimensional GARCH model with application to anomaly detection." *13th European Signal Processing Conf., Istanbul, Turkey*, IEEE, 1–4.

U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute. 2015. "United States Cancer Statistics 1999-2012 Incidence and Mortality Web-based Report."

Example ($n = 2$)

- Simple case for $n = 2 \rightsquigarrow$ two locations $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$

$$\boldsymbol{Y} \circ \boldsymbol{Y} = \begin{pmatrix} Y(\boldsymbol{s}_1)^2 \\ Y(\boldsymbol{s}_2)^2 \end{pmatrix} = \begin{pmatrix} h(\boldsymbol{s}_1)\,\varepsilon(\boldsymbol{s}_1)^2 \\ h(\boldsymbol{s}_2)\,\varepsilon(\boldsymbol{s}_2)^2 \end{pmatrix} .$$

- Taking a closer look on $Y(\boldsymbol{s}_1)^2$ one may see that

$$\begin{aligned} Y(\boldsymbol{s}_1)^2 &= \varepsilon(\boldsymbol{s}_1)^2 \left(\alpha_1 + w_{12} Y(\boldsymbol{s}_2)^2\right) \\ &= \alpha_1 \varepsilon(\boldsymbol{s}_1)^2 + w_{12} \varepsilon(\boldsymbol{s}_1)^2 \underbrace{h(\boldsymbol{s}_2)}_{=\alpha_2 + w_{21} Y(\boldsymbol{s}_1)^2} \varepsilon(\boldsymbol{s}_2)^2 \\ &= \alpha_1 \varepsilon(\boldsymbol{s}_1)^2 + w_{12} w_{21} \varepsilon(\boldsymbol{s}_1)^2 \varepsilon(\boldsymbol{s}_2)^2 Y(\boldsymbol{s}_1)^2 + \\ &\quad \alpha_2 w_{12} \varepsilon(\boldsymbol{s}_1)^2 \varepsilon(\boldsymbol{s}_2)^2 . \end{aligned}$$

- Thus, $Y(\boldsymbol{s}_1)^2 \geq 0$, if and only if $\boldsymbol{\alpha} \geq \mathbf{0}$, $w_{12} > 0$, $w_{21} > 0$ and

$$\varepsilon(\boldsymbol{s}_1)^2 \varepsilon(\boldsymbol{s}_2)^2 < \frac{1}{w_{12} w_{21}} .$$

Example ($n = 2$)

- Simple case for $n = 2 \rightsquigarrow$ two locations $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$

$$\boldsymbol{Y} \circ \boldsymbol{Y} = \begin{pmatrix} Y(\boldsymbol{s}_1)^2 \\ Y(\boldsymbol{s}_2)^2 \end{pmatrix} = \begin{pmatrix} h(\boldsymbol{s}_1)\,\varepsilon(\boldsymbol{s}_1)^2 \\ h(\boldsymbol{s}_2)\,\varepsilon(\boldsymbol{s}_2)^2 \end{pmatrix} .$$

- Taking a closer look on $Y(\boldsymbol{s}_1)^2$ one may see that

$$\begin{aligned} Y(\boldsymbol{s}_1)^2 &= \varepsilon(\boldsymbol{s}_1)^2 \left(\alpha_1 + w_{12} Y(\boldsymbol{s}_2)^2\right) \\ &= \alpha_1 \varepsilon(\boldsymbol{s}_1)^2 + w_{12} \varepsilon(\boldsymbol{s}_1)^2 \underbrace{h(\boldsymbol{s}_2)}_{=\alpha_2 + w_{21} Y(\boldsymbol{s}_1)^2} \varepsilon(\boldsymbol{s}_2)^2 \\ &= \alpha_1 \varepsilon(\boldsymbol{s}_1)^2 + w_{12} w_{21} \varepsilon(\boldsymbol{s}_1)^2 \varepsilon(\boldsymbol{s}_2)^2 Y(\boldsymbol{s}_1)^2 + \\ &\quad \alpha_2 w_{12} \varepsilon(\boldsymbol{s}_1)^2 \varepsilon(\boldsymbol{s}_2)^2 . \end{aligned}$$

- Thus, $Y(\boldsymbol{s}_1)^2 \geq 0$, if and only if $\boldsymbol{\alpha} \geq \boldsymbol{0}$, $w_{12} > 0$, $w_{21} > 0$ and

$$\varepsilon(\boldsymbol{s}_1)^2 \varepsilon(\boldsymbol{s}_2)^2 < \frac{1}{w_{12} w_{21}} .$$

Example ($n = 2$)

- Simple case for $n = 2 \rightsquigarrow$ two locations $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$

$$\boldsymbol{Y} \circ \boldsymbol{Y} = \begin{pmatrix} Y(\boldsymbol{s}_1)^2 \\ Y(\boldsymbol{s}_2)^2 \end{pmatrix} = \begin{pmatrix} h(\boldsymbol{s}_1)\,\varepsilon(\boldsymbol{s}_1)^2 \\ h(\boldsymbol{s}_2)\,\varepsilon(\boldsymbol{s}_2)^2 \end{pmatrix} .$$

- Taking a closer look on $Y(\boldsymbol{s}_1)^2$ one may see that

$$\begin{aligned} Y(\boldsymbol{s}_1)^2 &= \varepsilon(\boldsymbol{s}_1)^2 \left(\alpha_1 + w_{12} Y(\boldsymbol{s}_2)^2\right) \\ &= \alpha_1 \varepsilon(\boldsymbol{s}_1)^2 + w_{12} \varepsilon(\boldsymbol{s}_1)^2 \underbrace{h(\boldsymbol{s}_2)}_{=\alpha_2 + w_{21} Y(\boldsymbol{s}_1)^2} \varepsilon(\boldsymbol{s}_2)^2 \\ &= \alpha_1 \varepsilon(\boldsymbol{s}_1)^2 + w_{12} w_{21} \varepsilon(\boldsymbol{s}_1)^2 \varepsilon(\boldsymbol{s}_2)^2 Y(\boldsymbol{s}_1)^2 + \\ &\quad \alpha_2 w_{12} \varepsilon(\boldsymbol{s}_1)^2 \varepsilon(\boldsymbol{s}_2)^2 . \end{aligned}$$

- Thus, $Y(\boldsymbol{s}_1)^2 \geq 0$, if and only if $\boldsymbol{\alpha} \geq \mathbf{0}$, $w_{12} > 0$, $w_{21} > 0$ and

$$\varepsilon(\boldsymbol{s}_1)^2 \varepsilon(\boldsymbol{s}_2)^2 < \frac{1}{w_{12} w_{21}} .$$

- Let

$$\begin{aligned}\boldsymbol{\eta} &= \left(\alpha_i \varepsilon(\boldsymbol{s}_i)^2 + \varepsilon(\boldsymbol{s}_i)^2 \sum_{v=1}^{n} w_{iv} \varepsilon(\boldsymbol{s}_v)^2 \alpha_v\right)_{i=1,\ldots,n} \\ \mathbf{A} &= \text{diag}\left(\varepsilon(\boldsymbol{s}_1)^2, \ldots, \varepsilon(\boldsymbol{s}_n)^2\right) \mathbf{W} \\ \boldsymbol{Y}^{(2)} &= (Y(\boldsymbol{s}_1)^2, \ldots, Y(\boldsymbol{s}_n)^2)'.\end{aligned}$$

Theorem

Suppose that

$$\det\left(\mathbf{I} - \boldsymbol{A}^2\right) \neq 0\,, \qquad (3)$$

then $Y(\boldsymbol{s}_1)^2, \ldots, Y(\boldsymbol{s}_n)^2$ *are uniquely determined by* $\varepsilon(\boldsymbol{s}_1)^2, \ldots, \varepsilon(\boldsymbol{s}_n)^2$*. It holds that*

$$\boldsymbol{Y}^{(2)} = \left(\mathbf{I} - \mathbf{A}^2\right)^{-1} \boldsymbol{\eta} \quad \text{and} \quad \boldsymbol{h} = \boldsymbol{\alpha} + \mathbf{W}\left(\mathbf{I} - \mathbf{A}^2\right)^{-1} \boldsymbol{\eta}\,. \qquad (4)$$

- Let

$$
\begin{aligned}
\boldsymbol{\eta} &= \left(\alpha_i \varepsilon(\boldsymbol{s}_i)^2 + \varepsilon(\boldsymbol{s}_i)^2 \sum_{v=1}^{n} w_{iv} \varepsilon(\boldsymbol{s}_v)^2 \alpha_v\right)_{i=1,\ldots,n} \\
\mathbf{A} &= \text{diag}\left(\varepsilon(\boldsymbol{s}_1)^2, \ldots, \varepsilon(\boldsymbol{s}_n)^2\right) \mathbf{W} \\
\boldsymbol{Y}^{(2)} &= (Y(\boldsymbol{s}_1)^2, ..., Y(\boldsymbol{s}_n)^2)'.
\end{aligned}
$$

Theorem

Suppose that

$$\det\left(\mathbf{I} - \boldsymbol{A}^2\right) \neq 0\,, \tag{3}$$

then $Y(\boldsymbol{s}_1)^2, \ldots, Y(\boldsymbol{s}_n)^2$ *are uniquely determined by* $\varepsilon(\boldsymbol{s}_1)^2, \ldots, \varepsilon(\boldsymbol{s}_n)^2$. *It holds that*

$$\mathbf{Y}^{(2)} = \left(\mathbf{I} - \mathbf{A}^2\right)^{-1} \boldsymbol{\eta} \quad \textit{and} \quad \boldsymbol{h} = \boldsymbol{\alpha} + \mathbf{W}\left(\mathbf{I} - \mathbf{A}^2\right)^{-1} \boldsymbol{\eta}\,. \tag{4}$$

Literature

- Engle (1982); Bollerslev (1986) introduce ARCH/GARCH model for time series analysis
- Further extensions, generalizations
- Including multivariate, temporal GARCH models (e.g. CCC, DCC models)
- Borovkova und Lopuhaa (2012) and Caporin und Paruolo (2006) propose „spatial GARCH mode"
  - time series model
  - first temporal lag incorporates neighboring observations
- Bera und Simlai (2004) develop **spatial ARCH model** of order 1 (SARCH(1)) as „side effect"
- Noiboar und Cohen (2005): multidimensional GARCH-type model for detection of image anomalies

However,

1. there is no general approach, but the proposed models only deal with some special cases.
2. there is no theoretical analysis of the models.

Literature

- Engle (1982); Bollerslev (1986) introduce ARCH/GARCH model for time series analysis
- Further extensions, generalizations
- Including multivariate, temporal GARCH models (e.g. CCC, DCC models)
- Borovkova und Lopuhaa (2012) and Caporin und Paruolo (2006) propose „spatial GARCH mode"
  - time series model
  - first temporal lag incorporates neighboring observations
- Bera und Simlai (2004) develop **spatial ARCH model** of order 1 (SARCH(1)) as „side effect"
- Noiboar und Cohen (2005): multidimensional GARCH-type model for detection of image anomalies

However,

1. there is no general approach, but the proposed models only deal with some special cases.
2. there is no theoretical analysis of the models.

Literature

- Engle (1982); Bollerslev (1986) introduce ARCH/GARCH model for time series analysis
- Further extensions, generalizations
- Including multivariate, temporal GARCH models (e.g. CCC, DCC models)
- Borovkova und Lopuhaa (2012) and Caporin und Paruolo (2006) propose „spatial GARCH mode"
  - time series model
  - first temporal lag incorporates neighboring observations
- Bera und Simlai (2004) develop **spatial ARCH model** of order 1 (SARCH(1)) as „side effect"
- Noiboar und Cohen (2005): multidimensional GARCH-type model for detection of image anomalies

However,

1. there is no general approach, but the proposed models only deal with some special cases.
2. there is no theoretical analysis of the models.

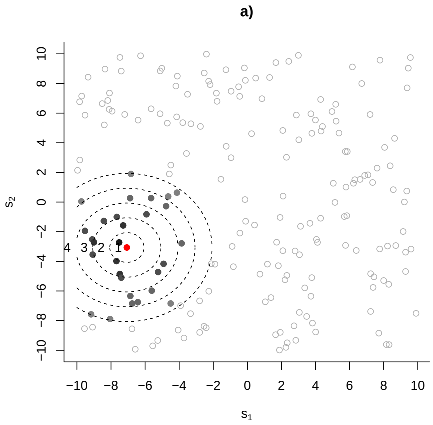

Figure: Representation of the positive elements in $\mathbf{W}$ (colored in grey, filled dots) for some location i (colored in red) regarding a) a spARCH(5) process and b) a radial process with center s_0.

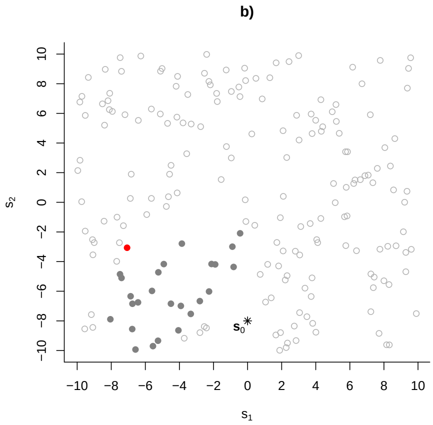

Figure: Representation of the positive elements in $\mathbf{W}$ (colored in grey, filled dots) for some location i (colored in red) regarding a) a spARCH(5) process and b) a radial process with center s_0.

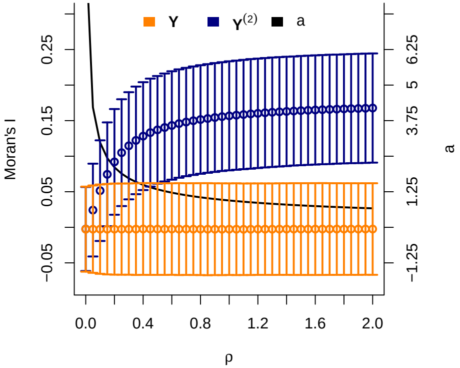

Figure: Moran's I of the observations $\boldsymbol{Y}$ and the squared observations $\boldsymbol{Y}^{(2)}$, including the asymptotic 95% confidence intervals of I for $\rho \in \{0, 0.05, \ldots, 2\}$. Moreover, the resulting bound a is plotted as a bold, black line.

Properties

- If the distribution of ε is sign-symmetric, the distribution of $\boldsymbol{Y}$ is sign-symmetric as well.
- If *either* ε is sign-symmetric and $\mathbf{W}$ is an upper or lower triangular matrix with non-negative elements and the $8r[(n-1)/2]$th moment of $\varepsilon(\boldsymbol{s}_i)$ exists *or* $||\mathbf{A}^2|| \leq \lambda < 1$ P-a.e. and the $2r$th moment of $\varepsilon(\boldsymbol{s}_i)$ exists then

$$E(Y(\boldsymbol{s}_i)^{2r}) < \infty \,,$$

$$E(Y(\boldsymbol{s}_i)^{2v-1}) = 0,\; E(Y(\boldsymbol{s}_i)^{2v-1}|Y(\boldsymbol{s}_j), j \neq i) = 0,\; v = 1, \ldots, r \,.$$

- If $\mathbf{W}$ is an upper or lower triangular matrix with non-negative elements, then

$$h(\boldsymbol{s}_i) = E(Y(\boldsymbol{s}_i)^2|Y(\boldsymbol{s}_j), j \neq i) \,.$$

- If $\mathbf{W}$ is neither an upper nor a lower triangular matrix, this result is not valid.

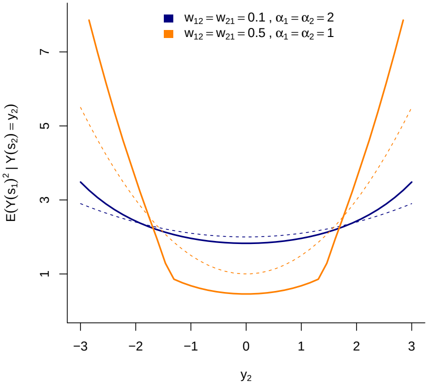

Figure: Conditional expectation of $Y(\boldsymbol{s}_1)^2$ given y_2 for $n = 2$, where $E(Y(\boldsymbol{s}_1)^2 | Y(\boldsymbol{s}_2) = y_2^2)$ is plotted as a solid line and $h(\boldsymbol{s}_1)$ as a dashed line.

Simulation and Visualization

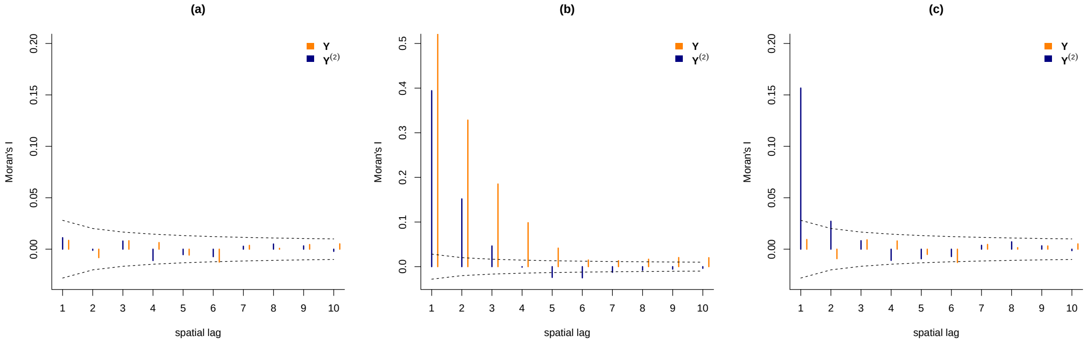

Figure: Spatial autocorrelation function (ACF) of the (a) Spatial White Noise, (b) the spatial autoregressive process, and (c) the spatial ARCH process.