Bayesian Auto-regressive Latent Growth Model

Spatial Growth Model

Zachary Roman Ph.D. - Zachary.Roman@psychologie.uzh.ch - Stan Connect 2022

Contents

Road Map

Spatial Models and Behavioral Sciences

Psychologists are getting better

1. Psychologists have been slow to adapt spatial models
2. Often ignore independence violations due to space
3. Social Network Auto-Regressive (SNAR) model (Valente et al., 1997)

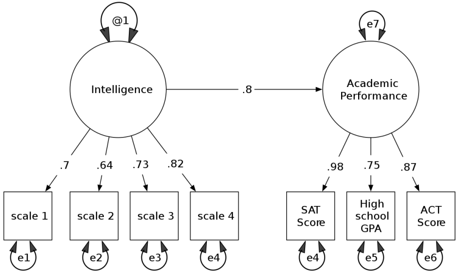
- SAR model with different conceptualization of W
- $y = \alpha + \rho \boldsymbol{W} y + \beta X + \epsilon$
- Connections in W represent “social distance”
- Strength of influence over another person (e.g., opinions)

Traditional Structural Equation Models

See (Bollen, 1989) social for more on SEM.

$$\eta = \alpha + \gamma \boldsymbol{\xi} + \zeta, \quad \zeta \sim N(0, \sigma_\zeta^2) \tag{1}$$

$$\begin{aligned} y_j &= \tau_{yj} + \lambda_{yj}\eta + \epsilon_j, & \epsilon_j &\sim N(0, \sigma_\epsilon^2) \\ x_k &= \tau_{xk} + \boldsymbol{\lambda}_{xk}\boldsymbol{\xi} + \delta_k, & \delta_k &\sim N(0, \sigma_\delta^2) \end{aligned} \tag{2}$$



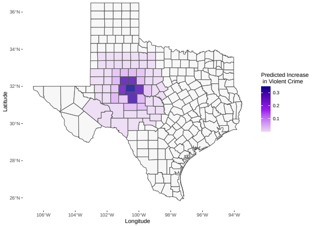
Bayesian Auto-Regressive Dependence Structural Equation Models (BARDSEM) (Roman & Brandt, 2021)

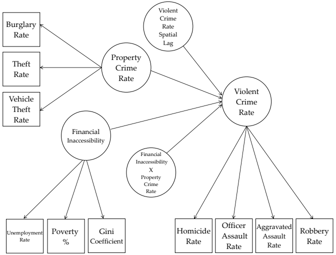
$$\eta = \alpha + \rho \mathbf{W} \eta + \gamma \boldsymbol{\xi} + \boldsymbol{\xi} + \zeta, \quad \zeta \sim N(0, \sigma^2_\zeta) \tag{3}$$

$$\begin{aligned} y_j &= \tau_{yj} + \lambda_{yj} \eta + \epsilon_j, & \epsilon_j &\sim N(0, \sigma^2_\epsilon) \\ x_k &= \tau_{xk} + \boldsymbol{\lambda}_{xk} \boldsymbol{\xi} + \delta_k, & \delta_k &\sim N(0, \sigma^2_\delta) \end{aligned} \tag{4}$$

- CPDM formula

$$\partial_\eta / \partial_\xi^{q\prime} = (\mathbf{I}_N - \rho \mathbf{W})^{-1} \boldsymbol{I}_N \gamma_m \tag{5}$$





Latent Growth Model

$$\mathbf{\Lambda}_x = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ \vdots & \vdots & \vdots \\ 1 & p-1 & (p-1)^2 \end{bmatrix} \tag{6}$$

$$\mathbf{\Psi} = \begin{bmatrix} \psi_{11} & & \\ \psi_{21} & \psi_{22} & \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix} \tag{7}$$

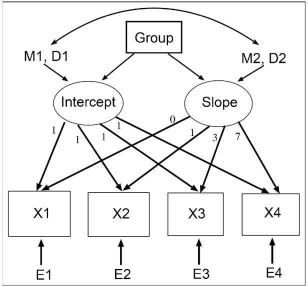
$$\mathbf{\Theta}_\epsilon = \begin{bmatrix} \Theta_\epsilon & & & & \\ 0 & \Theta_\epsilon & & & \\ 0 & 0 & \Theta_\epsilon & & \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \dots & \Theta_{\epsilon m} \end{bmatrix} \tag{8}$$

Sometimes we may want to take the "person" level estimates (latent slopes, latent intercepts etc.) and predict something with them.

$$Y = \alpha + \beta_k \boldsymbol{\xi_k} + \zeta, \quad \zeta \sim N(0, \sigma^2_\zeta) \tag{9}$$

Other times we may want to predict the slopes with covariates (Z).

$$\boldsymbol{\xi_s} = \alpha + \beta_k Z + \zeta, \quad \zeta \sim N(0, \sigma^2_\zeta) \tag{10}$$



Spatial Latent Growth

Growth as a predictor

$$Y = \alpha + \beta_k \boldsymbol{\xi_k} + \rho \boldsymbol{W} Y + \zeta, \quad \zeta \sim N(0, \sigma_\zeta^2) \tag{11}$$

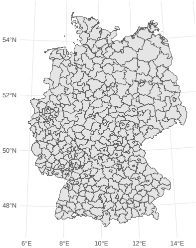
Growth as an outcome

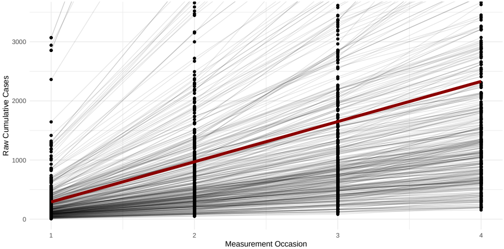
$$\xi_s = \alpha + \beta_k Z + \rho \boldsymbol{W} \xi_s + \zeta, \quad \zeta \sim N(0, \sigma_\zeta^2) \tag{12}$$

German Covid Data Example

Cases in the data are 423 “Stadtkreis” and “Landkreis” regions outlined in the map.

Modeling new cases, positive slope over time shows that the rate of infection is increasing.





- Covid cases are cumulative, in 14 day intervals.
- In this example we have 4 “measurement occasions” (MO).
- Intervals (yyy-mm-dd) :
- 2020-10-11 to 2020-10-25 (MO 1)
- 2020-10-26 to 2020-11-08 (MO 2)
- 2020-11-09 to 2020-11-22 (MO 3)
- 2020-11-23 to 2020-12-06 (MO 4)

Stan Specification and Prios

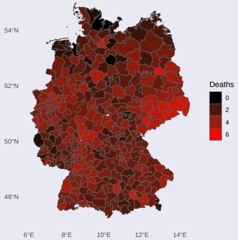
```
data {
 int<lower =0 > N;
 int<lower =0 > Kx;
 matrix [N, Kx] x;
 vector[N] Y;
 vector[N] pop1000;
 matrix<lower = 0, upper = 1>[N, N] W;
}

parameters {
 vector<lower=0>[Kx] sigmax;
 vector<lower=0>[2] sigmaxi;
 real<lower=0> sigma;  //Structural Error
 cholesky_factor_corr[2] L1;
 matrix [N ,2] zi;
 vector[2] muxi1; // Means of Int and Slope
 real beta1; //Slope of Cases predicting Death
 real beta2; //Slope of Pop1000 (covariate) predicting death
 real rho;
 real alpha;
}

model {
 matrix [N, Kx] mux ; //E [ x | xi ]
 matrix [N, 2] muxivector; //Means of Slope xi[2] and Int xi[1]
 matrix [N ,2] xi; //Xi[,1] = intercept, xi[,2] = slope

for(n in 1:N){for(j in 1:2){muxivector[n,j]=muxi1[j];}}
//Cholesky Decomp for xi
xi = muxivector + zi * diag_pre_multiply(sigmaxi,L1)';
```

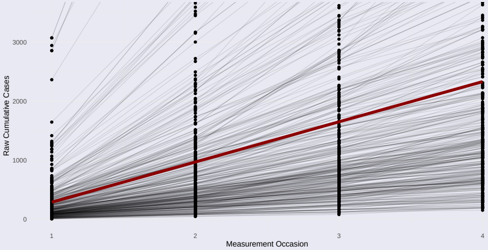
```
for(i in 1:Kx){
 mux [,i] = 1*xi[,1] + // Intercept
            (i-1)*xi[,2]; // Linear Slope
}

for(z in 1:Kx){x[,z] ~ normal(mux[,z], sigmax[z]);}
  to_vector(zi) ~ normal(0,1);
  sigmax ~ cauchy(0,2.5); //errors
  sigma ~ cauchy(0,2.5);
  sigmaxi ~ cauchy(0,2.5);
  L1 ~ lkj_corr_cholesky(2);
  muxi1[1] ~ normal(0,3); //Mean of Intercept
  muxi1[2] ~ normal(0,3); //Mean of Slope
  rho ~ uniform(-1,1);    //Spatial term
  beta1 ~ normal(0,2);    //Slope of case slope predicting death
  beta2 ~ normal(0,2);    //Slope of pop covariate
  Y ~ normal(alpha + beta1*xi[,2] + //Structural model
                   beta2*pop1000 + //Pop covariate
                   rho*(W*Y), sigma); //Spatial lag
}

generated quantities{
  matrix[2,2] phi; //Varcov matrix of Int and Slope Factors.
  phi = diag_pre_multiply(sigmaxi,L1)*diag_pre_multiply(sigmaxi,L1)'
}
```

Deaths

$$Y_{death5} = \alpha + \rho_{death} \boldsymbol{W} Y_{death5} + \beta \xi_s + \zeta \tag{13}$$

Regional Deaths at T5





Inputs

4,000 itterations, half warmup, 2 chains, no thinning

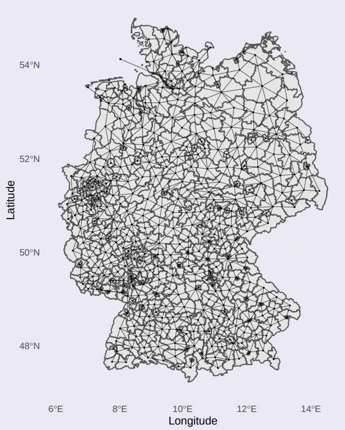
Parameter Estimates

Summary

	mean	se_mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
muxi1[1]	2.99	0.05	0.28	2.45	2.82	2.98	3.16	3.57	27.78	1.07
muxi1[2]	1.97	0.02	0.15	1.67	1.87	1.98	2.07	2.27	37.70	1.04
alpha	-3.34	0.07	1.62	-6.42	-4.47	-3.34	-2.23	-0.26	505.84	1.00
beta1	9.94	0.06	0.61	8.78	9.51	9.93	10.34	11.17	99.74	1.03
beta2	0.02	0.00	0.01	0.01	0.02	0.02	0.02	0.03	459.37	1.01
rho	0.33	0.00	0.03	0.26	0.31	0.33	0.36	0.40	1100.75	1.00
phi[1,1]	30.87	0.15	2.11	26.66	29.42	30.89	32.32	35.06	212.60	1.02
phi[2,1]	15.20	0.07	1.08	13.11	14.48	15.23	15.92	17.36	262.75	1.02
phi[2,2]	9.56	0.06	0.67	8.25	9.10	9.57	10.03	10.85	142.66	1.05

Weight Matrix Specification

Row normalized first order contiguity




```
##      [,1] [,2] [,3] [,4]
## 1       0    1    1    0
## 2       1    0    1    0
## 3       1    1    0    1
## 4       0    0    1    0
```

Interpretation

Spillover

The model implies cases deaths impact deaths of neighboring regions.
To compute predicted values, and understand the models interpretations, these interactions must be considered.
We account for them with spillover computation of the cross-partial derivatives matrix:

$$CPDM = (\mathbf{I}_N - \hat{\rho}\mathbf{W})^{-1} \boldsymbol{I}_N \hat{\beta}_x \tag{14}$$

The larger the slope (β), spatial effect (ρ), and connections to neighbors ($\boldsymbol{W}$) the larger the elements of the CPDM will be.

```
##              stuttgart boeblingen esslingen
## stuttgart        10.16       0.99      0.95
## boeblingen        0.57      10.17      0.56
## esslingen         0.76       0.79     10.15
```

Impacts

Direct Spillover (Impact)

- An increase in the growth rate $(+100)$ of covid cases in region i, corresponds with an expected increase of 10.18 deaths on average in all $\neq i$ regions.

Indirect Spillover (Impact)

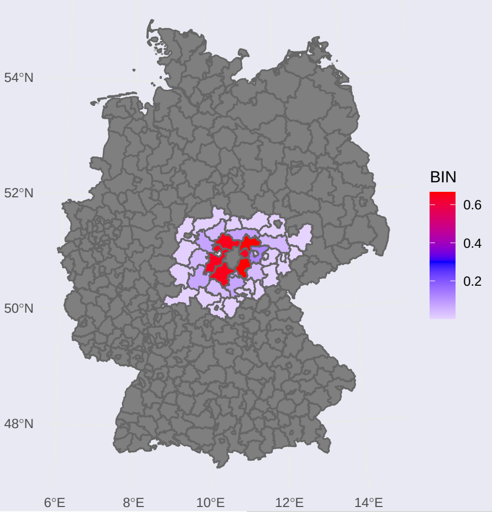
- An increase in the growth rate of covid cases in regions $\neq i$, corresponds with an expected increase of 4.72 in region i.

Total Spillover (Impact)

- An increase in the growth rate of covid cases in all regions, corresponds with an expected average increase of 14.9 in all regions.

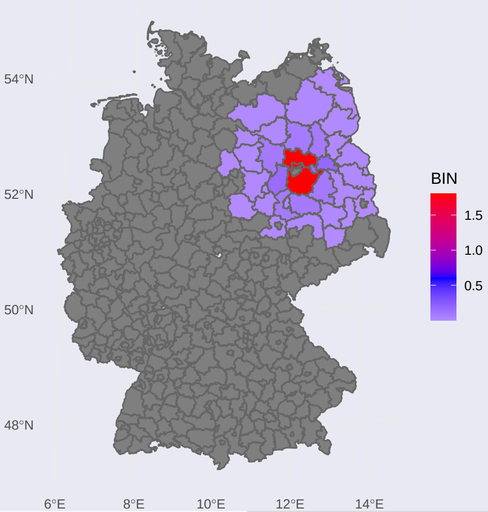
Spillover for cases of interest

Gotha



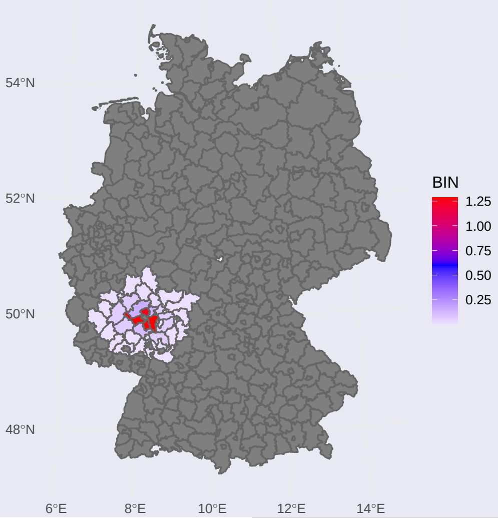
Spillover for cases of interest

Brandenburg An Der Havel



Spillover for cases of interest

Mainz

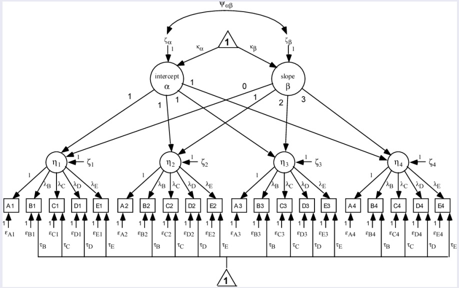


Future Directions

Spatial adaptions

- Spatial error lag
- Spatial predictor lag

Latent variable adaptions



References

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Roman, Z. J., & Brandt, H. (2021). A latent auto-regressive approach for bayesian structural equation modeling of spatially or socially dependent data. *Multivariate Behavioral Research*, 1–25. https://doi.org/10.1080/00273171.2021.1957663

Valente, T. W., Watkins, S. C., Jato, M. N., Van Der Straten, A., & Tsitsol, L.-P. M. (1997). Social network associations with contraceptive use among cameroonian women in voluntary associations. *Social Science & Medicine*, *45*(5), 677–687. https://doi.org/10.1016/S0277-9536(96)00385-1