# Spatial Analysis

March 4th, 2024

## Introduction

### Data loading and processing

```
rm(list = ls()) # remove all variables in the current workspace
library(sf); library(tmap); library(spdep); library(spatialreg); library(tidyverse)
setwd('~/Scrivania/geospatial/project')
```

Countries that are to be considered for the final analysis that were filtered according the availability of data.

```
common  <- c('CMR', 'GNB', 'ERI', 'MAR', 'LBY', 'ETH', 'GHA', 'DZA',
             'SWZ', 'TZA', 'ZAF', 'AGO', 'GNQ', 'GAB', 'MLI', 'TCD',
             'COG', 'NGA', 'BEN', 'CAF', 'TGO', 'GIN', 'GMB', 'MRT',
             'BWA', 'COD', 'SSD', 'ZWE', 'KEN', 'MWI', 'RWA', 'BDI',
             'NER', 'LBR', 'BFA', 'CIV', 'SEN', 'EGY', 'ZMB', 'NAM',
             'TUN', 'SOM', 'LSO', 'SDN', 'SLE', 'MOZ', 'UGA')
```

Loading the shapefile and filtering according to the countries "ISO" codes.

With `st_make_valid` the polygons' geometry are checked and fixed in order to correctly estimate the spatial lag matrix by means of contiguity (countries that share a common border).

```
africa <- st_make_valid(africa)
```

Neighbours `nb` object had been estimated by listing all contiguous countries that have a border in common.

```
nb  <- spdep::poly2nb(africa)
nbW <- nb2listw(nb, style="W", zero.policy = TRUE)
```

Predictors and the dependent variable are loaded.

```
predictors <- read.csv("./data/event_predictors.csv")
predictors <- predictors[order(predictors$ISO),]

stability <- read.csv("./data/stability_indexes.csv")
stability <- stability[order(stability$ISO),]
```

We finally defined the formula that will be used as follows in order to test Spatial autoregressive models. Specifically, the current purpose is to define a function that maps the relation between the amount of different event type occurring in each country within a given one-year span to the stability index.

```
formula <- stability ~ EventRoot5 + EventRoot6 + EventRoot7 +
  EventRoot10 + EventRoot12 + EventRoot13 + EventRoot14 +
  EventRoot19 + AvgTone + GoldsteinScale
```

# Analysing events occurring in 2020

## Subset selection according to `year`

Subset of records belonging to the current year under examination are filtered and merged into a final dataframe object (`df`).

```
year <- "2020"
pred <- predictors[predictors$Year==year,]
stab <- stability[stability$Year==year,]
df <- list(pred, stab) %>% reduce(full_join, by='ISO')
```

## Moran Index

We need first to evaluate whether the dependent variable may show some spatial correlation pattern over the current lag matrix representation. A Moran Index test is executed showing the following result:

```
moran.test(df$stability, nbW, randomisation=FALSE)
```

```
##
##  Moran I test under normality
##
## data:  df$stability
## weights: nbW
## n reduced by no-neighbour observations
##
## Moran I statistic standard deviate = 2.9674, p-value = 0.001501
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##        0.28939028       -0.02222222        0.01102717
```

A significant positive spatial autocorrelation is found (*Moran I = 0.28, p-value = 0.001*): this may allow to assume that a spatial dependency is occurring.

## Ordinary Least Square (OLS) estimation and spatial dependency

According to Elhorst (2010) procedure we need first to test whether the OLS model estimated below is found to yield a significant result in Lagrange Multiplier test (LM-test). This allow to proceed with the estimation of regression models considering also predictors' spatial components. A linear model is estimated and coefficients were computed according to the standard Least Square approach: coefficients and their relative level of significance are shown below:

```
l <- lm(formula, data=df)
summary(l)
```

```
##
## Call:
## lm(formula = formula, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.55674 -0.22889  0.01733  0.27483  1.49595
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -2.0813573  1.3137547  -1.584   0.1219
```

```
## EventRoot5      -0.0002377  0.0003028  -0.785   0.4376
## EventRoot6       0.0004174  0.0008641   0.483   0.6320
## EventRoot7       0.0010030  0.0007978   1.257   0.2168
## EventRoot10      0.0016422  0.0012640   1.299   0.2021
## EventRoot12      0.0007508  0.0013223   0.568   0.5737
## EventRoot13     -0.0030164  0.0023472  -1.285   0.2070
## EventRoot14     -0.0002528  0.0011634  -0.217   0.8292
## EventRoot19     -0.0005802  0.0002563  -2.264   0.0297 *
## AvgTone         -0.3498549  0.3806955  -0.919   0.3642
## GoldsteinScale   0.9956214  0.8990292   1.107   0.2754
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6885 on 36 degrees of freedom
## Multiple R-squared:  0.4521, Adjusted R-squared:    0.3
## F-statistic: 2.971 on 10 and 36 DF,  p-value: 0.007901
```

The only predictor that shows a significant effect is the amount of GDELT event associated to the `EventRootCode = 19` (mass-violence-related events). Specifically, the amount of violence/coercion-based events seem to have an impact on the stability index by a negative trend ($\beta = -0.5 \times 10^{-3}$).

The *Lagrange Multiplier test* (LM test) is then applied with the purpose of detecting which statistic may better explain any error or lag component (local or global spatial spillover effect respectively). As shown below a significant result (*p-value < 0.05*) over the `adjRSlag` statistic which refers to the test for lag dependence in the possible presence of a missing lagged dependent variable is found. This will further enforce the idea that the spatial effect may be described globally with autoregressive models as described in the following sections.

```
LMtest <- spdep::lm.RStests(l, nbW, zero.policy=TRUE, test="all")
summary(LMtest)
```

```
##  Rao's score (a.k.a Lagrange multiplier) diagnostics for spatial
##  dependence
## data:
## model: lm(formula = formula, data = df)
## test weights: nbW
##
##           statistic parameter p.value
## RSerr     0.0068881         1 0.93386
## RSlag     1.6116569         1 0.20426
## adjRSerr  3.5263398         1 0.06040 .
## adjRSlag  5.1311086         1 0.02350 *
## SARMA     5.1379967         2 0.07661 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Spatial autoregressive global spillover models

### Spatial Durbin Model (SDM)

Following the Elhorst criteria a SDM is estimated in order to test whether $\rho$ and $\theta$ coefficients will be considered significantly different from zero to further confirm this model against a simpler global spatial spillover model.

```
sdm <- lagsarlm(stability ~ EventRoot5 + EventRoot6 + EventRoot7 +
                EventRoot10 + EventRoot12 + EventRoot13 +
                EventRoot14 + EventRoot19 + AvgTone +
                GoldsteinScale, data = df,
```

```
                listw=nbW, type="mixed",
                zero.policy=TRUE)
#summary(sdm)
```

**Spatial Autoregressive Model (SAR)**

A SAR model is also estimated in order to be compared.

```
sar <- lagsarlm(stability ~ EventRoot5 + EventRoot6 + EventRoot7 +
                EventRoot10 + EventRoot12 + EventRoot13 +
                EventRoot14 + EventRoot19 + AvgTone +
                GoldsteinScale, data = df,
                listw=nbW,
                zero.policy=TRUE)
```

**Testing The Likelihood Ratio test (LRT)**

```
anova(sar, sdm)
```

```
##       Model df    AIC  logLik Test L.Ratio p-value
## sar       1 13 109.35 -41.673    1
## sdm       2 23 113.82 -33.908    2  15.529 0.11393
```

According to the result (*LRT p-value > 0.05*) the Null hypothesis ($H_0 : \theta = 0$) can not be rejected, meaning that the SDR model is not suitable to provide a good approximation of the global spatial spillover effect given that $\theta$ had been estimated to be not significantly different from zero.

Therefore, the SAR model is considered to be more suited to explain the spatial dependency. SAR-related coefficients are displayed as follows:

```
summary(sar)
```

```
##
## Call:lagsarlm(formula = stability ~ EventRoot5 + EventRoot6 + EventRoot7 +
##     EventRoot10 + EventRoot12 + EventRoot13 + EventRoot14 + EventRoot19 +
##     AvgTone + GoldsteinScale, data = df, listw = nbW, zero.policy = TRUE)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.537249 -0.254720  0.026916  0.352129  1.308348
##
## Type: lag
## Regions with no neighbours included:
##   28
## Coefficients: (asymptotic standard errors)
##                   Estimate  Std. Error z value Pr(>|z|)
## (Intercept)    -1.6142e+00  1.1224e+00 -1.4381  0.15041
## EventRoot5     -2.1740e-04  2.5612e-04 -0.8488  0.39597
## EventRoot6      2.5613e-05  7.3535e-04  0.0348  0.97221
## EventRoot7      1.2177e-03  6.7312e-04  1.8091  0.07044
## EventRoot10     7.8366e-04  1.0872e-03  0.7208  0.47102
## EventRoot12     8.3557e-04  1.1160e-03  0.7487  0.45402
## EventRoot13    -2.7916e-03  1.9797e-03 -1.4101  0.15850
## EventRoot14     7.9381e-05  9.8499e-04  0.0806  0.93577
## EventRoot19    -5.2319e-04  2.1730e-04 -2.4077  0.01605
## AvgTone        -3.3102e-01  3.2108e-01 -1.0310  0.30256
```

```
## GoldsteinScale  6.5745e-01  7.6716e-01  0.8570  0.39145
##
## Rho: 0.30582, LR test value: 2.4247, p-value: 0.11944
## Asymptotic standard error: 0.15259
##     z-value: 2.0041, p-value: 0.045054
## Wald statistic: 4.0166, p-value: 0.045054
##
## Log likelihood: -41.67283 for lag model
## ML residual variance (sigma squared): 0.33722, (sigma: 0.58071)
## Number of observations: 47
## Number of parameters estimated: 13
## AIC: 109.35, (AIC for lm: 109.77)
## LM test for residual autocorrelation
## test value: 0.90131, p-value: 0.34243
```

The only predictor that turned out to be significant is `EventRoot19` ($p$-value $< 0.05$) and its coefficient was found to be negative ($\beta = -5.23 \times 10^{-5}$) meaning that a negative change in one unit of the predictor can be associated to a positive Stability index increment measured in that country. This result can be explained in light of the LR test results computed over the spatial lag parameter $\rho$ for which it can be stated that the Null hypothesis ($H_0 : \rho = 0$) is accepted and thus the inclusion of the lagged values seems not to improve the model. Consequently, given that there is no sufficient evidence regarding the local spatial dependency from the Lagrange Multiplier test ($p_{adjRSerr} = 0.06$, the adjusted statistic referring to the Rao's Score test for error dependence) and since the Null hypotheses $H_0 : \theta = \rho = 0$ has been accepted, one may rely on the simpliest Linear Regression to better describe the relationship between the yearly event counts and the stability index.

To conclude, AIC criteria are displayed in order to compare these models according to their goodness of fit.

```
print("AIC:")
```

```
## [1] "AIC:"
```

```
print(paste("OLS:",round(AIC(l),3),
            " SDM:",round(AIC(sdm),3),
            " SAR:",round(AIC(sar),3),  sep=" "))
```

```
## [1] "OLS: 109.77  SDM: 113.816  SAR: 109.346"
```

It can be noticed how AIC estimated from the OLS and SAR models is similar. This can further support the choice of the former while mapping the current set of predictors to the countries' stability index.

# Analysing events occurring in 2021

## Subset selection according to `year`

```
year <- "2021"
pred <- predictors[predictors$Year==year,]
stab <- stability[stability$Year==year,]
df <- list(pred, stab) %>% reduce(full_join, by='ISO')
```

## Moran I

Again, a positive Moran Index (0.36) is found to be significant ($p$-value $= 0.0001$) and a spatial effect is occurring over the stability index estimated over this year.

```
moran.test(df$stability, nbW, randomisation=FALSE)
```

```
##
##  Moran I test under normality
##
## data:  df$stability
## weights: nbW
## n reduced by no-neighbour observations
##
## Moran I statistic standard deviate = 3.7036, p-value = 0.0001063
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation           Variance
##        0.36669774       -0.02222222         0.01102717
```

## Ordinary Least Square (OLS) estimation and spatial dependency

In the following code chunks the Elhorst procedure is tested again over the current subset of data. An OLS is tested and coefficients are shown below:

```
l <- lm(formula, data=df)
summary(l)
```

```
##
## Call:
## lm(formula = formula, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.05577 -0.57059  0.09742  0.49724  1.61596
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     8.940e-01  1.352e+00   0.661  0.51255
## EventRoot5      6.589e-05  4.312e-04   0.153  0.87939
## EventRoot6      1.925e-03  1.805e-03   1.067  0.29314
## EventRoot7     -1.104e-03  1.423e-03  -0.776  0.44295
## EventRoot10     5.630e-04  2.423e-03   0.232  0.81759
## EventRoot12    -3.857e-04  1.505e-03  -0.256  0.79923
## EventRoot13     3.113e-03  2.704e-03   1.151  0.25713
## EventRoot14     1.052e-03  1.079e-03   0.975  0.33590
## EventRoot19    -1.080e-03  3.172e-04  -3.404  0.00164 **
## AvgTone         6.492e-01  4.429e-01   1.466  0.15144
## GoldsteinScale -3.717e-01  8.734e-01  -0.426  0.67295
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7394 on 36 degrees of freedom
## Multiple R-squared:  0.3882, Adjusted R-squared:  0.2182
## F-statistic: 2.284 on 10 and 36 DF,  p-value: 0.03411
```

Again `EventRoot19` proved to be significant ($p\text{-value} = 0.001$) yielding a negative coefficient ($\beta = -1.08 \times 10^{-3}$). Following the result of the LM test which points out how RS test for both the default and adjusted statistics are significant ($p\text{-value} < 0.05$ in both conditions): consequently, the condition to test spatial regression models with global spillover effect are met.

```
LMtest <- spdep::lm.RStests(l, nbW, zero.policy=TRUE, test="all")
summary(LMtest)
```

```
##  Rao's score (a.k.a Lagrange multiplier) diagnostics for spatial
##  dependence
## data:
## model: lm(formula = formula, data = df)
## test weights: nbW
##
##           statistic parameter p.value
## RSerr        3.1881         1 0.07417 .
## RSlag        6.2888         1 0.01215 *
## adjRSerr     1.2465         1 0.26423
## adjRSlag     4.3471         1 0.03707 *
## SARMA        7.5353         2 0.02311 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Spatial autoregressive global spillover models

### Spatial Durbin Model (SDM)

```
sdm <- lagsarlm(stability ~ EventRoot5 + EventRoot6 + EventRoot7 +
                EventRoot10 + EventRoot12 + EventRoot13 +
                EventRoot14 + EventRoot19 + AvgTone +
                GoldsteinScale, data = df,
                listw=nbW, type="mixed",
                zero.policy=TRUE)
```

### Spatial Autoregressive Model (SAR)

```
sar <- lagsarlm(stability ~ EventRoot5 + EventRoot6 + EventRoot7 +
                EventRoot10 + EventRoot12 + EventRoot13 +
                EventRoot14 + EventRoot19 + AvgTone +
                GoldsteinScale, data = df,
                listw=nbW,
                zero.policy=TRUE)
```

### Testing The Likelihood Ratio test (LRT)

```
anova(sar, sdm)
```

```
##     Model df    AIC  logLik Test L.Ratio p-value
## sar     1 13 109.72 -41.863    1
## sdm     2 23 114.50 -34.251    2  15.224 0.12412
```

According to the LRT there is no sufficient evidence to further support the SDM given that the Null hypothesis has to be rejected and therefore $\theta$ is considered to be equal to zero.

By contrast, the SAR model summary points out that the LRT addressing the $H_0 : \rho = 0$ is rejected (*p-value < 0.01*) yielding therefore to accept the Alternative hypothesis $\rho \neq 0$ (shown below). This implies that the simpler SAR model can be considered as a solution that best represents the current stability effect in a global spillover manner. This evidence is further support by the AIC showing SAR as the model that achieved the smaller information criterion.

```
print("AIC:")
```

```
## [1] "AIC:"
```

```
print(paste("OLS:",round(AIC(l),3),
            " SDM:",round(AIC(sdm),3),
            " SAR:",round(AIC(sar),3),  sep=" "))
```

```
## [1] "OLS: 116.471  SDM: 114.501  SAR: 109.725"
```

In the following chunk SAR coefficients and previously discussed results from the LRT with rejected $H_0 : \rho = 0$ are shown.

```
summary(sar)
```

```
##
## Call:lagsarlm(formula = stability ~ EventRoot5 + EventRoot6 + EventRoot7 +
##     EventRoot10 + EventRoot12 + EventRoot13 + EventRoot14 + EventRoot19 +
##     AvgTone + GoldsteinScale, data = df, listw = nbW, zero.policy = TRUE)
##
## Residuals:
##      Min       1Q    Median       3Q      Max
## -1.09011 -0.52059  0.11047  0.45996  1.21995
##
## Type: lag
## Regions with no neighbours included:
##   28
## Coefficients: (asymptotic standard errors)
##                   Estimate  Std. Error z value  Pr(>|z|)
## (Intercept)     0.91566583  1.03666654  0.8833 0.3770855
## EventRoot5      0.00019338  0.00033067  0.5848 0.5586711
## EventRoot6      0.00020439  0.00138525  0.1475 0.8826987
## EventRoot7     -0.00046313  0.00109228 -0.4240 0.6715632
## EventRoot10    -0.00208784  0.00186061 -1.1221 0.2618085
## EventRoot12     0.00061279  0.00115489  0.5306 0.5956917
## EventRoot13     0.00317580  0.00207601  1.5298 0.1260757
## EventRoot14     0.00160226  0.00082850  1.9339 0.0531214
## EventRoot19    -0.00094664  0.00024334 -3.8902 0.0001002
## AvgTone         0.42088908  0.34080136  1.2350 0.2168312
## GoldsteinScale -0.49467885  0.66964451 -0.7387 0.4600779
##
## Rho: 0.54365, LR test value: 8.746, p-value: 0.0031029
## Asymptotic standard error: 0.13067
##     z-value: 4.1606, p-value: 3.1744e-05
## Wald statistic: 17.31, p-value: 3.1744e-05
##
## Log likelihood: -41.86257 for lag model
## ML residual variance (sigma squared): 0.32134, (sigma: 0.56687)
## Number of observations: 47
## Number of parameters estimated: 13
## AIC: 109.73, (AIC for lm: 116.47)
## LM test for residual autocorrelation
## test value: 0.033009, p-value: 0.85583
```

The `EventRoot19` predictor is still highlighting a negative effect ($\beta = -0.9 \times 10^{-3}$ with *p-value = 0.001*): this coefficient can be explained correctly according to the impacts evaluation as follows in the next section.

**Impacts**

```
impacts_sar <- spatialreg::impacts(sar, listw=nbW, R=1000)
s <- summary(impacts_sar, zstats=TRUE, short=TRUE)
im <- data.frame(s$res)
p <- data.frame(s$pzmat)
rownames(im) <- c("EventRoot5","EventRoot6","EventRoot7","EventRoot10",
                  "EventRoot12","EventRoot13","EventRoot14","EventRoot19",
                  "AvgTone","GoldsteinScale")
```

We are going to select those predictors whose probability associated to the simulation test is below 0.05:

```
rown <- rownames(p[p$Direct<0.05 | p$Indirect<0.05 | p$Total<0.05,])
```

the corresponding *impact effects* are shown below,

```
print("Displaying impact effects:")
```

```
## [1] "Displaying impact effects:"
```

```
im[rown, ]
```

```
##                  direct      indirect        total
## EventRoot19 -0.001032544 -0.001041819 -0.002074363
```

along with the corresponding *probability.*

```
print("Displaying impact probability:")
```

```
## [1] "Displaying impact probability:"
```

```
p[rown, ]
```

```
##               Direct   Indirect    Total
## EventRoot19 0.0002144206 0.1679178 0.03496806
```

From the results shown above a negative and significant direct impact (-0.001, *p-value = 0.0002*) implies that the higher the amount of events coded as belonging to the root code 19, the more likely its stability index is to be affected negatively in the same country's borders. No significance evidence has been found for indirect impact on the same predictor.

# Analysing events occurring in 2022

## Subset selection according to `year`

```
year <- "2022"
pred <- predictors[predictors$Year==year,]
stab <- stability[stability$Year==year,]
df <- list(pred, stab) %>% reduce(full_join, by='ISO')
```

## Moran I

As it was found for the past years, the subset of 2022 records shows a positive Moran Index (0.33) which indeed is significant (*p-value = 0.0002*).

```
moran.test(df$stability, nbW, randomisation=FALSE)
```

```
##
##  Moran I test under normality
```

```
##
## data:  df$stability
## weights: nbW
## n reduced by no-neighbour observations
##
## Moran I statistic standard deviate = 3.446, p-value = 0.0002845
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic        Expectation            Variance
##        0.33963919        -0.02222222          0.01102717
```

### Ordinary Least Square (OLS) estimation and spatial dependency

The current linear model fitted over the 2022 events highlighted a single significant predictor (`EventRoot6`, amount of event related to material cooperation) in a positive $\beta$ coefficient $(2 \times 10^{-3})$ (*p-value < 0.05*).

```
l <- lm(formula, data=df)
summary(l)
```

```
##
## Call:
## lm(formula = formula, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.40078 -0.46441 -0.01662  0.57880  1.85904
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -1.1675809  1.3097718  -0.891   0.3786
## EventRoot5      0.0007345  0.0007847   0.936   0.3555
## EventRoot6      0.0027781  0.0012915   2.151   0.0383 *
## EventRoot7     -0.0028906  0.0017804  -1.624   0.1132
## EventRoot10     0.0023302  0.0021695   1.074   0.2899
## EventRoot12    -0.0010956  0.0012059  -0.909   0.3696
## EventRoot13     0.0021079  0.0033689   0.626   0.5355
## EventRoot14    -0.0016007  0.0016298  -0.982   0.3326
## EventRoot19    -0.0004767  0.0004502  -1.059   0.2967
## AvgTone        -0.0640207  0.4202609  -0.152   0.8798
## GoldsteinScale  0.4505230  0.8065385   0.559   0.5799
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8434 on 36 degrees of freedom
## Multiple R-squared:  0.2235, Adjusted R-squared:  0.007771
## F-statistic: 1.036 on 10 and 36 DF,  p-value: 0.4343
```

Evidence of the the global spillover effect in the standard RS statistics (*p-value = 0.009*) is present along with a slightly significant probability associated to error dependencies (`RSerr`) allowing to reject the $H_0$ for both classes of models (global and local spillover). Given that the former is allows to reject the Null hypothesis with higher confidence, global spatial autoregressive models will be fitted in the next section.

```
LMtest <- spdep::lm.RStests(l, nbW, zero.policy=TRUE, test="all")
summary(LMtest)
```

```
##  Rao's score (a.k.a Lagrange multiplier) diagnostics for spatial
```

```
##   dependence
## data:
## model: lm(formula = formula, data = df)
## test weights: nbW
##
##           statistic parameter  p.value
## RSerr        3.9698         1 0.046322 *
## RSlag        6.7035         1 0.009622 **
## adjRSerr     1.0633         1 0.302472
## adjRSlag     3.7969         1 0.051346 .
## SARMA        7.7668         2 0.020581 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Spatial autoregressive global spillover models

### Spatial Durbin Model (SDM)

```
sdm <- lagsarlm(stability ~ EventRoot5 + EventRoot6 + EventRoot7 +
                EventRoot10 + EventRoot12 + EventRoot13 +
                EventRoot14 + EventRoot19 + AvgTone +
                GoldsteinScale, data = df,
                listw=nbW, type="mixed",
                zero.policy=TRUE)
```

### Spatial Autoregressive Model (SAR)

```
sar <- lagsarlm(stability ~ EventRoot5 + EventRoot6 + EventRoot7 +
                EventRoot10 + EventRoot12 + EventRoot13 +
                EventRoot14 + EventRoot19 + AvgTone +
                GoldsteinScale, data = df,
                listw=nbW,
                zero.policy=TRUE)
```

### Testing The Likelihood Ratio test (LRT)

Again from LRT it can be stated that SDM can be rejected in favor of SAR model considering that the Null hypothesis for which $H_0 : \theta = 0$ needs to be accepted (*p-value = 0.11*).

```
anova(sar, sdm)
```

```
##     Model df    AIC  logLik Test L.Ratio p-value
## sar     1 13 122.98 -48.491    1
## sdm     2 23 127.56 -40.781    2   15.42  0.1175
```

Moreover while comparing the AIC of each of the model here tested considering the 2022 subset we got that:

```
print("AIC:")
```

```
## [1] "AIC:"
```

```
print(paste("OLS:",round(AIC(l),3),
            " SDM:",round(AIC(sdm),3),
            " SAR:",round(AIC(sar),3),  sep=" "))
```

```
## [1] "OLS: 128.84  SDM: 127.562  SAR: 122.982"
```

Here SAR related coefficients are displayed and the LR test evaluating whether $\rho = 0$ under the Null hypothesis yield to accept the Alternative hypothesis considering the current probability (0.005). $\rho$ is significantly different from zero and thus SAR is correctly explaining the current spatial dependency.

```
summary(sar)
```

```
##
## Call:lagsarlm(formula = stability ~ EventRoot5 + EventRoot6 + EventRoot7 +
##     EventRoot10 + EventRoot12 + EventRoot13 + EventRoot14 + EventRoot19 +
##     AvgTone + GoldsteinScale, data = df, listw = nbW, zero.policy = TRUE)
##
## Residuals:
##       Min         1Q     Median         3Q        Max
## -1.525704 -0.478998   0.055004   0.525894   1.438762
##
## Type: lag
## Regions with no neighbours included:
##   28
## Coefficients: (asymptotic standard errors)
##                   Estimate  Std. Error z value Pr(>|z|)
## (Intercept)     -0.82832475  1.02812940 -0.8057  0.42044
## EventRoot5       0.00072221  0.00061364  1.1769  0.23923
## EventRoot6       0.00211736  0.00100935  2.0977  0.03593
## EventRoot7      -0.00245791  0.00139373 -1.7636  0.07781
## EventRoot10      0.00063150  0.00170044  0.3714  0.71036
## EventRoot12     -0.00067936  0.00094480 -0.7191  0.47211
## EventRoot13      0.00238606  0.00262796  0.9080  0.36390
## EventRoot14     -0.00096093  0.00127501 -0.7537  0.45105
## EventRoot19     -0.00040489  0.00035114 -1.1531  0.24888
## AvgTone         -0.12456062  0.32781476 -0.3800  0.70397
## GoldsteinScale   0.37683912  0.62941745  0.5987  0.54937
##
## Rho: 0.49299, LR test value: 7.8579, p-value: 0.0050598
## Asymptotic standard error: 0.14056
##     z-value: 3.5074, p-value: 0.00045253
## Wald statistic: 12.302, p-value: 0.00045253
##
## Log likelihood: -48.49093 for lag model
## ML residual variance (sigma squared): 0.4328, (sigma: 0.65787)
## Number of observations: 47
## Number of parameters estimated: 13
## AIC: 122.98, (AIC for lm: 128.84)
## LM test for residual autocorrelation
## test value: 0.14252, p-value: 0.70579
```

**Impacts**

```
impacts_sar <- spatialreg::impacts(sar, listw=nbW, R=1000)
s <- summary(impacts_sar, zstats=TRUE, short=TRUE)
im <- data.frame(s$res)
p <- data.frame(s$pzmat)
rownames(im) <- c("EventRoot5","EventRoot6","EventRoot7","EventRoot10",
                  "EventRoot12","EventRoot13","EventRoot14","EventRoot19",
                  "AvgTone","GoldsteinScale")
```

We are going to select those predictors whose probability associated to the simulation test is below 0.05:

```r
rown <- rownames(p[p$Direct<0.05 | p$Indirect<0.05 | p$Total<0.05,])
```

the corresponding *impact effects* are shown below,

```r
print("Displaying impact effects:")
```

```
## [1] "Displaying impact effects:"
```

```r
im[rown, ]
```

```
##                direct     indirect       total
## EventRoot6 0.002267364 0.001908774 0.004176138
```

along with the corresponding *probability*.

```r
print("Displaying impact probability:")
```

```
## [1] "Displaying impact probability:"
```

```r
p[rown, ]
```

```
##                Direct   Indirect      Total
## EventRoot6 0.03901375 0.3039651 0.1291187
```

A positive and (slightly) significant direct impact (0.002, *p-value = 0.04*) on the amount of events that represent any form of material cooperation (`EventRoot6`) is displayed as above showing that the higher the occurrence of goods trading promotion, the higher the stability index of that country on average.