Relevant readings for the lecture are Waller & Gotway, Chapter 9, and Golpher , A. B., and P. R. Voss. How to Interpret the Coefficients of Spatial Models: Spillovers, Direct and Indirect Effects. Spatial Demography (2016) 4:175-205.

Classification of Gaussian spatial regression models

- First order effects in all spatial regression models are as usually written as linear regression equation $E(\mathbf{y}\,|\,\mathbf{X}) = \mathbf{X}\cdot\boldsymbol{\beta}$ in a set of independent variables $\mathbf{X}$.
- Second order effects are reflected by a dependency pattern among the regression disturbances $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\cdot\boldsymbol{\beta}$.
- Note: this discussion is based on the unobserved disturbances $\boldsymbol{\varepsilon}$ and the population regression parameters $\boldsymbol{\beta}$.
- The covariance matrix of the disturbances $Cov(\boldsymbol{\varepsilon}) = \sigma^2 \cdot \boxed{\boldsymbol{\Omega}(\rho)}$ is exhibiting some form of spatial autocorrelation where σ^2 is the variance of the disturbances (i.e., we assume homoscedasticity) and $\boldsymbol{\Omega}(\rho)$ autocorrelation matrix, which depends on the autocorrelation coefficient ρ.
- Depending on the explicit specification of autocorrelation matrix $\boldsymbol{\Omega}(\rho)$ we get three different Gaussian spatial processes:
    - The ***simultaneous autoregressive*** spatial process (SAR) $\boldsymbol{\Omega}(\rho) = (\mathbf{I} - \rho\cdot\mathbf{V})^{-1}\cdot(\mathbf{I} - \rho\cdot\mathbf{V}^T)^{-1}$ or equivalently $\boldsymbol{\Omega}(\rho) = \left[(\mathbf{I} - \rho\cdot\mathbf{V}^T)\cdot(\mathbf{I} - \rho\cdot\mathbf{V})\right]^{-1}$
    - The ***moving average*** spatial process (MA) $\boldsymbol{\Omega}(\rho) = (\mathbf{I} + \rho\cdot\mathbf{V})\cdot(\mathbf{I} + \rho\cdot\mathbf{V}^T)$.

- The ***conditional autoregressive*** spatial process (CAR)
  $\mathbf{\Omega}(\rho) = (\mathbf{I} - \rho \cdot \mathbf{V})^{-1}$.
  Here the matrix $\mathbf{I} - \rho \cdot \mathbf{V}$ must be ***positive definite*** and, therefore, the coded spatial link matrix **V** has to be ***symmetric***.

- The ***feasible value range*** of the spatial autocorrelation coefficient ρ is restricted in all specification.
  In essence, its range depends on the eigenvalues of the coded spatial link matrix **V**.
  Any autocorrelation coefficient ρ outside the feasible range leads to a ***non-stationary*** spatial process.
- For the moving average and the simultaneous autoregressive spatial processes, the matrix $\mathbf{\Omega}(\rho)$ automatically satisfies symmetry and positive definiteness (the inverse matrix exists) criteria for any feasible autocorrelation coefficients.
- Most popular in spatial autocorrelation modeling is the *simultaneous autoregressive* process with a *first order spatial lag* matrix in the *W*-coding scheme. – This is mostly for practical reasons.
- Consequently, in the subsequent discussion we will focus only on $\mathbf{\Omega}(\rho) = (\mathbf{I} - \rho \cdot \mathbf{W})^{-1} \cdot (\mathbf{I} - \rho \cdot \mathbf{W}^T)^{-1}$

The simultaneous autoregressive process

- The model structure of a simultaneous autoregressive process is
  $\mathbf{y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon}$ where $\boldsymbol{\varepsilon} = (\mathbf{I} - \rho \cdot \mathbf{W})^{-1} \cdot \boldsymbol{\eta} \quad \Leftrightarrow \quad \boldsymbol{\varepsilon} = \rho \cdot \mathbf{W} \cdot \boldsymbol{\varepsilon} + \boldsymbol{\eta}$.
  - $\boldsymbol{\eta}$ is the ***independent white noise*** disturbance, which is the random input into the stochastic process
  - $(\mathbf{I} - \rho \cdot \mathbf{W})^{-1}$ is the linking mechanism of the stochastic process
  - $\boldsymbol{\varepsilon}$ is the spatially autocorrelated output of the stochastic process.

- This model can also be written in different but equivalent forms:

$$\mathbf{y} = \mathbf{X}\cdot\boldsymbol{\beta} + \boxed{(\mathbf{I}-\rho\cdot\mathbf{W})^{-1}\cdot\boldsymbol{\eta}}$$ error term

$$\Leftrightarrow \quad (\mathbf{I}-\rho\cdot\mathbf{W})\cdot\mathbf{y} = (\mathbf{I}-\rho\cdot\mathbf{W})\cdot\mathbf{X}\cdot\boldsymbol{\beta} + \boldsymbol{\eta}$$ (this is the specification of the FGLS estimator)

$$\Leftrightarrow \quad \mathbf{y} = \rho\cdot\mathbf{W}\cdot\mathbf{y} + \mathbf{X}\cdot\boldsymbol{\beta} - \rho\cdot\mathbf{W}\cdot\mathbf{X}\cdot\boldsymbol{\beta} + \boldsymbol{\eta}$$

- The terms $\mathbf{W}\cdot\mathbf{y}$ and $\mathbf{W}\cdot\mathbf{X}$ denote the spatial averages (spill-over effects) of the dependent and independent variables, respectively, around the reference locations.

  not random

- In OLS the right side of the regression equation comprises solely of exogenous information and a random error term.
  Now, however, we also have the random component $\rho\cdot\mathbf{W}\cdot\mathbf{y}$ on the right hand side, which therefore will be correlated with the error term $\boldsymbol{\eta}$.
  This violates the fundamental OLS assumptions that the exogenous information is independent from the random error term.

- Note: As long as the underlying data generating process is Gaussian auto-regressive the estimated OLS regression coefficients $\boxed{\hat{\boldsymbol{\beta}} \text{ remain unbiased.}}$
  However, if the underlying data generation process deviates for the Gaussian auto-regressive model the estimated OLS regression coefficients will become biased.

- The estimated coefficients $\hat{\boldsymbol{\beta}}_*$ of the term $\mathbf{W}\cdot\mathbf{X}$ have to satisfy the ***common factor constraint*** $\hat{\boldsymbol{\beta}}_* = -\hat{\rho}\cdot\hat{\boldsymbol{\beta}}$ for the ***simultaneous autoregressive model***.
  The violation of this constraint leads to the Durbin model in spatial econometrics.

- Commonly used alternative models are:
  - Estimates may indicate, however, a violation of the common factor constraint, that is, $\widehat{\boldsymbol{\beta}}_* \neq -\rho \cdot \widehat{\boldsymbol{\beta}}$. Then the model no longer a simultaneous autoregressive.
  - Alternatively, one may constrain $\boldsymbol{\beta}_* = \mathbf{0}$ explicitly to be zero. This leads to the econometric spatial ***lag model*** $\mathbf{y} = \rho \cdot \mathbf{W} \cdot \mathbf{y} + \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\eta}$.
- Both models can be distinguished be the Lagrange multiplier tests.
  - The Lagrange multiplier test for the ***simultaneous autoregressive*** error model is asymptotically χ^2-distributed with one degree of freedom. It is asymptotically equivalent to Moran's *I*, however, Moran's *I* is more precise.
  - The Lagrange multiplier test for the *lag* model is asymptotically χ^2-distributed with one degree of freedom.
  - If the Lagrange multiplier test for the lag model is more significant than the Lagrange multiplier test for the spatial autoregressive model ($\alpha_{LAG} < \alpha_{SAR}$ or $\chi^2_{SAR} < \chi^2_{LAG}$, respectively) then the underlying data generating process has more in common with the lag model and we choose the lag model specification.
  - Alternatively, $\boxed{\text{if } \alpha_{LAG} > \alpha_{SAR} \text{ or } \chi^2_{SAR} > \chi^2_{LAG}}$, respectively, then we choose the simultaneous autoregressive model as underlying data generating process.
  - If neither χ^2_{SAR} nor χ^2_{LAG} is significant then the data generating process is neither LAG nor SAR.
- Evaluate the different models with the R-script **`SpatialRegModels.R`**.

Excurse: Calculation of the Spatial Covariance Matrix

- Bailey and Gatrell (p 284) derive the covariance matrix of the simultaneous autoregressive process explicitly. Recall that $\boldsymbol{\varepsilon} = (\mathbf{I} - \rho \cdot \mathbf{W})^{-1} \cdot \boldsymbol{\eta}$. Thus:

$$
\begin{aligned}
Cov(\boldsymbol{\varepsilon}) &= E(\boldsymbol{\varepsilon} \cdot \boldsymbol{\varepsilon}^T) \text{ because the mean of } E(\boldsymbol{\varepsilon}) = \mathbf{0} \\
&= E\left[(\mathbf{I} - \rho \cdot \mathbf{W})^{-1} \cdot \boldsymbol{\eta} \cdot \boldsymbol{\eta}^T \cdot (\mathbf{I} - \rho \cdot \mathbf{W}^T)^{-1}\right] \\
&= (\mathbf{I} - \rho \cdot \mathbf{W})^{-1} \cdot \underbrace{E(\boldsymbol{\eta} \cdot \boldsymbol{\eta}^T)}_{=\sigma^2 \cdot \mathbf{I}} \cdot (\mathbf{I} - \rho \cdot \mathbf{W}^T)^{-1} \\
&= \sigma^2 \cdot \left[(\mathbf{I} - \rho \cdot \mathbf{W})^{-1} \cdot (\mathbf{I} - \rho \cdot \mathbf{W}^T)^{-1}\right] \\
&= \sigma^2 \cdot \left[(\mathbf{I} - \rho \cdot \mathbf{W}^T) \cdot (\mathbf{I} - \rho \cdot \mathbf{W})\right]^{-1} \text{ because } \mathbf{B}^{-1} \cdot \mathbf{A}^{-1} = \left[\mathbf{A} \cdot \mathbf{B}\right]^{-1} \\
&= \sigma^2 \cdot \boldsymbol{\Omega}(\rho)
\end{aligned}
$$

- If heteroscedasticity is present, that is, $E(\boldsymbol{\eta} \cdot \boldsymbol{\eta}^T) \neq \sigma^2 \cdot \mathbf{I}$, a diagonal weights matrix can be entered in-between the terms $(\mathbf{I} - \rho \cdot \mathbf{W})^{-1}$.
- Unfortunately, by default the matrix $\boldsymbol{\Omega}(\rho) = (\mathbf{I} - \rho \cdot \mathbf{W})^{-1} \cdot (\mathbf{I} - \rho \cdot \mathbf{W}^T)^{-1}$ is not constant on the diagonal. Thus by design the spatial process has a varying variance and, therefore, is not perfectly stationary. Also its off-diagonal elements may differ for a given spatial lag.
- The variance stabilizing *S*-coding scheme of a spatial link matrix alleviates this problem (See Tiefelsdorf, Griffith, Boots, 1999, Environment and Planning A)

Excurse: Estimation of the regression model under spatial autocorrelation

- If we would know the full specification of the spatial correlation matrix $\mathbf{\Omega}(\rho)$ in advance, then we could use General Least Squares to estimate the parameters $\boldsymbol{\beta}$ by
  $\hat{\boldsymbol{\beta}}_{gls} = (\mathbf{X}^T \cdot \mathbf{\Omega}^{-1}(\rho) \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot \mathbf{\Omega}^{-1}(\rho) \cdot \mathbf{y}$
- In the case of a simultaneous autoregressive process this equation simplifies to
  $\hat{\boldsymbol{\beta}}_{gls} = [\mathbf{X}^T \cdot (\mathbf{I} - \rho \cdot \mathbf{W}^T) \cdot (\mathbf{I} - \rho \cdot \mathbf{W}) \cdot \mathbf{X}]^{-1} \cdot \mathbf{X}^T \cdot (\mathbf{I} - \rho \cdot \mathbf{W}^T) \cdot (\mathbf{I} - \rho \cdot \mathbf{W}) \cdot \mathbf{y}$
  because the inverse of $\mathbf{\Omega}^{-1}(\rho) = (\mathbf{I} - \rho \cdot \mathbf{W}^T) \cdot (\mathbf{I} - \rho \cdot \mathbf{W})$.
- Thus no calculation of the inverse spatial autocorrelation matrix is required which is a main advantage of the simultaneous autoregressive model.
- However, since the autocorrelation level ρ is unknown, one would need to use feasible general least squares and estimate the spatial autocorrelation coefficient ρ by maximum likelihood.
- Assuming that $\mathbf{y} \sim N(\mathbf{X} \cdot \boldsymbol{\beta}, \sigma^2 \cdot \mathbf{\Omega}(\rho))$, that is, $\mathbf{y}$ is normal distributed with an expectation of $\mathbf{X} \cdot \boldsymbol{\beta}$ and a co-variance of $\sigma^2 \cdot \mathbf{\Omega}(\rho)$, the log-likelihood function is
  $\ell(\boldsymbol{\beta}, \sigma^2, \rho \mid \mathbf{y}, \mathbf{X}) \propto -\ln\left|\sigma^2 \cdot \mathbf{\Omega}(\rho)\right| - (\mathbf{y} - \mathbf{X} \cdot \boldsymbol{\beta})^T \cdot \frac{1}{\sigma^2} \cdot \mathbf{\Omega}^{-1}(\rho) \cdot (\mathbf{y} - \mathbf{X} \cdot \boldsymbol{\beta})$
- The term within the determinant of $\ln\left|\sigma^2 \cdot \mathbf{\Omega}(\rho)\right|$ is called the log-Jacobian. It ensured that we are dealing with a proper density function that integrates to one. This term complicates the estimation of any autoregressive spatial model substantially.

Misconceptions in Bailey and Gatrell:

- p 284: $-1 \leq \rho \leq 1$ does not hold, the bounds depend on the specific spatial structure, coding scheme and spatial process.
- p 285: For positive spatial autocorrelation the standard deviation of the estimated regression parameters is inflated. This statement is wrong; the direction of the bias cannot be given *a priori*.

Further reading on model specifications and interpretations:

Anselin, Luc (2003). Spatial Externalities, Spatial Multipliers, and Spatial Econometrics. *International Regional Science Review* **26**:153-166

Golgher, Andre Braz & Paul R. Voss (2016) How to Interpret the Coefficients of Spatial Models: Spillovers, Direct and Indirect Effects. *Spatial Demography*, **4**:175-205

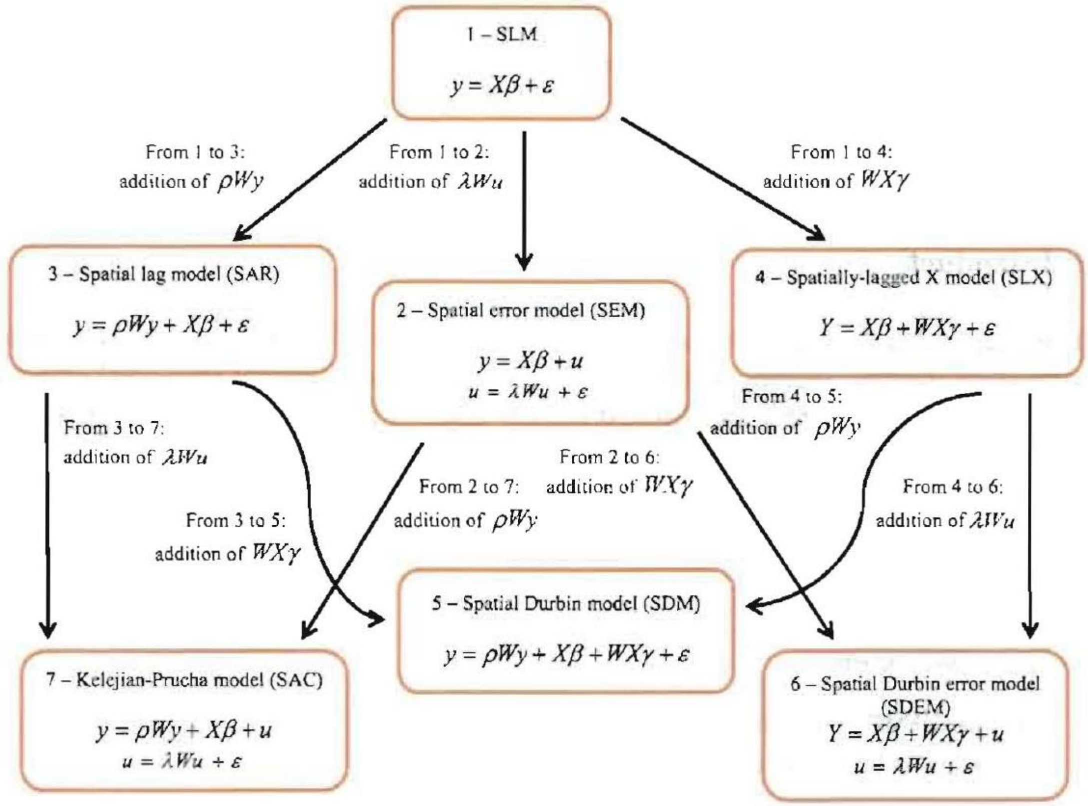

Fig. 1 Some spatial models