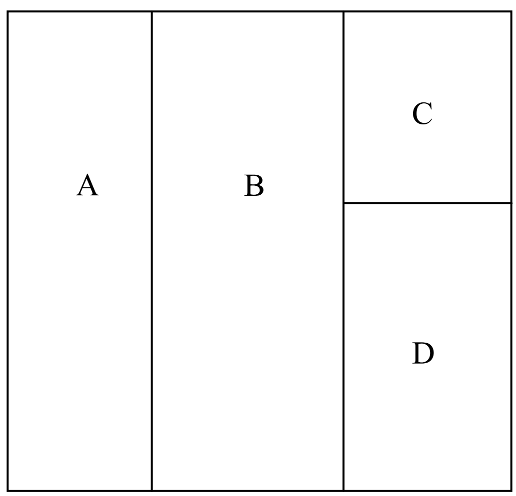
SPATIAL CORRELATION

[1] Motivation

- Potential correlations in error terms are generally ignored in the analysis of cross-section data.
- But error terms could be correlated even among cross-section units.
  - Crime rate of a county could be influenced by the crime rates of the adjacent counties.



  - The demand for rice of a village would be influenced by the demands of adjacent villages (Case, 1991, Econometrica).
  - The expenditures of a local government are likely to be correlated with the expenditures of the governments in a near area.

[2] Spatial Correlations in Error Terms

- Consider a following regression model:

$$y = X\beta + \varepsilon = 1_T\beta_1 + X^*\beta^* + \varepsilon ,$$

where, $E(X'\varepsilon) = 0$,

$y = (y_1,...,y_T)'$; $\varepsilon = (\varepsilon_1,...,\varepsilon_T)'$;

$X = (x_1,...,x_T)'$, x_t is k×1;

1_T = T×1 vector of ones.

(1) Equicorrelations in all errors

$$Cov(\varepsilon) = \sigma^2\Omega \equiv \sigma^2 \begin{pmatrix} 1 & \rho & \rho & ... & \rho \\ \rho & 1 & \rho & ... & \rho \\ \rho & \rho & 1 & ... & \rho \\ : & : & : & & : \\ \rho & \rho & \rho & ... & 1 \end{pmatrix}; \ 0 < \rho < 1.$$

$$\begin{aligned} \Omega &= (1-\rho)I_T + \rho e_T e_T' = (1-\rho)(P_T + M_T) + T\rho P_T \\ &= (1-\rho+T\rho)P_T + (1-\rho)M_T , \end{aligned}$$

where,

$e_T = (1,...,1)'$, $P_T = I_T - e_T(e_T'e_T)^{-1}e_T'$ and $M_T = I_T - P_T$.

$$\rightarrow \ \Omega^{-1} = \frac{1}{1-\rho+T\rho}P_T + \frac{1}{1-\rho}M_T .$$

- Consequences:
  - $\hat{\beta}^*_{OLS} = \hat{\beta}^*_{GLS}$ and $Cov(\hat{\beta}^*_{OLS}) = \sigma^2(1-\rho)(X^{*\prime}M(1_T)X^*)^{-1}$

    E[SSE/(T-k)] = $\sigma^2(1-\rho)$, where SSE is from OLS.

    → Can be shown by using Frisch-Waugh Theorem.
  - Not detectable by standard tests for autocorrelations (such as Durbin-Watson).
  - Need not to worry about this type of correlations. The OLS procedure ignoring this type of correlation would not bias statistic inferences.

(2) Equicorrelations within a group

- Assume two stochastically independent groups: {1,...,T/2} and {T/2+1,...,T}.
- Two models:

  (A) $y = 1_T\beta_1 + X^*\beta^* + \varepsilon$; (B) $y = Z_1\beta_1 + Z_2\beta_2 + X^{**}\beta^{**} + \varepsilon$,

  where $Z_1 = \begin{pmatrix} 1_{(T/2)} \\ 0_{(T/2)} \end{pmatrix}$ and $Z_2 = \begin{pmatrix} 0_{(T/2)} \\ 1_{(T/2)} \end{pmatrix}$ are the vectors of two group dummy variables.

- $$Cov\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_{T/2}\end{pmatrix} = \sigma^2\begin{pmatrix}1 & \rho_1 & \cdots & \rho_1\\ \rho_1 & 1 & \cdots & \rho_1\\ \vdots & \vdots & & \vdots\\ \rho_1 & \rho_1 & \cdots & 1\end{pmatrix};$$
  $$Cov\begin{pmatrix}\varepsilon_{T/2+1}\\ \varepsilon_{T/2+2}\\ \vdots\\ \varepsilon_T\end{pmatrix} = \sigma^2\begin{pmatrix}1 & \rho_2 & \cdots & \rho_2\\ \rho_2 & 1 & \cdots & \rho_2\\ \vdots & \vdots & & \vdots\\ \rho_2 & \rho_2 & \cdots & 1\end{pmatrix}$$

  $\rightarrow$

  $$Cov\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_{T/2}\\ \varepsilon_{T/2+1}\\ \varepsilon_{T/2+2}\\ \vdots\\ \varepsilon_T\end{pmatrix} = \sigma^2\begin{pmatrix}1 & \rho_1 & .. & \rho_1 & 0 & 0 & \cdots & 0\\ \rho_1 & 1 & \cdots & \rho_1 & 0 & 0 & \cdots & 0\\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots\\ \rho_1 & \rho_1 & \cdots & 1 & 0 & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0 & 1 & \rho_2 & \cdots & \rho_2\\ 0 & 0 & \cdots & 0 & \rho_2 & 1 & \cdots & \rho_2\\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 0 & \rho_2 & \rho_2 & \cdots & 1\end{pmatrix}.$$
  $$= \sigma^2\left[\begin{pmatrix}(1-\rho_1)I_{T/2} & 0_{(T/2)\times(T/2)}\\ 0_{(T/2)\times(T/2)} & (1-\rho_2)I_{T/2}\end{pmatrix} + \rho_1 Z_1 Z_1' + \rho_2 Z_2 Z_2'\right]$$

- OLS on Model (A) [Moulton (1990)]:
  - $\hat{\beta}^*_{OLS} \neq \hat{\beta}^*_{GLS}$.
  - Estimation of $Cov(\hat{\beta}^*_{OLS})$ becomes complicated.

- OLS on Model (B):
  - If $\rho_1 = \rho_2 = \rho$,

    $\hat{\beta}^{**}_{OLS} = \hat{\beta}^{**}_{GLS}$;

    $Cov(\hat{\beta}^{*}_{OLS}) = \sigma^2(1-\rho)(X^{**\prime}M(Z)X^{**})^{-1}$;

    E[SSE/(T-k)] = $\sigma^2(1-\rho)$,

    where Z = [Z_1,Z_2] and SSE is from OLS.
  - If $\rho_1 \neq \rho_2$,

    $\hat{\beta}^{**}_{OLS} \neq \hat{\beta}^{**}_{GLS}$;

    $Cov(\hat{\beta}_{OLS})$ should be estimated by the White method.
  - When model (B) is used, OLS with the White-corrected covariance matrix can be used.

- Lessons:
  - Equicorrelations within groups can be controlled by using group dummy variables as regressors.

(3) Spatial Error Dependence (Anselin and Bera, 1998)

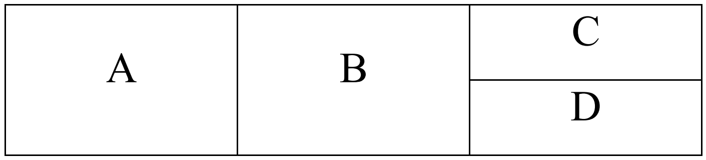
- Weight Matrix (T×T)
  - Rule 1: Zeros on diagonals.
  - Rule 2: Insert 1 on adjacent observations.
  - Rule 3: Normalize each row such that sum = 1.

[Example 1] 3 observations

A	B	C

$$W = \begin{pmatrix} 0 & & \\ & 0 & \\ & & 0 \end{pmatrix} \rightarrow W = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \rightarrow W = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}.$$

[Example 2] 4 observations



$$W = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \rightarrow W = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 1/3 & 1/3 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 & 0 \end{pmatrix}$$

$\rightarrow$ W is not necessarily symmetric.

- Model:

$$y_{T\times 1} = X_{T\times k}\beta_{k\times 1} + \varepsilon_{T\times 1}; \quad \varepsilon = \rho W\varepsilon + v,$$

where $v = (v_1,...,v_T)'$ and the v_t are iid with $(0,\sigma^2)$.

- Cov(ε):
  - $(I - \rho W)\varepsilon = v \rightarrow \varepsilon = (I - \rho W)^{-1}v$.
  - $Cov(\varepsilon) = \sigma^2(I - \rho W)^{-1}(I - \rho W')^{-1}$.
  - If ρ is known, β can be estimated by GLS.

- Estimation of ρ (Kelejian and Prucha)
  - Let:
    - $\bar{\varepsilon} = W\varepsilon$; $\ddot{\varepsilon} = WW\varepsilon$.
  - Define:
    - $\Gamma_T = \begin{pmatrix} \frac{2}{T}\varepsilon'\bar{\varepsilon} & -\frac{1}{T}\bar{\varepsilon}'\bar{\varepsilon} & 1 \\ \frac{2}{T}\ddot{\varepsilon}'\bar{\varepsilon} & -\frac{1}{T}\ddot{\varepsilon}'\ddot{\varepsilon} & \frac{1}{T}tr(W'W) \\ \frac{1}{T}(\varepsilon'\ddot{\varepsilon} + \bar{\varepsilon}'\bar{\varepsilon}) & -\frac{1}{T}\bar{\varepsilon}'\ddot{\varepsilon} & 0 \end{pmatrix}$; $\gamma_T = \begin{pmatrix} \frac{1}{T}\varepsilon'\varepsilon \\ \frac{1}{T}\bar{\varepsilon}'\bar{\varepsilon} \\ \frac{1}{T}\varepsilon'\bar{\varepsilon} \end{pmatrix}$;

    $g_T(\rho, \sigma^2) = \Gamma_N \begin{pmatrix} \rho \\ \rho^2 \\ \sigma^2 \end{pmatrix} - \gamma_N$.

    - Then, it can be shown:

$$E[g_T(\rho, \sigma^2)] = 0 \qquad (*)$$

    - Do GMM using the moment conditions (*) and replacing ε by OLS residuals.

- Difficulty in GLS
    - Often, (I-ρW) is too big to compute its inverse.

(4) General Spatial Correlations in Errors.

- Conley (JEC, 1999).
- Can be used for OLS and GMM.

[3] Spatial Lag Dependence (Spatial Autoregressive) Model

- Model:

$$y = \rho W y + X\beta + v,$$

where $v = (v_1, ..., v_T)'$ and the v_t are iid.

- Example: 3 observations

A	B	C

$$W = \begin{pmatrix} 0 & w_{12} & 0 \\ w_{21} & 0 & w_{23} \\ 0 & w_{32} & 0 \end{pmatrix} \rightarrow W = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}.$$

$$y_1 = \rho w_{12} y_2 + x_1'\beta + \varepsilon_1;$$

$$\rightarrow \quad y_2 = \rho w_{21} y_1 + \rho w_{23} y_3 + x_2'\beta + \varepsilon_2;$$

$$y_3 = \rho w_{32} y_2 + x_3'\beta + \varepsilon_2.$$

$$\rightarrow \quad y_2 = \frac{(\rho w_{21} x_1 + x_2 + \rho w_{23} x_3)'\beta}{1 - \rho^2 w_{21} w_{12} - \rho^2 w_{23} w_{32}} + error.$$

- Estimation:

1) OLS: Inconsistent.

2) MLE

  - Assume that v is $N(0_{T\times 1},\sigma^2 I_T)$.
  - $y \sim N((I-\rho W)^{-1}X\beta,\Omega)$, where $\Omega = \sigma^2(I-\rho W)^{-1}(I-\rho W')^{-1}$.
  - $$f(y|X) = \frac{1}{(2\pi)^{T/2}|\Omega|^{1/2}} \times [y-(I-\rho W)^{-1}X\beta]'\Omega^{-1}[y-(I-\rho W)^{-1}X\beta] = \frac{1}{(2\pi)^{T/2}(\sigma^2)^{T/2}|(I-\rho W)^{-1}|} \times \frac{1}{2\sigma^2}(y-\rho Wy - X\beta)'(y-\rho Wy - X\beta).$$
  - $|I-\rho W| = \prod_{t=1}^{T}(1-\rho w_t)$,

    where the w_t are the eigenvalues of W.
  - $$\ell_T(\beta,\rho,\sigma^2) = \Sigma_{t=1}^{T}\ln(1-\rho w_t) - \frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln(\sigma^2) - \frac{(y-\rho Wy-X\beta)'(y-\rho Wy-X\beta)}{2\sigma^2}.$$

3) GMM

  - Using Z = [WX,X] as instruments.