Lecture 3: Tests of Hypotheses for Spatial Linear Regression Model

Zhenlin Yang

School of Economics, Singapore Management University
zlyang@smu.edu.sg
http://www.mysmu.edu.sg/faculty/zlyang/

ECON747: Spatial Econometric Models and Methods
Term I, 2024-25

3.1. Introduction

This lecture introduces tests of hypothesis in spatial linear regression (SLR) models including:

1. tests for covariate effect (CE) or SD effect in all SLR models;
2. tests for SE effect in SLR-SE model;
3. tests for SL effect in SLR-SL model;
4. tests for SE effect and/or SL effect in SLR-SLE model:
    - joint test for the existence/non-existence of both SL and SE effects,
    - marginal test for SE effect allowing the existence of SL effect,
    - marginal test for SL effect allowing the existence of SE effect.

The test in (2) is also called the conditional test of SE given no SL.
The test in (3) is also called the conditional test of SL given no SE.

Three fundamental principles for hypothesis tests

There are three fundamental principles one can follow when constructing tests of hypothesis concerning parameters θ in a model:

- Likelihood ratio principle,

  (Jerzy Neyman and Egon S. Pearson, 1928, 1933; Samuel S. Wilks, 1938)
- Wald principle, (Abraham Wald, 1943)
- Score or Lagrange multiplier principle,

  (C. R. Rao, 1948; J. Aitchison and S. D. Silvey, 1958)

These three tests are referred to in statistical literature on testing of hypotheses as the Holy Trinity or the three classical tests.

- The three principles are equivalent to the first order of asymptotics – their limiting null distribution is chi-square with d.f. being the number of restrictions imposed by the null hypothesis,
- but differ to some extent in the second order properties.

Linear hypotheses/restrictions:

$$H_0: \; R\theta_0 = r, \tag{3.1}$$

where r is a $q \times 1$ nonrandom vector and R is a $q \times p$ nonstochastic matrix, $q \leq p$, with $\text{rank}(R) = q$.

- With proper choices of R and r, various hypotheses corresponding to the SLR-SLE models can be formulated, e.g., $\beta_{10} = \beta_{20}$, $\beta_{30} = \beta_{40} = 0, \;\; \lambda_0 = \rho_0 = 0, \;\; \lambda_0 = 0$ (single parameter hypothesis).
- The test of $\beta_3 = \beta_{40} = 0$ may correspond to spatial Durbin effects.

Nonlinear hypotheses/restrictions:

$$H_0: \; c(\theta_0) = 0, \tag{3.2}$$

where c: $\Theta \to \mathbb{R}^q$ is a continuously differentiable function on the parameter space $\Theta \subset \mathbb{R}^q$ and θ_0 is assumed to lie in the interior of Θ. Nonlinear hypotheses are of less interest in the SLR context.

We follow the general notations of Lecture 2: θ, $\ell_n(\theta)$, $S_n(\theta)$, $\mathcal{I}_n(\theta)$, and $\mathcal{J}_n(\theta)$, are, respectively, $p \times 1$ vector of parameters, (quasi) loglikelihood, (quasi) score, VC matrix of (quasi) score, and expected negative Hessian.

Recall from Lecture 2: under the general quasi ML framework,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N\big[0,\ \lim_{n\to\infty} n\mathcal{J}_n^{-1}(\theta_0)\mathcal{I}_n(\theta_0)\mathcal{J}_n^{-1}(\theta_0)\big], \tag{3.3}$$

which can equivalently be written as

$$\hat{\theta}_n \overset{a}{\sim} N\big[\theta_0,\ \mathcal{J}_n^{-1}(\theta_0)\mathcal{I}_n(\theta_0)\mathcal{J}_n^{-1}(\theta_0)\big],$$

where $\overset{a}{\sim}$ denotes “distributed asymptotically as”.

Note: the establishment of (3.3) requires the following result:

$$\tfrac{1}{\sqrt{n}} S_n(\theta_0) \xrightarrow{D} N\big[0, \lim_{n\to\infty} \tfrac{1}{n}\mathcal{I}_n(\theta_0)\big]. \tag{3.4}$$

Under correct specification, $\mathcal{I}_n(\theta_0) = \mathcal{J}_n(\theta_0)$, $\Rightarrow$ $\hat{\theta}_n \overset{a}{\sim} N\big[\theta_0,\ \mathcal{I}_n^{-1}(\theta_0)\big]$.
Now, $\mathcal{I}_n(\theta_0) =$ **Fisher Information (FI)** = variance of efficient score.

When error distribution is correctly specified, i.e., when $\ell_n(\theta)$ is the genuine loglikelihood function, the LR test statistic is defined as

$$\mathrm{LR}_n = -2[\ell_n(\tilde{\theta}_n) - \ell_n(\hat{\theta}_n)], \tag{3.5}$$

where $\tilde{\theta}_n$ is the restricted MLE of θ under H_0 and $\hat{\theta}_n$ is the MLE of θ under the full model. Under H_0, $\mathrm{LR}_n \overset{a}{\sim} \chi^2_q$.

- When error distribution is misspecified, $\ell_n(\theta)$ becomes a quasi loglikelihood and the result (3.5) no longer holds in general, except,
- when $\mathcal{I}_n = \alpha\mathcal{J}_n$ (generalized IME), then we have a quasi LR (QLR) test: $\mathrm{QLR}_n = -2[\ell_n(\tilde{\theta}_n) - \ell_n(\hat{\theta}_n)]/\hat{\alpha}_n$.
- For general QML estimation, with $\mathcal{I}_n \neq \alpha\mathcal{J}_n$, the LR test statistic can be distributed as a weighted sum of chi-squares (see Cameron & Trivedi, 2005, Sec. 8.5.3).

The Wald test is the most versatile one among the three as it only requires that $\hat{\theta}_n \xrightarrow{a} N(\theta_0, \mathcal{V}_n)$ where $n\mathcal{V}_n$ is positive definite for large enough n, where $\hat{\theta}_n$ can be any estimator with an asymptotic normal distribution.

- **Test of individual coefficient.** Let $\widehat{\mathcal{V}}_n$ be a consistent estimator of $\mathcal{V}_n$ and $\hat{\nu}_{njj}$ be the jth diagonal element of $\widehat{\mathcal{V}}_n$. Then, the asymptotic standard error of $\hat{\theta}_{nj}$ (the $j$the component of $\hat{\theta}_n$) is $\mathrm{se}(\hat{\theta}_{nj}) = \sqrt{\hat{\nu}_{njj}}$. Wald test statistic is simply $\mathrm{T}_n = (\hat{\theta}_{nj} - \theta_0)/\sqrt{\hat{\nu}_{njj}}$.
- **Test of linear hypothesis**, H_0: $R\theta_0 = r$. Wald statistic takes the form:

$$\mathrm{T}_n = (R\hat{\theta}_n - r)'(R\widehat{\mathcal{V}}_n R')^{-1}(R\hat{\theta}_n - r) \overset{a}{\sim} \chi^2_q, \text{ under } H_0. \tag{3.6}$$

- **Test of nonlinear hypothesis**, H_0: $c(\theta) = 0$. Wald statistic is:

$$\mathrm{T}_n = c'(\hat{\theta}_n)\left[C(\hat{\theta}_n)\widehat{\mathcal{V}}_n C'(\hat{\theta}_n)\right]^{-1} c(\hat{\theta}_n) \overset{a}{\sim} \chi^2_q, \text{ under } H_0, \tag{3.7}$$

where $C(\theta) = \nabla_\theta c(\theta)$ is the $q \times p$ Jacobian of the $c(\theta)$ function.

When error distribution is correctly specified, i.e., when $S_n(\theta)$ is the genuine score function, the LM or Score test statistic is defined simply as

$$\mathrm{LM}_n^{\mathtt{FI}} = S_n(\tilde{\theta}_n)\mathcal{I}_n^{-1}(\tilde{\theta}_n)S_n(\tilde{\theta}_n) \overset{a}{\sim} \chi_q^2, \text{ under } H_0, \tag{3.8}$$

where FI denotes Fisher Information, and $\tilde{\theta}_n$ is the MLE of θ under H_0.

- LM test is most preferred as it requires estimating only the null model. This is especially so when the null model is an OLS regression.
- $\mathcal{I}_n(\tilde{\theta}_n)$ can be replaced by $-\frac{\partial}{\partial\theta'}S_n(\theta)\big|_{\theta=\tilde{\theta}_n}$, to give an observed information ($\mathtt{OI}$) variant of LM test, denoted by $\mathrm{LM}_n^{\mathtt{OI}}$.
- It can also replaced by $\sum_{i=1}^n g_{ni}(\tilde{\theta}_n)g_{ni}'(\tilde{\theta}_n)$, **if** $S_n(\theta_0) = \sum_{i=1}^n g_{ni}(\theta_0)$ where $\{g_{ni}(\theta_0)\}$ form a martingale difference (M.D.) sequence, to give an M.D. variant of LM test, denoted by $\mathrm{LM}_n^{\mathrm{MD}}$.
- These tests are asymptotically equivalent, valid for any type of H_0, but may not be robust against distributional misspecification (DM).

When error distribution is misspecified, LM tests may not be valid as IME may fail to hold. However, it can be modified to allow for DM.

Robust LM test for model reduction. Let $\theta = (\vartheta', \varphi')'$ and consider

$$H_0 : \varphi_0 = 0.$$

Partition $S_n(\theta) = [S'_{n\vartheta}(\vartheta, \varphi), S'_{n\varphi}(\vartheta, \varphi)]'$, corresponding to (ϑ, φ), $\mathcal{I}_n = [\mathcal{I}_{n,\vartheta\vartheta}, \mathcal{I}_{n,\vartheta\varphi}; \mathcal{I}_{n,\varphi\vartheta}, \mathcal{I}_{n,\varphi\varphi}]$, and $\mathcal{J}_n = [\mathcal{J}_{n,\vartheta\vartheta}, \mathcal{J}_{n,\vartheta\varphi}; \mathcal{J}_{n,\varphi\vartheta}, \mathcal{J}_{n,\varphi\varphi}]$.

As $S_{n\vartheta}(\tilde{\vartheta}_n, 0) = 0$ at the null estimate $\tilde{\vartheta}_n$ of ϑ_0, the construction of robust LM test depends on $S_{n\varphi}(\tilde{\vartheta}_n, 0)$. A Taylor expansion leads to

$$\frac{1}{\sqrt{n}} S_{n\varphi}(\tilde{\vartheta}_n, 0) = \frac{1}{\sqrt{n}} S_{n\varphi}(\vartheta_0, 0) - \frac{1}{\sqrt{n}} \Pi_n(\vartheta_0) S_{n\vartheta}(\vartheta_0, 0) + o_p(1), \qquad (3.9)$$

where $\Pi_n(\vartheta_0) = \mathcal{J}_{n,\varphi\vartheta}(\vartheta_0, 0)\mathcal{J}^{-1}_{n,\vartheta\vartheta}(\vartheta_0, 0)$. It follows that

$$\mathrm{Var}\big[\tfrac{1}{\sqrt{n}} S_{n,\varphi}(\tilde{\theta}_n, 0)\big] = \tfrac{1}{n}\big[\mathcal{I}_{n,\varphi\varphi}(\theta_0, 0) - \mathcal{I}_{n,\varphi\vartheta}(\vartheta_0, 0)\Pi'_n(\vartheta_0) - \Pi_n(\vartheta_0)\mathcal{I}_{n,\vartheta\varphi}(\vartheta_0, 0) + \Pi_n(\vartheta_0)\mathcal{I}_{n,\vartheta\vartheta}(\vartheta_0, 0)\Pi'_n(\vartheta_0)\big] + o(1).$$

A general LM test for model reduction, robust to DM, is given as:

$$\mathrm{LM}_n^* = \tilde{S}'_{n,\varphi}\big(\tilde{\mathcal{I}}_{n,\varphi\varphi} - 2\tilde{\mathcal{I}}_{n,\varphi\vartheta}\tilde{\Pi}'_n + \tilde{\Pi}_n\tilde{\mathcal{I}}_{n,\vartheta\vartheta}\tilde{\Pi}'_n\big)^{-1}\tilde{S}_{n,\varphi}, \tag{3.10}$$

where $\tilde{S}_{n,\varphi} = S_{n,\varphi}(\tilde{\vartheta}_n, 0)$ and similarly are the other tilde-quantities defined. The limiting null distribution of LM_n^* is $\chi^2_{\mathrm{dim}(\varphi)}$. **Note:**

- When error distribution is correctly specified, $\mathcal{I}_n(\theta_0) = \mathcal{J}_n(\theta_0)$, and

$$\mathcal{I}_{n,\varphi\varphi} - 2\mathcal{I}_{n,\varphi\theta}\tilde{\Pi}'_n + \Pi_n\mathcal{I}_{n,\theta\theta}\Pi'_n = \mathcal{I}_{n,\varphi\varphi} - \mathcal{I}_{n,\varphi\vartheta}\mathcal{I}_{n,\vartheta\vartheta}\mathcal{I}'_{n,\varphi\vartheta}.$$

- On the other hand, for testing $\varphi = 0$, $\mathrm{LM}_n^{\mathtt{FI}}$ given in (3.8) reduces to

$$\mathrm{LM}_n^{\mathtt{FI}} = \tilde{S}'_{n,\varphi}\big(\tilde{\mathcal{I}}_n^{-1}\big)_{\varphi\varphi}\tilde{S}_{n,\varphi}, \tag{3.11}$$

  where $(\cdot)_{\varphi\varphi}$ denotes the $\varphi\varphi$-block of the corresponding matrix. Using the inverse of a partitioned matrix, we have

$$\big(\tilde{\mathcal{I}}_n^{-1}\big)_{\varphi\varphi} = (\mathcal{I}_{n,\varphi\varphi} - \mathcal{I}_{n,\varphi\vartheta}\mathcal{I}_{n,\vartheta\vartheta}\mathcal{I}'_{n,\varphi\vartheta})^{-1}. \tag{3.12}$$

- LM_n^* reduces to $\mathrm{LM}_n^{\mathtt{FI}}$ when error distribution is correctly specified.

If $S_n(\theta_0) = \sum_{i=1}^n g_{ni}(\theta_0)$, an M.D. representation, then by (3.9) $S_{n\varphi}(\tilde{\vartheta}_n, 0)$ has an asymptotic M.D. representation:

$$\frac{1}{\sqrt{n}} S_{n,\varphi}(\tilde{\vartheta}_n, 0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[g_{ni,\varphi}(\vartheta_0, 0) - \Pi_n(\vartheta_0) g_{ni,\vartheta}(\vartheta_0, 0) \right] + o_p(1), \qquad (3.13)$$

where $(g'_{ni,\vartheta}, g'_{ni,\varphi})' = g_{ni}$, and $\{g_{ni,\varphi} - \Pi_n g_{ni,\vartheta}\}$ form an M.D. sequence.

An OPMD (outer-product-of-martingale-difference) variant of LM_n^* is:

$$\mathrm{LM}_n^\dagger = \left(\sum_{i=1}^n \tilde{g}_{ni,\varphi}^\dagger\right)' \left(\sum_{i=1}^n \tilde{g}_{ni,\varphi}^\dagger \tilde{g}_{ni,\varphi}^{\dagger\,\prime}\right)^{-1} \left(\sum_{i=1}^n \tilde{g}_{ni,\varphi}^\dagger\right), \qquad (3.14)$$

where $\tilde{g}_{ni,\varphi}^\dagger = \tilde{g}_{ni,\varphi} - \tilde{\Pi}_n \tilde{g}_{ni,\vartheta}$.

- $\mathrm{LM}_n^\dagger$ can be obtained by directly working on LM_n^*,
- $\mathrm{LM}_n^\dagger \stackrel{a}{=} \mathrm{LM}_n^*$, where $\stackrel{a}{=}$ denotes asymptotic equivalence.
- For many spatial models, an M.D. representation for quasi score can be made, and an OPMD variant of robust LM test can be derived.
- MD and OPMD are natural methods leading to LM tests robust against both DM (distributional misspecification) and UH (unknown heteroskedasticity).

An important point: although score test starts from the likelihood setup, the score principle is applicable to any problem where the estimators solve a first-order conditions, including the general class of M-estimators.

LM test for a general hypothesis, linear or nonlinear, can also be extended to allow for DM. We give a general form of robust LM test for nonlinear hypothesis, H_0: $c(\theta_0) = 0$, as linear hypothesis is a special case.

- First, if $\tilde{\theta}_n$ solves $S_n(\theta) = 0$ s.t. $c(\theta_0) = 0$, it must be that $c(\tilde{\theta}_n) = 0$.
- Apply the mean value theorem to jth element $c_j(\tilde{\theta}_n)$ of $c(\tilde{\theta}_n)$,

$$0 = \sqrt{n}c_j(\tilde{\theta}_n) = \sqrt{n}c_j(\theta_0) + \sqrt{n}C_j(\bar{\theta}_n)(\tilde{\theta}_n - \theta_0),$$

where $\bar{\theta}_n$ lies between $\tilde{\theta}_n$ and θ_0, $C_j(\theta)$ = jth row of $C(\theta) = \frac{\partial c(\theta)}{\partial \theta'}$.

- $\Rightarrow \sqrt{n}C_j(\bar{\theta}_n)(\tilde{\theta}_n - \theta_0) = 0$, and as $C_j(\bar{\theta}_n) - C_j(\theta_0) = o_P(1)$, $\forall j$,

$$\sqrt{n}C(\theta_0)(\tilde{\theta}_n - \theta_0) = o_p(1).$$

By a Taylor expansion, we have under H_0,

$$\frac{1}{\sqrt{n}}S_n(\tilde{\theta}_n) = \frac{1}{\sqrt{n}}S_n(\theta_0) + \frac{1}{n}\mathcal{J}_n\sqrt{n}(\tilde{\theta}_n - \theta_0)] + o_p(1). \quad (3.15)$$

Premultiply $C(n\mathcal{J}_n^{-1})$ through (3.15), we have under H_0,

$$\begin{aligned} C(n\mathcal{J}_n^{-1})\frac{1}{\sqrt{n}}S_n(\tilde{\theta}_n) &= C(n\mathcal{J}_n^{-1})[\frac{1}{\sqrt{n}}S_n(\theta_0) + \frac{1}{n}\mathcal{J}_n\sqrt{n}(\tilde{\theta}_n - \theta_0)] + o_p(1) \\ &= C(n\mathcal{J}_n^{-1})[\frac{1}{\sqrt{n}}S_n(\theta_0)] + o_p(1). \end{aligned}$$

$$\Rightarrow \quad Var[C(n\mathcal{J}_n^{-1})\frac{1}{\sqrt{n}}S_n(\tilde{\theta}_n)] = nC\mathcal{J}_n^{-1}\mathcal{I}_n\mathcal{J}_n^{-1}C' + o(1).$$

A generalized LM test for testing H_0: $c(\theta_0) = 0$, robust against DM, is

$$\mathrm{LMG}_n = \tilde{S}_n'\tilde{\mathcal{J}}_n^{-1}\tilde{C}'\left(\tilde{C}\tilde{\mathcal{J}}_n^{-1}\tilde{\mathcal{I}}_n\tilde{\mathcal{J}}_n^{-1}\tilde{C}'\right)^{-1}\tilde{C}\tilde{\mathcal{J}}_n^{-1}\tilde{S}_n, \quad (3.16)$$

where $\tilde{S}_n = S_n(\tilde{\theta}_n)$, and similarly are the other tilde-quantities defined. The limiting null distribution of LMG_n is χ_q^2.

In the Gaussian QML framework, the QS functions that LM-type tests depend upon are linear-quadratic (LQ) forms in model errors. This allows an M.D. decomposition of the QS functions and thereby the construction of an OPMD variant of DM-robust LM test, as shown by the following lemma.

Lemma 3.1. Let A_n be an $n \times n$ non-stochastic matrix with elements $a_{n,ij}$, and b_n an $n \times 1$ non-stochastic vector with elements b_{ni}. Let ϵ_n be an $n \times 1$ random vector of iid elements, $\{\epsilon_{ni}\}$, with mean 0, variance σ_0^2, skewness γ_0 and finite excess kurtosis κ_0. Define $Q_n(\epsilon_n) = \epsilon_n' A_n \epsilon_n + b_n' \epsilon_n$. We have,

(i) $Q_n(\epsilon_n) - \mathrm{E}[Q_n(\epsilon_n)] = \sum_{i=1}^n [\epsilon_{ni}\xi_{ni} + a_{n,ii}(\epsilon_{ni}^2 - \sigma^2) + b_{ni}\epsilon_{ni}] \equiv \sum_{i=1}^n g_{ni}$, where $\xi_n = ((A_n^u)' + A_n^l)\epsilon_n$ with elements $\{\xi_{ni}\}$, and $A_n = A_n^u + A_n^l + A_n^d$, sum of upper triangular, lower triangular and diagonal matrices of A_n.

(ii) $\{g_{ni}, \mathcal{F}_{n,i}\}$ form a martingale difference sequence w.r.t. $\mathcal{F}_{n,i}$, the increasing σ-field generated by $\{\epsilon_{n1}, \ldots, \epsilon_{ni}\}$, i.e. $\mathrm{E}(g_{ni}|\mathcal{F}_{n,i-1}) = 0$.

(iii) $\{g_{ni}\}$ are uncorrelated and $\mathrm{Var}[Q_n(\epsilon_n)] = \sum_{i=1}^n \mathrm{E}(g_{ni}g_{ni}')$.

To understand Lemma 3.1 (i), letting $a_n = \texttt{diagv}(A_n)$, we have

$$
\begin{aligned}
Q_n(\epsilon_n) - \mathrm{E}[Q_n(\epsilon_n)] &= \epsilon_n'(A_n^u + A_n^l + A_n^d)\epsilon_n + b_n'\epsilon_n - \sigma^2 1_n' a_n \\
&= \epsilon_n'(A_n^u + A_n^l)\epsilon_n + a_n'\epsilon_n^2 + b_n'\epsilon_n - \sigma^2 1_n' a_n, \\
&= \epsilon_n'((A_n^u)' + A_n^l)\epsilon_n + a_n'(\epsilon_n^2 - \sigma^2 1_n) + b_n'\epsilon_n \\
&\equiv \epsilon_n'\xi_n + a_n'(\epsilon_n^2 - \sigma^2 1_n) + b_n'\epsilon_n \\
&= \textstyle\sum_{i=1}^n (\epsilon_{ni}\xi_{ni} + a_{n,ii}(\epsilon_{ni}^2 - \sigma^2) + b_{ni}\epsilon_{ni}) \equiv \sum_{i=1}^n g_{ni}.
\end{aligned}
$$

The results of Lemma 3.1 can be extended to a k-vector LQ forms:

$$
\mathbf{Q}_n(\epsilon_n) = \begin{cases} \epsilon_n' A_{1n}\epsilon_n + b_{1n}'\epsilon_n \\ \vdots \\ \epsilon_n' A_{kn}\epsilon_n + b_{kn}'\epsilon_n \end{cases} = \sum_{i=1}^n \mathbf{g}_{ni},
$$

where $\mathbf{g}_{ni} = (g_{1,ni}, \ldots, g_{k,ni})'$. This gives an M.D. decomposition of $\mathbf{Q}_n(\epsilon_n)$, as $\{\mathbf{g}_{ni}, \mathcal{F}_{n,i}\}$ form a vector M.D. sequence. Thus, $\mathrm{Var}[Q_n(\epsilon_n)] = \sum_{i=1}^n \mathrm{E}(\mathbf{g}_{ni}\mathbf{g}_{ni}')$. See Baltagi and Yang (2013b) for details.

3.2. Tests of Hypotheses for Spatial Error Model

Recall: linear regression model with SE dependence given in (2.1):

$$Y_n = X_n\beta + u_n, \quad u_n = \rho W_n u_n + \epsilon_n, \tag{3.17}$$

where $X_n\beta$ may contain spatial Durbin (SD) term, and SAR process for SE dependence can be replaced by SMA process: $u_n = \rho W_n \epsilon_n + \epsilon_n$,

- Y_n: $n \times 1$ vector of observations on n spatial units,
- X_n: an $n \times k$ matrix containing the values of k regressors,
- W_n: $n \times n$ matrix summarizing interactions among n spatial units, called the spatial weight matrix or the spatial interaction matrix,
- ϵ_n: $n \times 1$ vector of independent and identically distributed (iid) idiosyncratic errors with mean zero and variance σ^2,
- ρ: the spatial error parameter,
- β: $k \times 1$ vector of regression coefficients.

Hypotheses on Model (3.17) that are of main interest concern (i) covariate effect (CE) and (ii) spatial error (SE) effect:

- $H_0^{\mathrm{CE}}: R\beta_0 = r$, some regressors can be merged or dropped,
- $H_0^{\mathrm{SE}}: \rho_0 = 0$, standard liner regression model suffices,

where R is a $q \times k$ constant matrix, and $q \leq k$.

When the R matrix is designed so the each row contains a sole non-zero value of one corresponding to the Durbin effects, then a test of H_0^{CE} gives a test of no Durbin effects.

When the rows of R sum to zero, then a test of H_0^{CE} is a test of linear contrasts in β, e.g., $H_0^{\mathrm{CE}}: \beta_{10} = \beta_{20}$ and $\beta_{30} = \beta_{40}$.

Recall: the loglikelihood function $\ell_n(\theta)$ given in (2.4) and rewritten as:

$$\ell_n(\theta) = -\frac{n}{2}\ln(2\pi\sigma^2) + \log|B_n(\rho)| - \frac{1}{2\sigma^2}\|\epsilon_n(\beta,\rho)\|^2. \quad (3.18)$$

where $\epsilon_n(\beta,\rho) = B_n(\rho)(Y_n - X_n\beta)$, $B_n(\rho) = I_n - \rho W_n$, and $\|\cdot\|$ denotes Euclidean norm. The LR test statistic takes the general form (Sec. 3.1):

$$\mathrm{LR}_n = -2[\ell_n(\tilde{\theta}_n) - \ell_n(\hat{\theta}_n)],$$

where $\tilde{\theta}_n$ and $\hat{\theta}_n$ are the restricted and unrestricted MLEs of θ.

For testing either H_0^{CE} or H_0^{SE} in the SE model, it can be shown that the LR test statistic takes a simple common form:

$$\mathrm{LR}_{\varpi} = n\ln(\tilde{\sigma}_n^2\,\hat{\sigma}_n^{-2}) - 2\log|B_n(\tilde{\rho}_n)B_n^{-1}(\hat{\rho}_n)|, \quad \text{for } \varpi = \mathrm{CE}, \mathrm{SE}, \quad (3.19)$$

due to the fact that $\hat{\sigma}_n^{-2}\|\epsilon_n(\hat{\beta}_n,\hat{\rho}_n)\|^2 = \tilde{\sigma}_n^{-2}\|\epsilon_n(\tilde{\beta}_n,\tilde{\rho}_n)\|^2 = \frac{n}{2}$.

- Under H_0^{SE}, $\mathrm{LR}_{\mathrm{CE}} \overset{a}{\sim} \chi_q^2$; under H_0^{SE}, $\tilde{\rho}_n = 0$ and $\mathrm{LR}_{\mathrm{SE}} \overset{a}{\sim} \chi_1^2$.
- Note: $\mathrm{LR}_{\mathrm{SE}} = -2[\ell_n^c(0) - \ell_n^c(\hat{\lambda}_n)]$, where $\ell_n^c(\lambda)$ is given in (2.8).

Recall: from (2.10) and (2.11), the expected negative Hessian matrix:

$$\mathcal{J}_{\mathrm{SE}}(\theta_0) = \begin{pmatrix} \frac{1}{\sigma_0^2} X_n' B_n' B_n X_n & 0 & 0 \\ \sim & \frac{n}{2\sigma_0^4} & \frac{1}{\sigma_0^2}\mathrm{tr}(G_n) \\ \sim & \sim & \mathrm{tr}(G_n^s G_n) \end{pmatrix}; \tag{3.20}$$

and the VC matrix of score, written as $\mathcal{I}_{\mathrm{SE}}(\theta_0) = \mathcal{J}_{\mathrm{SE}}(\theta_0) + \mathcal{K}_{\mathrm{SE}}(\theta_0)$, where,

$$\mathcal{K}_{\mathrm{SE}}(\theta_0) = \begin{pmatrix} 0 & \frac{1}{2\sigma_0^3}\gamma_0 X_n' B_n' \iota_n & \frac{1}{\sigma_0}\gamma_0 X_n' B_n' g_n \\ \sim & \frac{n}{4\sigma_0^4}\kappa_0 & \frac{1}{2\sigma_0^2}\kappa_0 \mathrm{tr}(G_n) \\ \sim & \sim & \kappa_0 g_n' g_n \end{pmatrix}, \tag{3.21}$$

ι_n is a vector of ones, γ_0 and κ_0 are the measures of skewness and excess kurtosis of $\epsilon_{n,i}$, $g_n = \mathrm{diagv}(G_n)$, and $G_n^s = G_n(\rho) + G_n'(\rho)$.

With the QMLE $\hat{\theta}_n$, and the plug-in estimators $\hat{\mathcal{I}}_{\mathrm{SE}} = \mathcal{I}_{\mathrm{SE}}(\hat{\theta}_n)$ and $\hat{\mathcal{J}}_{\mathrm{SE}} = \mathcal{J}_{\mathrm{SE}}(\hat{\theta}_n)$ given in Sec. 2.2.1, various Wald tests can be constructed.

Wald test for covariates effects. From Theorem 2.1, we see that $\mathcal{J}_{\mathrm{SE}}$ is block diagonal, and hence $\hat{\beta}_n \overset{a}{\sim} N(\beta_0, \sigma_0^2(X_n' B_n' B_n X_n)^{-1})$.

- Wald test for testing H_0^{CE}: $R\beta_0 = r$ has the expression:

$$\mathrm{T}_{\mathrm{CE}} = (R\hat{\beta}_n - r)'[\hat{\sigma}_n^2 R'(X_n'\hat{B}_n'\hat{B}_n X_n)^{-1} R]^{-1}(R\hat{\beta}_n - r). \tag{3.22}$$

  Asymptotic null distribution of the statistic T_{CE} is χ_q^2.

- When R is a row vector ($q = 1$), Wald test reduces to a t-test:

$$\mathrm{t}_{\mathrm{CE}} = \frac{R\hat{\beta}_n - r}{\sqrt{\hat{\sigma}_n^2 R(X_n'\hat{B}_n'\hat{B}_n X_n)^{-1} R'}} \overset{a}{\sim} N(0,1), \text{ under } H_0^{\mathrm{CE}}. \tag{3.23}$$

Wald test for spatial effect. Of particular interest is the test of H_0^{SE}: $\rho = 0$. A t-ratio, for confidence interval (CI) and test, takes the general form:

$$\mathrm{t}_{\mathrm{SE}} = \frac{\hat{\rho}_n - \rho_0}{\sqrt{(\hat{\mathcal{J}}_{\mathrm{SE}}^{-1}\hat{\mathcal{I}}_{\mathrm{SE}}\hat{\mathcal{J}}_{\mathrm{SE}}^{-1})_{\rho\rho}}} \overset{a}{\sim} N(0,1), \text{ under } H_0^{\mathrm{SE}}, \tag{3.24}$$

where $(\cdot)_{\rho\rho}$ denotes the ρ-ρ element of the corresponding matrix.

The univariate Wald statistic, or *t*-statistic, t_{CE}, is asymptotically $N(0, 1)$, and hence inferences for $R\beta_0$ is carried out by referring to the $N(0, 1)$ critical values.

- By construction, this test is robust against nonnormality of the errors.
- However, its finite sample performance may be poor.
- Liu and Yang (2015b) presented improved tests based on bias-correction.

Similarly, the *t*-statistic, t_{SE}, is asymptotically $N(0, 1)$, and hence inferences for ρ_0 is carried out by referring to the $N(0, 1)$ critical values.

- Again, this test is robust against nonnormality.
- The denominator of t_{SE} canbe made more 'explicit'.
- However, its finite sample property is not clear.
- It would be interesting to make a Monte Carlo comparison between t_{SE} given in (3.24) $\mathrm{LR}_{\mathrm{SE}}$ given in (3.19).

Recall: the (quasi) score function $S_n(\theta)$ given in (2.5):

$$S_{\text{SE}}(\theta) = \begin{cases} \frac{1}{\sigma^2} X_n' B_n'(\rho) B_n(\rho) u_n(\beta), \\ \frac{1}{2\sigma^4} u_n'(\beta) B_n'(\rho) B_n(\rho) u_n(\beta) - \frac{n}{2\sigma^2}, \\ \frac{1}{\sigma^2} u_n'(\beta) B_n'(\rho) W_n u_n(\beta) - \text{tr}[G_n(\rho)]. \end{cases} \tag{3.25}$$

where $G_n(\rho) = W_n B_n^{-1}(\rho)$.

Under normality: noting $\mathcal{I}_{\text{SE}} = \mathcal{J}_{\text{SE}}$, following (3.8) and discussions below it, we obtain three variants of LM tests for a general hypothesis on θ:

$$\text{LM}_{\text{FI}} = S_{\text{SE}}'(\tilde\theta_n) \mathcal{I}_{\text{SE}}(\tilde\theta_n)^{-1} S_{\text{SE}}(\tilde\theta_n), \tag{3.26}$$

$$\text{LM}_{\text{OI}} = S_{\text{SE}}'(\tilde\theta_n) \Big[-\frac{\partial}{\partial\theta'} S_{\text{SE}}(\theta)\big|_{\theta=\tilde\theta_n} \Big]^{-1} S_{\text{SE}}(\tilde\theta_n), \tag{3.27}$$

$$\text{LM}_{\text{MD}} = S_{\text{SE}}'(\tilde\theta_n) \Big[\textstyle\sum_{i=1}^n \tilde g_{ni} \tilde g_{ni}' \Big]^{-1} S_{\text{SE}}(\tilde\theta_n), \tag{3.28}$$

if in (3.28), $S_{\text{SE}}(\theta_0) = \sum_{i=1}^n g_{ni}(\theta_0)$ and $\{g_{ni}(\theta_0)\}$ is an M.D. sequence. All three statistics are χ_q^2 distributed under H_0 with q restrictions.

Among the three classical tests, the LM is of particular interest as its implementation requires only the estimation of the null model.

- Its advantage is clear if the null hypothesis specifies that $\rho = 0$, and hence the implementation of the LM test requires only OLS estimates.
- Many tests of this type are available in the literature.

We thus focus on the LM tests for $\texttt{SE}$ effect in linear regression model.

Moran's I test: To see if there exists spatial correlation among the observations, $Y_1, Y_2, \cdots, Y_n$, Moran (1950) propose a test of the form:

$$\mathtt{I} = \frac{\sum_i \sum_j w_{ij}(Y_i - \bar{Y})(Y_j - \bar{Y})}{\sum_i (Y_i - \bar{Y})^2}, \tag{3.29}$$

where $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_{ni}$. If the $Y'_{ni}s$ are iid normal when there is no spatial correlation, then the limiting null distribution of Moran's I statistic is normal.

Cliff-Ord test: Cliff and Ord (1972) extend Moran I to linear regression: $Y_n = X_n\beta + u_n$, and give a test of no spatial correlation among u_{ni} as:

$$\mathrm{I}_{\mathtt{SE}} = \frac{\tilde{u}_n' W_n \tilde{u}_n}{\tilde{u}_n' \tilde{u}_n}, \tag{3.30}$$

where $\tilde{u}_n$ is a vector of OLS residuals from regressing Y_n on X_n.
If $u_{ni}'s$ are normal, then the distribution of $\mathtt{I}_{\mathtt{SE}}$ under the null hypothesis of no spatial error dependence is asymptotically $N(\mu_{\mathtt{I}}, \sigma_{\mathtt{I}}^2)$, where

$$\mu_{\mathrm{I}} = \frac{1}{n-k}\mathrm{tr}(M_n W_n),$$
$$\sigma_{\mathtt{I}}^2 = \frac{\mathrm{tr}(M_n W_n M_n W_n') + \mathrm{tr}((M_n W_n)^2) - \frac{2}{n-k}[\mathrm{tr}(M_n W_n)]^2}{(n-k)(n-k+2)}.$$

Here $M_n = I_n - X_n(X_n'X_n)^{-1}X_n'$. This leads to a standardized Moran's I test for non-existence of spatial error correlation as:

$$\mathrm{I}_{\mathrm{SE}}^{\circ} = \frac{\mathtt{I}_{\mathtt{SE}} - \mu_{\mathtt{I}}}{\sigma_{\mathtt{I}}} \overset{a}{\sim} N(0,1), \quad \text{under the null.} \tag{3.31}$$

Clearly, (3.29) is a special case of (3.30) with $X_n = \iota_n$. Therefore, The final working version of Moran's I test is taken to be $\mathrm{I}_{\mathrm{SE}}^{\circ}$.

FI-based LM test: Burridge (1980) presented an LM test for testing H_0^{SE}: $\rho = 0$, based on Fisher information (FI) matrix $\mathcal{I}_n$ given in (3.20):

$$\text{LM}_{\text{SE}}^{\text{FI}} = \frac{n}{\sqrt{S_0}} \frac{\tilde{u}_n' W_n \tilde{u}_n}{\tilde{u}_n' \tilde{u}_n}, \tag{3.32}$$

where $S_0 = \text{tr}(W_n' W_n + W_n^2)$, having asymptotic null distribution $N(0, 1)$.

OPMD-based LM test: Born and Breitung (2011) derived an OPMD variant of Burridge's LM test of H_0^{SE}: $\rho = 0$:

$$\text{LM}_{\text{SE}}^{\text{MD}} = \frac{\tilde{u}_n' W_n \tilde{u}_n}{\sqrt{(\tilde{u}_n \odot \tilde{u}_n)'(\tilde{\xi}_n \odot \tilde{\xi}_n)}}, \tag{3.33}$$

where $\odot$ denotes Hadamard product, $\tilde{\xi}_n = (W_n^l + W_n^{u\prime})\tilde{u}_n$, W_n^l and W_n^u are the lower and upper triangular matrices of W_n, and $\text{LM}_{\text{SE}}^{\text{MD}}|_{H_0} \xrightarrow{D} N(0, 1)$.

To derive (3.33), note: $\tilde{u}_n' W_n \tilde{u}_n \stackrel{a}{=} u_n' W_n u_n$, $u_n' W_n u_n = u_n' \xi_n = \sum_{i=1}^n u_{ni} \xi_{ni}$, and $\{u_{ni} \xi_{ni}\}$ are uncorrelated if $\{u_{ni}\}$ are independent under H_0: $\rho = 0$.

Baltagi and Yang (2013a) commented both $\mathrm{LM}_{\mathrm{SE}}^{\mathtt{FI}}$ and $\mathrm{LM}_{\mathrm{SE}}^{\mathtt{MD}}$ may have poor finite sample performance as they are neither centered, nor standardized.

Baltagi and Yang (2013a) give a standardized version of $\mathrm{LM}_{\mathrm{SE}}^{\mathtt{FI}}$:

$$\mathrm{SLM}_{\mathrm{SE}}^{\circ} = \frac{n\tilde{u}_n'(W_n - S_1 I_N)\tilde{u}_n}{(\tilde{\kappa}_n S_2 + S_3)^{\frac{1}{2}}\,\tilde{u}_n'\tilde{u}_n}, \tag{3.34}$$

where $S_1 = \frac{1}{n-k}\mathrm{tr}(W_n M_n)$, $S_2 = \sum_{i=1}^n a_{ii}^2$, and $S_3 = \mathrm{tr}(A_n A_n' + A_n^2)$, $A_n = M_n W_n M_n - S_1 M_n$, a_{ii} are the diagonal elements of A_n, and $\tilde{\kappa}_n$ is the excess sample kurtosis of OLS residuals $\tilde{u}_n$.

To see (3.34), under H_0^{SE}, $\tilde{u}_n = M_n u_n$ and $\mathrm{E}(\tilde{u}_n' W_n \tilde{u}_n) = \sigma_0^2 \mathrm{tr}(W_n M_n) \neq 0$. This motivates the use of $\tilde{u}_n' W_n \tilde{u}_n - \sigma_0^2 \mathrm{tr}(W_n M_n)$ or its feasible version:

$$\tilde{u}_n' W_n \tilde{u}_n - \frac{1}{n-k}\tilde{u}_n'\tilde{u}_n \mathrm{tr}(W_n M_n) \equiv u_n' A_n u_n.$$

Finding $\mathrm{Var}(u_n' A_n u_n)$ and replacing σ_0^2 by $\tilde{\sigma}_n^2$ leads to the result.

Baltagi and Yang (2013a) also give a standardized version of $\mathrm{LM}_{\mathrm{SE}}^{\mathrm{MD}}$:

$$\mathrm{SLM}_{\mathrm{SE}}^{\mathrm{MD}} = \frac{\tilde{u}_n'(W_n - S_1 I_N)\tilde{u}_n}{\sqrt{(\tilde{u}_n \odot \tilde{u}_n)'[\tilde{\zeta}_n \odot \tilde{\zeta}_n + (A_n^d \tilde{u}_n) \odot (A_n^d \tilde{u}_n)]}}, \tag{3.35}$$

where $\tilde{\zeta}_n = (A_n^l + A_n^{u\prime})\tilde{u}_n$, and A_n^l, A_n^u and A_n^d are, respectively the lower, upper and diagonal matrices of A_n defined in (3.34).

The result follows almost immediately Lemma 3.1, by noticing that the numerator of (3.35) has the form $u_n' A_n u_n$, where $u_n = \epsilon_n$ under the null.

- A very important feature of the SLM tests is that their derivations do not depend on normality of errors, and thus robust against non-normality (NN).
- Another important feature of SLM-OPMD variant is that its variance estimate is also robust to unknown heteroskedasticity (UH).
- The ideas of **standardization** and **M.D. decomposition** can be extended to give SLM tests for other models (for robustness and better performance).
- See Baltagi and Yang (2013a,b) for details on these important ideas.

Assumption A1: *The errors* $\{\epsilon_{ni}\}$ *are iid with mean 0, variance* σ_0^2*, and excess kurtosis* κ_0*. Also, the moment* $E|\epsilon_{ni}|^{4+\eta}$ *exists for some* $\eta > 0$*.*

Assumption A2: *i The elements* $\{w_{ij}\}$ *of* W_n *are at most of order* h_n^{-1} *uniformly for all* i, j*, with the rate sequence* $\{h_n\}$ *satisfying* $h_n/n \to 0$ *as* $n \to \infty$*, (ii)* $w_{ii} = 0$ *and* $\sum_j w_{ij} = 1$ *for all i, and (iii) The the row and column sums of* W_n *in absolute value are uniformly bounded.*

Assumption A3: *The elements of the* $n \times k$ *matrix* X_n *are uniformly bounded for all n, and* $\lim_{n\to\infty} \frac{1}{n} X_n' X_n$ *exists and is nonsingular.*

Theorem 3.1. Under Assumptions 1-3, Model (3.17) and H_0: $\rho = 0$,

(*i*) $\mathrm{SLM}_{\mathrm{SE}}^{\circ} \xrightarrow{D} N(0,1)$, as $n \to \infty$,

(*ii*) $\mathrm{SLM}_{\mathrm{SE}}^{\mathrm{MD}} \xrightarrow{D} N(0,1)$, as $n \to \infty$,

(*iii*) $\mathrm{I}_{\mathrm{SE}}^{\circ}$, $\mathrm{LM}_{\mathrm{SE}}^{\mathrm{FI}}$, $\mathrm{LM}_{\mathrm{SE}}^{\mathrm{MD}}$, $\mathrm{SLM}_{\mathrm{SE}}^{\circ}$, and $\mathrm{SLM}_{\mathrm{SE}}^{\mathrm{MD}}$ are asymptotically equivalent.

Neighborhood Crime. See Sec. 2.2.4 for a description of the data and variables, and the construction of the spatial weight matrix.

Consider an SE model for `Crime` with regressors: `constant`, `Income`, `House`, `East`, W_n`Income`, and W_n`House`.

- $\mathrm{LR}_{\mathtt{SE}} =$.
- $\mathrm{t}_{\mathtt{SE}} =$.
- $\mathrm{T}^{\mathrm{MD}}_{\mathtt{CE}} =$, where CE = $\{W_n\mathtt{Income},\ W_n\mathtt{House}\}$
- $\mathrm{I}^{\circ}_{\mathtt{SE}} =$.
- $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SE}} =$.
- $\mathrm{LM}^{\mathrm{MD}}_{\mathtt{SE}} =$.
- $\mathrm{SLM}^{\circ}_{\mathtt{SE}} =$.
- $\mathrm{SLM}^{\mathrm{MD}}_{\mathtt{SE}} =$.

Boston House Price. See Sec. 2.2.4 for a description of the data and variables, and the construction of the spatial weight matrix.

Consider an SE model for `MEDV` including all the regressors, and adding SD-`room` and adding SD-`access`.

- $\mathrm{LR}_{\mathtt{SE}} =$.
- $\mathrm{t}_{\mathtt{SE}} =$.
- $\mathrm{T}^{\mathrm{MD}}_{\mathtt{CE}} =$, where $\mathtt{CE}$ = {SD-`room`, SD-`access`}
- $\mathrm{I}^{\circ}_{\mathtt{SE}} =$.
- $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SE}} =$.
- $\mathrm{LM}^{\mathrm{MD}}_{\mathtt{SE}} =$.
- $\mathrm{SLM}^{\circ}_{\mathtt{SE}} =$.
- $\mathrm{SLM}^{\mathrm{MD}}_{\mathtt{SE}} =$.

SE model may be the simplest in the SLR framework, but there are still results that are desired but unavailable in the literature. These include,

- an LM test of H_0^{SE} based on the robust LM principle described around (3.10).
- an LM test of H_0^{SE} based on the robust LM principle described around (3.14).
- finite sample of these two robust LM tests, expected to perform not as well as the two SLM tests as the principles in (3.10) and (3.14) do not lead to mean corrections.
- finite sample performance of Moran's I, I°_{SE}.
- finite sample performance of $\mathrm{LR}_{\mathrm{SE}}$ and t_{SE}.

Recall: liner regression model with SL dependence given in (2.14):

$$Y_n = \lambda W_n Y_n + X_n\beta + \epsilon_n, \tag{3.36}$$

where Y_n, X_n, and W_n are as in (3.17). The errors $\epsilon_{n,i}$ are iid$(0, \sigma^2)$.

The hypotheses of interest for Model (3.36) concern (i) covariate effect (CE), and (ii) spatial lag (SL) effect:

- $H_0^{\texttt{CE}} : R\beta_0 = r$, some regressors can be merged or dropped,
- $H_0^{\texttt{SL}} : \lambda_0 = 0$, standard liner regression model suffices,

where R is a $q \times k$ constant matrix, and $q \leq k$.

The $H_0^{\texttt{CE}}$ hypothesis covers: insignificance of some covariates effects, non-existence of spatial Durbin effects, etc., with proper choices of the linear restriction matrix R and the vector r.

Recall: the loglikelihood function of $\theta = (\beta', \sigma^2, \lambda)$ given in (2.14):

$$\ell_n(\theta) = -\tfrac{n}{2}\log(2\pi\sigma^2) + \log|A_n(\lambda)| - \tfrac{1}{2\sigma^2}\|\epsilon_n(\beta,\lambda)\|^2, \qquad (3.37)$$

where $\epsilon_n(\beta,\lambda) = A_n(\lambda)Y_n - X_n\beta$, and $A_n(\lambda) = I_n - \lambda W_n$.

From the general form of LR statistic: $\mathrm{LR}_n = -2[\ell_n(\tilde{\theta}_n) - \ell_n(\hat{\theta}_n)]$, where $\tilde{\theta}_n$ and $\hat{\theta}_n$ are the restricted and unrestricted MLEs of θ, we have, similar to the SE model, a simple common form of the LR test statistic for testing either H_0^{CE} or H_0^{SL} in the SL model:

$$\mathrm{LR}_{\varpi} = n\ln(\tilde{\sigma}_n^2\,\hat{\sigma}_n^{-2}) - 2\log|A_n(\tilde{\lambda}_n)A_n^{-1}(\hat{\lambda}_n)|, \quad \text{for } \varpi = \mathrm{CE}, \mathrm{SL}, \qquad (3.38)$$

due to the fact that $\hat{\sigma}_n^{-2}\|\epsilon_n(\hat{\beta}_n,\hat{\lambda}_n)\|^2 = \tilde{\sigma}_n^{-2}\|\epsilon_n(\tilde{\beta}_n,\tilde{\lambda}_n)\|^2 = \frac{n}{2}$.

- Under H_0^{CE}, $\mathrm{LR}_{\mathrm{CE}} \overset{a}{\sim} \chi_q^2$; under H_0^{SL}, $\lambda = 0$ and $\mathrm{LR}_{\mathrm{SL}} \overset{a}{\sim} \chi_1^2$.
- Note: $\mathrm{LR}_{\mathrm{SL}} = -2[\ell_n^c(0) - \ell_n^c(\hat{\lambda}_n)]$, where $\ell_n^c(\lambda)$ is given in (2.19).

Recall: from (2.21) and (2.22), the expected negative Hessian:

$$\mathcal{J}_{\mathrm{SL}} = \begin{pmatrix} \frac{1}{\sigma_0^2} X_n' X_n & 0 & \frac{1}{\sigma_0} X_n' \eta_n \\ 0 & \frac{n}{2\sigma_0^4} & \frac{1}{\sigma_0^2} \mathrm{tr}(F_n) \\ \frac{1}{\sigma_0} \eta_n' X_n & \frac{1}{\sigma_0^2} \mathrm{tr}(F_n) & \eta_n' \eta_n + \mathrm{tr}(F_n^s F_n) \end{pmatrix}, \tag{3.39}$$

and the VC matrix of quasi score function: $\mathcal{I}_{\mathrm{SL}} = \mathcal{J}_{\mathrm{SL}} + \mathcal{K}_{\mathrm{SL}}$, where

$$\mathcal{K}_{\mathrm{SL}} = \begin{pmatrix} 0 & \frac{1}{2\sigma_0^3} \gamma_0 X_n' \iota_n & \frac{1}{\sigma_0} \gamma_0 X_n' f_n \\ \sim & \frac{n}{4\sigma_0^4} \kappa_0 & \frac{1}{2\sigma_0^2} \gamma_0 \iota_n' \eta_n + \frac{1}{2\sigma_0^2} \kappa_0 \mathrm{tr}(F_n) \\ \sim & \sim & \kappa_0 f_n' f_n + 2\gamma_0 f_n' \eta_n \end{pmatrix}, \tag{3.40}$$

where $f_n = \mathrm{diagv}(F_n)$, $F_n^s = F_n + F_n'$, $\eta_n = \sigma_0^{-1} G_n X_n \beta_0$, and γ_0 and κ_0 are the skewness and excess kurtosis of $\epsilon_{n,i}$.

With the QMLE $\hat{\theta}_n$, and the plug-in estimators $\hat{\mathcal{I}}_{\mathrm{SL}} = \mathcal{I}_{\mathrm{SL}}(\hat{\theta}_n)$ and $\hat{\mathcal{J}}_{\mathrm{SL}} = \mathcal{J}_{\mathrm{SL}}(\hat{\theta}_n)$ given in Sec. 2.2.2, various Wald tests can be constructed.

Wald tests for covariate effects. Let $\widehat{V}_{\text{SL}} = \hat{\mathcal{J}}_{\text{SL}}^{-1}\hat{\mathcal{I}}_{\text{SL}}\hat{\mathcal{J}}_{\text{SL}}^{-1}$. Let $\widehat{V}_{\text{SL},\beta\beta}$ and $\hat{V}_{\text{SL},\lambda\lambda}$ be, respectively, the β-β and λ-λ diagonal blocks of $\widehat{V}_{\text{SL}}$.

- Wald test for testing H_0^{CE}: $R\beta_0 = r$ has the expression:

$$\mathrm{T}_{\text{CE}} = (R\hat{\beta}_n - r)'(RV_{\text{SL},\beta\beta}R')^{-1}(R\hat{\beta}_n - r). \tag{3.41}$$

  Under H_0^{CE}, $\mathrm{T}_{\text{CE}} \xrightarrow{D} \chi_q^2$, where q is the number of rows of R.

- When R is a row vector ($q = 1$), Wald test reduces to a t-test:

$$\mathrm{t}_{\text{CE}} = \frac{R\hat{\beta}_n - r}{\sqrt{R\hat{V}_{\text{SL},\beta\beta}R'}} \xrightarrow{D} N(0,1), \ \text{under } H_0^{\text{CE}}. \tag{3.42}$$

Finite sample performance of T_{CE} and t_{CE} can be poor, because $\hat{\lambda}_n$ is downward biased (Yang, 2015), which passes to $\hat{\beta}_n$ as seen from below:

$$\hat{\beta}_n = \tilde{\beta}_n(\hat{\lambda}_n) = \beta_0 + (\lambda_0 - \hat{\lambda}_n)(X_n'X_n)^{-1}X_n'G_nX_n\beta_0 + o_p(1),$$

and thus causes the variance estimate $\hat{V}_{\text{SL},\beta\beta}$ to be biased.

The poor finite sample performance of T_{CE} and t_{CE} is confirmed by the Monte Carlo results given in Liu and Yang (2015a), where improved inferences methods are provided.

Wald Test for spatial effect. Similarly, with the λ-λ element of $\hat{V}_{\mathrm{SL}}$, $\hat{V}_{\mathrm{SL},\lambda\lambda}$, we obtain a Wald statistic for λ, which is asymptotic $N(0,1)$,

$$\mathrm{t}_{\mathrm{SL}}(\lambda_0) = \frac{\hat{\lambda}_n - \lambda_0}{\sqrt{\hat{V}_{\mathrm{SL},\lambda\lambda}}}, \tag{3.43}$$

- The statistic $t_{\mathrm{SL}}(\lambda_0)$ can be used to test H_0: $\lambda = 0$.
- Finite sample property of $\mathrm{t}_{\mathrm{SL}}(\lambda_0)$ is studied by Yang (2015),
- along with the improved tests based on bias-corrections.

Recall: the (quasi) score function given in (2.17):

$$S_{\mathrm{SL}}(\theta) = \begin{cases} \frac{1}{\sigma^2} X_n' \epsilon_n(\beta, \lambda), \\ \frac{1}{2\sigma^4} \epsilon_n'(\beta, \lambda)\epsilon_n(\beta, \lambda) - \frac{n}{2\sigma^2}, \\ \frac{1}{\sigma^2} Y_n' W_n' \epsilon_n(\beta, \lambda) - \mathrm{tr}[F_n(\lambda)], \end{cases} \tag{3.44}$$

where $F_n(\lambda) = W_n A_n^{-1}(\lambda)$.

Under normality: we have IME, $\mathcal{I}_n = \mathcal{J}_n$, and we obtain three variants of LM tests for a general hypothesis on θ as those for SE model:

$$\mathrm{LM_{FI}} = S_{\mathrm{SL}}'(\tilde{\theta}_n)\mathcal{I}_{\mathrm{SL}}(\tilde{\theta}_n)^{-1} S_{\mathrm{SL}}(\tilde{\theta}_n), \tag{3.45}$$

$$\mathrm{LM_{OI}} = S_{\mathrm{SL}}'(\tilde{\theta}_n)\Big[- \frac{\partial}{\partial \theta'} S_{\mathrm{SL}}(\theta)\Big|_{\theta=\tilde{\theta}_n}\Big]^{-1} S_{\mathrm{SL}}(\tilde{\theta}_n), \tag{3.46}$$

$$\mathrm{LM_{MD}} = S_{\mathrm{SL}}'(\tilde{\theta}_n)\Big[\sum_{i=1}^n \tilde{g}_{ni}\tilde{g}_{ni}'\Big]^{-1} S_{\mathrm{SL}}(\tilde{\theta}_n), \tag{3.47}$$

if in (3.47), $S_{\mathrm{SL}}(\theta_0) = \sum_{i=1}^n g_{ni}(\theta_0)$ and $\{g_{ni}(\theta_0)\}$ is an M.D. sequence. All three statistics are χ_q^2 distributed under H_0 with q restrictions.

LM tests for spatial effect. For testing H_0^{SL}: $\lambda = 0$, LM tests are preferred as they require only the OLS estimates.

Anselin (1988) presents an LM test based on the expected information:

$$\mathrm{LM}_{\mathrm{SL}}^{\mathrm{FI}} = \frac{\tilde{\epsilon}_n' W_n Y_n}{\tilde{\sigma}_n^2 \sqrt{\tilde{D}_n + T_n}}, \tag{3.48}$$

where $T_n = \mathrm{tr}[(W_n + W_n')W_n]$, $\tilde{D}_n = \tilde{\sigma}_n^{-2}(W_n X_n \tilde{\beta}_n)' M_n W_n X_n \tilde{\beta}_n$, $\tilde{\beta}_n$ and $\tilde{\sigma}_n^2$ are the OLS estimates, and $\tilde{\epsilon}_n$ are the OLS residuals.

- Derivation of $\mathrm{LM}_{\mathrm{SL}}^{\mathrm{FI}}$ is based on $S_{\mathrm{SL},\lambda}(\theta)$ and $[\mathcal{J}_{\mathrm{SL}}^{-1}(\theta)]_{\lambda\lambda}$, evaluated at $\tilde{\theta}_n = (\tilde{\beta}_n'(0), \tilde{\sigma}_n^2(0), 0)'$, where $\tilde{\beta}_n'(0)$ and $\tilde{\sigma}_n^2(0)$ are given in (2.18).
- Although $\mathrm{LM}_{\mathrm{SL}}^{\mathrm{FI}}$ is derived under normality, one can shown that it is robust against non-normality (NN).
- Its finite sample performance can be poor as the effect of estimating β and σ^2 is not taken into account.
- The OI (observed information) variant can easily be derived as well.

Paralleled with the OPMD-based LM test (3.33), Born and Breitung (2011) also derived an OPMD-based LM test for testing H_0^{SL}:

$$\mathrm{LM}_{\mathrm{SL}}^{\mathrm{MD}} = \frac{\tilde{\epsilon}_n' W_n Y_n}{\sqrt{(\tilde{\epsilon}_n \odot \tilde{\epsilon}_n)'(\tilde{\xi}_n \odot \tilde{\xi}_n)}}, \tag{3.49}$$

where $\tilde{\xi}_n = (W_n^l + W_n^{u\prime})\tilde{\epsilon}_n$, W_n^l and W_n^u are the lower and upper triangular matrices such that $W_n^l + W_n^u = W_n$, and $\mathrm{LM}_{\mathrm{SL}}^{\mathrm{MD}}|_{H_0} \xrightarrow{D} N(0,1)$.

- Again, this test is robust against NN.
- The denominator of the test statistic can also been seen to be robust against unknown heteroskedasticity (UH).
- It finite sample performance may be poor for the same reason.
- LM-type tests with better finite sample performance are desired.

The standardized version of $\mathrm{LM}_{\mathtt{SL}}^{\mathtt{FI}}$, denoted as $\mathrm{SLM}_{\mathtt{SL}}^{\mathtt{FI}}$, in the spirit of $\mathrm{SLM}_{\mathtt{SE}}^{\mathtt{FI}}$ statistic given in (3.34), is of interest but unavailable.

To derive $\mathrm{SLM}_{\mathtt{SL}}^{\mathtt{FI}}$, note that $S_{\mathtt{SL},\lambda}(\theta)|_{\theta=\tilde{\theta}_n}$ is proportional to

$$\tilde{\epsilon}_n M_n W_n Y_n = \epsilon_n' M_n W_n \epsilon_n + \epsilon_n' M_n W_n X_n \beta.$$

Standardize this quantity, we obtain:

$$\mathrm{SLM}_{\mathtt{SL}}^{\mathtt{FI}} = ???, \tag{3.50}$$

where $? \cdots ?$, and $\tilde{\gamma}_n$ and $\tilde{\kappa}_n$ are the sample skewness and the excess of OLS residuals $\tilde{\epsilon}_n$.

You are strongly encouraged to complete the derivation for $\mathrm{SLM}_{\mathtt{SL}}^{\mathtt{FI}}$.

An OPMD variant of SLM test can be obtained along the line of (3.35):

$$\mathrm{SLM}_{\mathrm{SL}}^{\mathrm{MD}} = ???, \tag{3.51}$$

where $? \cdots ?$. You are strongly encouraged to complete this derivation.

As commented for the SE model, the tests $\mathrm{SLM}_{\mathrm{SL}}^{\mathrm{FI}}$ and $\mathrm{SLM}_{\mathrm{SL}}^{\mathrm{FI}}$ suggested above may not be truly "standardized LM tests", as mean corrections are not made.

Theorem 3.2. Under Assumptions 1-3, Model (3.17) and H_0: $\rho = 0$,

(*i*) $\mathrm{SLM}_{\mathrm{SL}}^{\circ} \xrightarrow{D} N(0, 1)$, as $n \to \infty$,

(*ii*) $\mathrm{SLM}_{\mathrm{SL}}^{\mathrm{MD}} \xrightarrow{D} N(0, 1)$, as $n \to \infty$,

(*iii*) $\mathrm{LM}_{\mathrm{SL}}^{\mathrm{FI}}$, $\mathrm{LM}_{\mathrm{SL}}^{\mathrm{MD}}$, $\mathrm{SLM}_{\mathrm{SL}}^{\circ}$, and $\mathrm{SLM}_{\mathrm{SL}}^{\mathrm{MD}}$ are asymptotically equivalent.

Neighborhood Crime. See Sec. 2.2.4 for a description of the data and variables, and the construction of the spatial weight matrix.

Consider an SL model for `Crime` with regressors: `constant`, `Income`, `House`, `East`, W_n`Income`, and W_n`House`.

- $\mathrm{LR}_{\mathtt{SL}} =$.
- $\mathrm{t}_{\mathtt{SL}} =$.
- $\mathrm{T}^{\mathrm{MD}}_{\mathtt{CE}} =$, where `CE` = $\{W_n$`Income`, W_n`House`$\}$
- $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SL}} =$.
- $\mathrm{LM}^{\mathrm{MD}}_{\mathtt{SL}} =$.
- $\mathrm{SLM}^{\circ}_{\mathtt{SL}} =$.
- $\mathrm{SLM}^{\mathrm{MD}}_{\mathtt{SL}} =$.

Boston House Price. See Sec. 2.2.4 for a description of the data and variables, and the construction of the spatial weight matrix.

Consider an SL model for MEDV including all the regressors, and adding SD-room and adding SD-access.

- $\mathrm{LR}_{\mathtt{SL}} =$.
- $\mathrm{t}_{\mathtt{SL}} =$.
- $\mathrm{T}^{\mathrm{MD}}_{\mathtt{CE}} =$, where CE = {SD-room, SD-access}
- $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SL}} =$.
- $\mathrm{LM}^{\mathrm{MD}}_{\mathtt{SL}} =$.
- $\mathrm{SLM}^{\circ}_{\mathtt{SL}} =$.
- $\mathrm{SLM}^{\mathrm{MD}}_{\mathtt{SL}} =$.

The SL model is another simple model in the SLR framework and is more popular than the SE model. However, there are more results that are desired but unavailable in the literature. These include,

- an LM test of H_0^{SL} by the robust LM principle described around (3.10).
- an LM test of H_0^{SL} by the robust LM principle described around (3.14).
- an SLM test along the line of (3.34).
- an SLM test along the line of (3.35).
- a Monte Carlo comparison of these tests for their finite sample performance, to recommend to practitioners a simple and reliable test for spatial lag dependence.

Recall: The SLR model with both SL and SE (SLE) given in (2.25):

$$Y_n = \lambda W_{1n} Y_n + X_n\beta + u_n, \quad u_n = \rho W_{2n} u_n + \epsilon_n, \tag{3.52}$$

where all quantities are defined as in the SL and SE models. This model involves two spatial weight matrices W_{1n} and W_{2n}, which can be the same.

The hypotheses of interest for Model (3.52) concern (i) covariate effect (CE), and (ii) spatial lag (SL) and/or spatial error (SE) effects:

- $H_0^{\mathtt{CE}} : R\beta_0 = r$, some regressors can be merged or dropped,
- $H_0^{\mathtt{SLE}} : \delta_0 = (\lambda_0, \rho_0)' = 0$, standard liner regression model suffices,
- $H_0^{\mathtt{SL|SE}} : \lambda_0 = 0$, SE model suffices,
- $H_0^{\mathtt{SE|SL}} : \rho_0 = 0$, SL model suffices,

where R is a $q \times k$ constant matrix, and $q \leq k$.

Recall: the log-likelihood function of $\theta = (\beta', \sigma^2, \lambda, \rho)'$ given in (2.26):

$$\ell_n(\theta) = -\frac{n}{2}\log(2\pi\sigma^2) + \log|A_n(\lambda)| + \log|B_n(\rho)| - \frac{1}{2\sigma^2}\|\epsilon_n(\beta,\delta)\|^2, \quad (3.53)$$

- $\epsilon_n(\beta,\delta) = \mathbb{Y}_n(\delta) - \mathbb{X}_n(\rho)\beta$, $\mathbb{Y}_n(\delta) = B_n(\rho)A_n(\lambda)Y_n$, and $\mathbb{X}_n(\rho) = B_n(\rho)X_n$.

Based on the LR principle introduced in Sec. 3.1., we show that the LR statistic for testing H_0^{CE}, H_0^{SLE}, $H_0^{\text{SL|SE}}$, or $H_0^{\text{SE|SL}}$ takes a common form:

$$\text{LR}_\varpi = n\ln(\tilde{\sigma}_n^2\,\hat{\sigma}_n^{-2}) - 2\log|A_n(\tilde{\lambda}_n)A_n^{-1}(\hat{\lambda}_n)| - 2\log|B_n(\tilde{\rho}_n)B_n^{-1}(\hat{\rho}_n)|, \quad (3.54)$$

for $\varpi = \text{CE}, \text{SLE}, \text{SL|SE}, \text{SE|SL}$, where $\tilde{\omega}_n$ and $\hat{\omega}_n$ are the restricted (under the null hypothesis) and unrestricted MLEs of a parameter ω, respectively.

- The limiting null distribution of LR_ϖ is χ^2_{df}, with $df = q, 2, 1$ and 1, respectively, for the four tests.
- The LR tests for spatial effects can be formulated directly from $\ell_n^c(\delta)$ given in (2.19), e.g., $\text{LR}_{\text{SLE}} = -2[\ell_n^c(0) - \ell_n^c(\hat{\delta}_n)]$.

Recall: from (2.33) and (2,34), the expected negative Hessian:

$$\mathcal{J}_{\texttt{SLE}} = \begin{pmatrix} \frac{1}{\sigma_0^2}\mathbb{X}_n'\mathbb{X}_n & 0 & \frac{1}{\sigma_0}\mathbb{X}_n'\mu_n & 0 \\ \sim & \frac{n}{2\sigma_0^4} & \frac{1}{\sigma_0^2}\mathrm{tr}(F_n) & \frac{1}{\sigma_0^2}\mathrm{tr}(G_n) \\ \sim & \sim & \mu_n'\mu_n + \mathrm{tr}(\bar{F}_n^s\bar{F}_n) & \mathrm{tr}(G_n^s\bar{F}_n) \\ \sim & \sim & \sim & \mathrm{tr}(G_n^s G_n) \end{pmatrix}, \quad (3.55)$$

and the VC matrix of QS function: $\mathcal{I}_{\texttt{SLE}} = \mathcal{J}_{\texttt{SLE}} + \mathcal{K}_{\texttt{SLE}}$, where

$$\mathcal{K}_{\texttt{SLE}} = \begin{pmatrix} 0 & \frac{\gamma_0}{2\sigma_0^3}\mathbb{X}_n'\iota_n & \frac{\gamma_0}{\sigma_0}\mathbb{X}_n'\bar{f}_n & \frac{\gamma_0}{\sigma_0}\mathbb{X}_n' g_n \\ \sim & \frac{n\kappa_0}{4\sigma_0^4} & \frac{\kappa_0}{2\sigma_0^2}\mathrm{tr}(F_n) + \frac{\gamma_0}{2\sigma_0^2}\iota_n'\mu_n & \frac{\kappa_0}{2\sigma_0^2}\mathrm{tr}(G_n) \\ \sim & \sim & \kappa_0\bar{f}_n'\bar{f}_n + 2\gamma_0\bar{f}_n'\mu_n & \kappa_0 g_n'\bar{f}_n + \gamma_0 g_n'\mu_n \\ \sim & \sim & \sim & \kappa_0 g_n' g_n \end{pmatrix}, \quad (3.56)$$

$\mu_n = \sigma_0^{-1}B_nF_nX_n\beta_0$, $\bar{F}_n = B_nF_nB_n^{-1}$, $\bar{f}_n = \mathrm{diag}(\bar{F}_n)$, $\bar{F}_n^s = \bar{F}_n + \bar{F}_n'$, and others quantities are defined earlier.

Similar to the LR test, the Wald test for any hypothesis concerning θ in the SLE model can be written in the following general form. Denote the plug-in estimate of asymptotic VC matrix of the QMLE $\hat{\theta}_n$ as

$$\widehat{V}_{\mathtt{SLE}} = \mathcal{J}_{\mathtt{SLE}}^{-1}(\hat{\theta}_n)\mathcal{I}_{\mathtt{SLE}}(\hat{\theta}_n)\mathcal{J}_{\mathtt{SLE}}^{-1}(\hat{\theta}_n).$$

First, a Wald statistic for inference for a general linear combination $c'\theta_0$, where c is a $(k+2)\times 1$ constant vector, is a univariate t-statistic:

$$\mathrm{t}_{\mathtt{SLE}}(\theta_0) = \frac{c'\hat{\theta}_n - c'\theta_0}{\sqrt{c'\widehat{V}_{\mathtt{SLE}}c}} \overset{a}{\sim} N(0,1), \tag{3.57}$$

- $\mathrm{t}_{\mathtt{SLE}}(\theta_0)$ can be used to construct a CI for $c'\theta_0$.
- It can also be used to test the hypothesis of $c'\theta_0 = 0$.
- When c contains a single non-zero value 1, $\mathrm{t}_{\mathtt{SLE}}(\theta_0)$ is a t-statistic for a single parameter, and can be used to test $H_0^{\mathtt{SL|SE}}$ or $H_0^{\mathtt{SE|SL}}$ (?).
- When the elements of c adds to 0, it provides inference method for a linear contrast on θ_0, e.g., $\beta_{10} - \beta_{20}$.

Let $\widehat{V}_{\texttt{SLE},\beta\beta}$ and $\widehat{V}_{\texttt{SLE},\delta\delta}$ be the β-β and δ-δ diagonal blocks of $\widehat{V}_{\texttt{SLE}}$.

Wald test for testing H_0^{CE}: $R\beta_0 = r$ has the expression:

$$\mathrm{T}_{\texttt{CE}} = (R\hat{\beta}_n - r)'(RV_{n\beta\beta}R')^{-1}(R\hat{\beta}_n - r). \qquad (3.58)$$

Under H_0^{CE}, $\mathrm{T}_{\texttt{CE}} \xrightarrow{D} \chi^2_q$, where q is the number of rows of R.

Similarly, a Wald statistic for inferences for δ is

$$\mathrm{T}_{\texttt{SLE}}(\delta_0) = (\hat{\delta}_n - \delta_0)'\widehat{V}^{-1}_{\texttt{SLE},\delta\delta}(\hat{\delta}_n - \delta_0) \overset{a}{\sim} \chi^2_2. \qquad (3.59)$$

This test statistic can be used to test H_0: $\delta = 0$. One directional tests for λ (allowing ρ) and for ρ (allowing λ) can easily be formulated.

For testing H_0: $\delta = 0$, LM tests are preferred as they require only the estimates of the null model, which in this case is the OLS regression.

Recall: the (quasi) score function given in (2.27):

$$S_{\mathrm{SLE}}(\theta) = \begin{cases} \frac{1}{\sigma^2}\mathbb{X}_n'(\rho)\epsilon_n(\beta,\delta), \\ \frac{1}{2\sigma^4}\epsilon_n'(\beta,\lambda)\epsilon_n(\beta,\delta) - \frac{n}{2\sigma^2}, \\ \frac{1}{\sigma^2}Y_n'W_{1n}'B_n'(\rho)\epsilon_n(\beta,\delta) - \mathrm{tr}[F_n(\lambda)], \\ \frac{1}{\sigma^2}\epsilon_n'(\beta,\delta)G_n(\rho)\epsilon_n(\beta,\delta) - \mathrm{tr}[G_n(\rho)], \end{cases} \tag{3.60}$$

where $F_n(\lambda) = W_{1n}A_n^{-1}(\lambda)$ and $G_n(\rho) = W_{2n}B_n^{-1}(\rho)$.

Under normality: the LM principle introduced in Sec. 3.1.3 again leads to three variants of LM tests for a general hypothesis on θ:

$$\mathrm{LM}_{\mathrm{FI}} = S_{\mathrm{SLE}}'(\tilde{\theta}_n)\mathcal{I}_{\mathrm{SLE}}(\tilde{\theta}_n)^{-1}S_{\mathrm{SLE}}(\tilde{\theta}_n), \tag{3.61}$$

$$\mathrm{LM}_{\mathrm{OI}} = S_{\mathrm{SLE}}'(\tilde{\theta}_n)\Big[-\frac{\partial}{\partial\theta'}S_{\mathrm{SLE}}(\theta)\Big|_{\theta=\tilde{\theta}_n}\Big]^{-1}S_{\mathrm{SLE}}(\tilde{\theta}_n), \tag{3.62}$$

$$\mathrm{LM}_{\mathrm{MD}} = S_{\mathrm{SLE}}'(\tilde{\theta}_n)\Big[\sum_{i=1}^n \tilde{g}_{ni}\tilde{g}_{ni}'\Big]^{-1}S_{\mathrm{SLE}}(\tilde{\theta}_n), \tag{3.63}$$

if in (3.65), $S_{\mathrm{SLE}}(\theta_0) = \sum_{i=1}^n g_{ni}(\theta_0)$ and $\{g_{ni}(\theta_0)\}$ is an M.D. sequence.

FI-based LM test. Anselin (1988) gives an LM test of H_0^{SLE}, based on Fisher information matrix $\mathcal{J}_{\mathrm{SLE}}$ given in (3.55):

$$\mathrm{LM}_{\mathrm{SLE}}^{\mathrm{FI}} = \frac{1}{\tilde{\sigma}_n^4} \begin{pmatrix} \tilde{\epsilon}_n' W_{1n} Y_n \\ \tilde{\epsilon}_n' W_{2n} \tilde{\epsilon}_n \end{pmatrix}' \begin{pmatrix} T_{1n} + \tilde{D}_n & T_{3n} \\ T_{3n} & T_{2n} \end{pmatrix}^{-1} \begin{pmatrix} \tilde{\epsilon}_n' W_{1n} Y_n \\ \tilde{\epsilon}_n' W_{2n} \tilde{\epsilon}_n \end{pmatrix}, \quad (3.64)$$

where $T_{1n} = \mathrm{tr}[(W_{1n} + W_{1n}')W_{1n}]$, $\quad T_{2n} = \mathrm{tr}[(W_{2n} + W_{2n}')W_{2n}]$,
$T_{3n} = \mathrm{tr}[(W_{2n} + W_{2n}')W_{1n}]$, $\quad \tilde{D}_n$ is defined in (3.48).

OPMD-based LM test. Born and Breitung (2011) give an OPMD variant:

$$\mathrm{LM}_{\mathrm{SLE}}^{\mathrm{MD}} = \begin{pmatrix} \tilde{\varepsilon}_n' W_{1n} Y_n \\ \tilde{\varepsilon}_n' W_{2n} \tilde{\varepsilon}_n \end{pmatrix}' \begin{pmatrix} \tilde{\epsilon}_n^{2\,\prime} \tilde{\xi}_{1n}^2 & \tilde{\epsilon}_n^{2\,\prime}(\tilde{\xi}_{1n} \odot \tilde{\xi}_{2n}) \\ \tilde{\epsilon}_n^{2\,\prime}(\tilde{\xi}_{1n} \odot \tilde{\xi}_{2n}) & \tilde{\epsilon}_n^{2\,\prime} \tilde{\xi}_{2n}^2 \end{pmatrix}^{-1} \begin{pmatrix} \tilde{\varepsilon}_n' W_{1n} Y_n \\ \tilde{\varepsilon}_n' W_{2n} \tilde{\varepsilon}_n \end{pmatrix}, \quad (3.65)$$

- **square of a vector:** e.g., $\tilde{\varepsilon}_n^2 = \tilde{\varepsilon}_n \odot \tilde{\varepsilon}_n$,
- $\tilde{\xi}_{1n} = (W_{1n}^{u\prime} + W_{1n}^{l})\tilde{\epsilon}_n + M_n W_n X_n \tilde{\beta}_n$, $\ \tilde{\xi}_{2n} = (W_{2n}^{u\prime} + W_{2n}^{l})\tilde{\epsilon}_n$,
- W_{rn}^u and W_{rn}^l: the upper and lower triangular matrices of $W_{rn}, r = 1, 2$.

A general formulation of FI-based LM tests for testing $H_0^{\mathtt{SLE}}$, $H_0^{\mathtt{SL|SE}}$ and $H_0^{\mathtt{SE|SL}}$ can be obtained by applying (3.11) and (3.12):

$$\mathrm{LM}_{\mathtt{SLE}}^{\mathtt{FI}}(\delta) = \tilde{S}'_{\mathtt{SLE},\delta}(\delta)[\tilde{\mathcal{J}}_{\mathtt{SLE}}^{-1}(\delta)]_{\delta\delta}\tilde{S}_{\mathtt{SLE},\delta}(\delta), \quad (3.66)$$

where the *tilde* quantities are the estimates at $\tilde{\beta}_n(\delta)$ and $\tilde{\sigma}_n^2(\delta)$ given in (2.28) and (2.29).

- Taking $\delta = 0$, $\mathrm{LM}_{\mathtt{SLE}}^{\mathtt{FI}}(\delta)$ reduces to the joint test $\mathtt{LM}_{\mathtt{SLE}}^{\mathtt{FI}}$ given in (3.64).
- Taking $\delta = (0, \tilde{\rho}_n)'$, it gives a marginal test, $\mathtt{LM}_{\mathtt{SL|SE}}^{\mathtt{FI}}$, for testing $H_0^{\mathtt{SL|SE}}$: $\lambda = 0$ in the SLE model.
- Taking $\delta = (\tilde{\lambda}, 0)'$, it gives a marginal test, $\mathtt{LM}_{\mathtt{SE|SL}}^{\mathtt{FI}}$, for testing $H_0^{\mathtt{SE|SL}}$: $\rho = 0$ in the SLE model.
- The two marginal tests $\mathtt{LM}_{\mathtt{SL|SE}}^{\mathtt{FI}}$ and $\mathtt{LM}_{\mathtt{SE|SL}}^{\mathtt{FI}}$ are unlikely robust to NN.

The quantity, $[\tilde{\mathcal{J}}_{\mathtt{SLE}}^{-1}(\delta)]_{\delta\delta}$, can be simplified!

You are strongly encouraged to complete the derivations for the test statistics $\mathtt{LM}_{\mathtt{SL|SE}}^{\mathtt{FI}}$ and $\mathtt{LM}_{\mathtt{SE|SL}}^{\mathtt{FI}}$ to make them as compact as possible!

The standardized LM tests for the joint hypothesis, H_0^{SLE}: $\delta = 0$, can be obtained by combining the ideas behind $\mathrm{SLM}^{\circ}_{\mathrm{SE}}$ and $\mathrm{SLM}^{\circ}_{\mathrm{SL}}$, and the ideas behind $\mathrm{SLM}^{\mathrm{DM}}_{\mathrm{SE}}$ and $\mathrm{SLM}^{\mathrm{DM}}_{\mathrm{SL}}$. **Note:**

$$\tilde{S}'_{\mathrm{SLE},\delta}(\delta)|_{\delta=0} \propto \begin{cases} \tilde{\epsilon}'_n W_{1n} Y_n = \epsilon'_n M_n W_{1n} \epsilon_n + \epsilon'_n M_n W_{1n} X_n \beta_0, \\ \tilde{\epsilon}'_n W_{2n} \tilde{\epsilon}_n = \epsilon'_n M_n W_{2n} M_n \epsilon_n. \end{cases}$$

This would lead to the two variants of SLM tests for H_0^{SLE}: $\delta = 0$:

$$\mathrm{SLM}^{\circ}_{\mathrm{SLE}} = ??? \tag{3.67}$$

$$\mathrm{SLM}^{\mathrm{MD}}_{\mathrm{SLE}} = ??? \tag{3.68}$$

These tests are robust against NN, by construction.

You are encouraged to complete these derivations, and perhaps conduct a Monte Carlo study on the finite sample performance of these two tests.

The ideas for deriving an SLM tests introduced above may not be easily extended to give marginal SLM tests for the hypotheses, $H_0^{\mathtt{SL|SE}}$: $\lambda_0 = 0$ in the SLE model and $H_0^{\mathtt{SE|SL}}$: $\rho_0 = 0$ in the SLE model. This is because the 'working QS function', $\tilde{S}_{\mathtt{SLE},\delta}(\tilde{\delta}_n)$, is not linear-quadratic in ϵ_n, either when $\tilde{\delta}_n = (0, \tilde{\rho}_n)'$ under $H_0^{\mathtt{SL|SE}}$ or when $\tilde{\delta}_n = (\tilde{\lambda}_n, 0)'$ under $H_0^{\mathtt{SE|SL}}$.

However, following the principle laid out around (3.10), one obtains:

$$\mathrm{SLM}^{\circ}_{\mathtt{SL|SE}} = ??? \quad (3.69)$$

$$\mathrm{SLM}^{\circ}_{\mathtt{SE|SL}} = ??? \quad (3.70)$$

Following the principle laid out around (3.14), one obtains:

$$\mathrm{SLM}^{\mathrm{MD}}_{\mathtt{SL|SE}} = ??? \quad (3.71)$$

$$\mathrm{SLM}^{\mathrm{MD}}_{\mathtt{SE|SL}} = ??? \quad (3.72)$$

You may like to consider a rigorous study on these conditional SLM tests, under the topic: *"Robust LM Tests for Marginal Spatial Effects in Spatial Linear Regression Models"*, as a Research Paper for the course.

Neighborhood Crime. See Sec. 2.2.4 for a description of the data and variables, and the construction of the spatial weight matrix.

Consider an SLE model for `Crime` with regressors: `constant`, `Income`, `House`, `East`, W_n`Income`, and W_n`House`.

- $\mathrm{LR}_{\mathtt{SLE}} =$.
- $\mathrm{T}_{\mathtt{SLE}} =$.
- $\mathrm{T}^{\mathrm{MD}}_{\mathtt{CE}} =$, where $\mathtt{CE} = \{W_n\mathtt{Income},\ W_n\mathtt{House}\}$
- $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SLE}} =$.
- $\mathrm{LM}^{\mathrm{MD}}_{\mathtt{SLE}} =$.
- $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SL|SE}} =$.
- $\mathrm{LM}^{\mathtt{FI}}_{\mathtt{SE|SL}} =$.
- $\mathrm{SLM}^{\circ}_{\mathtt{SLE}} =$.
- $\mathrm{SLM}^{\mathrm{MD}}_{\mathtt{SLE}} =$.

Boston House Price. See Sec. 2.2.4 for a description of the data and variables, and the construction of the spatial weight matrix.

Consider an SLE model for `MEDV` including all the regressors, and adding SD-`room` and adding SD-`access`.

- $\mathrm{LR}_{\mathtt{SLE}} =$.
- $\mathrm{t}_{\mathtt{SLE}} =$.
- $\mathrm{T}_{\mathtt{CE}}^{\mathrm{MD}} =$, where `CE` = {SD-`room`, SD-`access`}
- $\mathrm{LM}_{\mathtt{SLE}}^{\mathtt{FI}} =$.
- $\mathrm{LM}_{\mathtt{SLE}}^{\mathrm{MD}} =$.
- $\mathrm{SLM}_{\mathtt{SLE}}^{\circ} =$.
- $\mathrm{SLM}_{\mathtt{SLE}}^{\mathrm{MD}} =$.

The SL model is another simple model in the SLR framework and is more popular than the SE model. However, there are more results that are desired but unavailable in the literature. These include,

- an LM test of $H_0^{\mathtt{SLE}}$ based on the robust LM principle (3.10).
- an LM test of $H_0^{\mathtt{SLE}}$ based on the robust LM principle (3.14).
- an SLM test of $H_0^{\mathtt{SLE}}$ in line of (3.34) and (3.50).
- an SLM test of $H_0^{\mathtt{SLE}}$ in line of (3.35) and (3.51).
- an LM test of $H_0^{\mathtt{SL|SE}}$ based on the robust LM principle (3.10).
- an LM test of $H_0^{\mathtt{SE|SL}}$ based on the robust LM principle (3.14).
- a full development of the four marginal spatial tests suggested in (3.69-3.72) including mean corrections, which may be a publishable research topic.

References

- Anselin, L., 1988. *Spatial Econometrics: Methods and Models*. The Netherlands: Kluwer Academic Publishers.
- Baltagi, B., Yang, Z. L., 2013a. Standardized LM tests for spatial error dependence in linear or panel regressions. *The Econometrics Journal* 16 103-134.
- Baltagi, B., Yang, Z. L., 2013b. Heteroskedasticity and non-normality robust LM tests of spatial dependence. *Regional Science and Urban Economics* 43, 725-739.
- Born, B., Breitung, J., 2011. Simple regression based tests for spatial dependence. *Econometrics Journal* 14, 330-342.
- Burridge, P. 1980. On the Cliff-Ord test for spatial autocorrelation. *Journal of the Royal Statistical Society* B 42, 107-8.
- Cameron, A. C., Trivedi, P. K., 2005. *Microeconometrics: Methods and Applications*. Cambridge University Press.

- Cliff, A., Ord, J. K., 1972. Testing for spatial autocorrelation among regression residuals. *Geographical Analysis* 4, 267-284.
- Liu, S. F., Yang, Z. L., 2015a. Asymptotic distribution and finite-sample bias correction of QML estimators for spatial error dependence model. *Econometrics* 3, 376-411.
- Liu, S. F., Yang, Z. L., 2015b. Improved inferences for spatial regression models. *Regional Science and Urban Economics* 55, 55-67.
- Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. *Biometrika* 37, 17-33.
- Yang, Z. (2015). A general method for third-order bias and variance correction on a nonlinear estimator. *Journal of Econometrics* 186, 178-200.