Asymptotically efficient root estimators for spatial autoregressive models with spatial autoregressive disturbances

Fei Jin*[1] and Lung-fei Lee[2]

[1]School of Economics, Fudan University, and Shanghai Institute of International Finance and Economics, Shanghai 200433, China

[2]Department of Economics, The Ohio State University, Columbus, OH 43210, USA

July 23, 2020

Abstract

This paper considers closed-form root estimators for spatial autoregressive models with spatial autoregressive disturbances (SARAR model). We first derive a simple consistent closed-form estimator. Then we construct feasible moment conditions that are quadratic in the spatial lag and spatial error dependence parameters separately, which generate root estimators with closed forms. We consider both the cases with homoskedastic and unknown heteroskedastic disturbances. In the homoskedastic case, the root estimator can be asymptotically as efficient as the quasi-maximum likelihood estimator (QMLE); in the heteroskedastic case, it can be asymptotically as efficient as a method of moments estimator (MME) that sets adjusted quasi-maximum likelihood scores to zero, where the adjusted scores have zero means at the true parameters. The root estimators and their associated standard errors can avoid the computation of any matrix determinants or inverses, so they are computationally simple without iterations, especially valuable for big data where the sample size is large.

Keywords: Spatial autoregression, root estimator, efficiency, heteroskedasticity

JEL classification: C13, C21, R15

1 Introduction

The spatial autoregressive (SAR) model with spatial error (SE) is known as an SARAR model, which is popular in empirical studies. This paper considers efficient root estimators for the SARAR model, which generalizes those in Jin and Lee (2012) of the SAR or SE models. The root estimators have closed-form expressions and are computationally and asymptotically efficient.

For the SARAR model, the quasi maximum likelihood (QML) estimator is consistent, but computationally intensive with large sample sizes. When disturbances are normal, it is the maximum likelihood

*Corresponding author. E-mail addresses: jin.fei@live.com (F. Jin), lee.1777@osu.edu (L.-F. Lee).

estimator (MLE) and thus asymptotically efficient.[1] Kelejian and Prucha (1998) propose a generalized spatial two-sage least squares estimator (GS2SLSE) for the spatial lag parameter and coefficients of regressors, but quadratic moments that capture spatial correlation are not used. The GS2SLSE is less efficient than the QMLE. Lee (2007a) and Liu et al. (2010) propose a generalized method of moments estimator (GMME), which uses both linear and quadratic moments. Based on concentration, Lee (2007b) proposes a GMME that reduces the estimation to only the spatial lag parameter. With properly chosen moments, the GMMEs can be asymptotically as efficient as the MLE for normal disturbances, and asymptotically more efficient than the QMLE for non-normal disturbances. Although relatively simpler in computation, the above GMMEs have no closed forms and searching over a parameter space is needed. Jin and Lee (2012) propose a root estimator for the SAR model, which is computationally simple since it has a closed form.[2]

The approach in Jin and Lee (2012) cannot be applied to the SARAR model, since it generates two nonlinear moment conditions with two unknown parameters, which do not have simple attractive closed-form solutions. However, a sequential GMM approach in Jin and Lee (2020) can be applied and can generate simple closed-form root estimators. We estimate the spatial lag and spatial error dependence parameters separately by using properly estimated quadratic moments. We specify conditions on selecting the consistent root from the two roots of quadratic moments. With homoskedastic disturbances, the root estimator can be asymptotically as efficient as the QMLE; with heteroskedastic disturbances, it is asymptotically efficient as a method of moments estimator (MME) that sets jointly the adjusted QML scores to zero, where the adjusted scores have mean zero at the true parameter values (Liu and Yang, 2015). The root estimator and its standard error can avoid the computation of any matrix inverse. Our computationally simple root estimator for the SARAR model can be useful, in particular with big data.

Section 2 considers a simple initial consistent estimator. Section 3 investigates root estimators. Monte Carlo (MC) results are reported in Section 4. Section 5 concludes. Proofs are in an online supplementary file.

2 Initial consistent estimator

Consider the SARAR model:

$$y_n = \lambda_0 W_n y_n + X_n \beta_0 + u_n, \quad u_n = \rho_0 M_n u_n + \epsilon_n, \tag{2.1}$$

where y_n is an $n \times 1$ vector of observed dependent variables, X_n is an $n \times k_x$ matrix of exogenous variables, W_n and M_n are $n \times n$ zero-diagonal spatial weights matrices, which may or may not be the same, λ_0 and ρ_0 are scalar parameters, β_0 is a $k_x \times 1$ vector of coefficients, and $\epsilon_n = [\epsilon_{n1}, \dots, \epsilon_{nn}]'$ is an $n \times 1$ vector of independent disturbances with zero means and finite variances. Let $\sigma_{ni}^2 = \mathrm{E}(\epsilon_{ni}^2)$ and

[1] The MLE for the SAR model has been considered in Ord (1975). Lee (2004) and Jin and Lee (2013) provide asymptotic properties of the QMLE for, respectively, the SAR and SARAR models.

[2] Ord (1975) introduced a root estimator for the SAR model, but no theoretical analysis was provided.

$\Sigma_n = \mathrm{diag}(\sigma_{n1}^2, \cdots, \sigma_{nn}^2)$ be a diagonal matrix of σ_{ni}^2's. When ϵ_{ni}'s are homoskedastic, $\sigma_{ni}^2 = \sigma_0^2$ and $\Sigma_n = \sigma_0^2 I_n$.

A consistent estimator $\tilde{\eta}$ of $\eta_0 = [\lambda_0, \beta_0']'$ can be a 2SLSE, as in Kelejian and Prucha (1998). Let the IV matrix for $Z_n = [W_n y_n, X_n]$ be Q_n, e.g., a matrix consisting of independent columns of $[X_n, W_n X_n, W_n^2 X_n]$. The 2SLSE with Q_n is $\tilde{\eta} = [Z_n' Q_n (Q_n' Q_n)^{-1} Q_n' Z_n]^{-1} Z_n' Q_n (Q_n' Q_n)^{-1} Q_n' y_n$.

To derive an initial estimator of ρ_0, we consider a root estimator as in Ord (1975) and Jin and Lee (2012). With a zero-diagonal $n \times n$ matrix P_n, $\mathrm{E}(\epsilon_n' P_n \epsilon_n) = \mathrm{tr}(P_n \Sigma_n) = 0$. A zero-diagonal P_n can be M_n or $M_n^2 - \mathrm{diag}(M_n^2)$, where $\mathrm{diag}(A)$ denotes a diagonal matrix formed by the diagonal elements of A. The $\tilde{u}_n = y_n - Z_n \tilde{\eta}$ is an estimate of u_n. Denote $A^{(s)} = A + A'$ for any square matrix A. We consider roots of the quadratic equation:

$$0 = \tilde{u}_n' (I_n - \rho M_n)' P_n (I_n - \rho M_n) \tilde{u}_n = a_n \rho^2 + b_n \rho + c_n, \tag{2.2}$$

where $a_n = \tilde{u}_n' M_n' P_n M_n \tilde{u}_n$, $b_n = -\tilde{u}_n' P_n^{(s)} M_n \tilde{u}_n$, and $c_n = \tilde{u}_n' P_n \tilde{u}_n$. Under regularity conditions, $\frac{1}{n^2} b_n^2 - \frac{4}{n^2} a_n c_n = (\frac{d_n}{n})^2 + o_p(1)$, where $d_n = \tilde{u}_n' (I_n - \rho_0 M_n') P_n^{(s)} M_n \tilde{u}_n$. The consistent root estimator $\tilde{\rho}$ of ρ_0 is:

$$\tilde{\rho} = \frac{-b_n - \sqrt{b_n^2 - 4a_n c_n}}{2a_n} \text{ if } \operatorname*{plim}_{n\to\infty} \frac{1}{n} d_n \geq 0; \text{ and } \tilde{\rho} = \frac{-b_n + \sqrt{b_n^2 - 4a_n c_n}}{2a_n} \text{ if } \operatorname*{plim}_{n\to\infty} \frac{1}{n} d_n < 0. \tag{2.3}$$

Since $\frac{1}{n^2} b_n^2 - \frac{4}{n^2} a_n c_n = (\frac{d_n}{n})^2 + o_p(1)$, $b_n^2 - 4a_n c_n \geq 0$ with probability approaching one if $\mathrm{plim}_{n\to\infty} \frac{1}{n} d_n \neq 0$. In finite samples, if $b_n^2 - 4a_n c_n < 0$, it is proper to set $\tilde{\rho} = -\frac{b_n}{2a_n}$.

As $\frac{1}{n} d_n$ involves ρ_0, one may use a consistently estimated ρ_0 to determine the consistent root. A possible way is to compute the roots for two elected P_n, say P_n^* and P_n^{**}. Let the roots corresponding to P_n^* be $\tilde{\rho}_1^*$ and $\tilde{\rho}_2^*$, and the coefficients of that quadratic equation be a_n^*, b_n^* and c_n^*. Similarly, with P_n^{**}, we have $\tilde{\rho}_1^{**}$, $\tilde{\rho}_2^{**}$, a_n^{**}, b_n^{**} and c_n^{**}. Then we compute $|\tilde{\rho}_1^* - \tilde{\rho}_1^{**}|$, $|\tilde{\rho}_1^* - \tilde{\rho}_2^{**}|$, $|\tilde{\rho}_2^* - \tilde{\rho}_1^{**}|$ and $|\tilde{\rho}_2^* - \tilde{\rho}_2^{**}|$. A consistent root $\tilde{\rho}$ can be either one of the two roots corresponding to the difference with the smallest absolute value. The limits of the four absolute values are 0, $|-\mathrm{plim}_{n\to\infty} \frac{b_n^*}{a_n^*} - 2\rho_0|$, $|-\mathrm{plim}_{n\to\infty} \frac{b_n^{**}}{a_n^{**}} - 2\rho_0|$, and $|\mathrm{plim}_{n\to\infty} \frac{b_n^*}{a_n^*} - \mathrm{plim}_{n\to\infty} \frac{b_n^{**}}{a_n^{**}}|$, where the first one corresponds to two consistent roots. This method can locate a consistent root as long as $\mathrm{plim}_{n\to\infty} \frac{b_n^*}{a_n^*} \neq \mathrm{plim}_{n\to\infty} \frac{b_n^{**}}{a_n^{**}}$, i.e, $\lim_{n\to\infty} \frac{\mathrm{tr}(R_n'^{-1} P_n^{*(s)} M_n R_n^{-1} \Sigma_n)}{\mathrm{tr}(R_n'^{-1} M_n' P_n^* M_n R_n^{-1} \Sigma_n)} \neq \lim_{n\to\infty} \frac{\mathrm{tr}(R_n'^{-1} P_n^{**(s)} M_n R_n^{-1} \Sigma_n)}{\mathrm{tr}(R_n'^{-1} M_n' P_n^{**} M_n R_n^{-1} \Sigma_n)}$, which would hold generally. The P_n^* and P_n^{**} can be, e.g., $M_n + \kappa M_n^2 + \kappa^2 M_n^3 - \mathrm{diag}(M_n + \kappa M_n^2 + \kappa^2 M_n^3)$ for $\kappa = 0.2$ or 0.6 in our experience.

We maintain two assumptions for the $\sqrt{n}$-consistency of $\tilde{\eta}$ and $\tilde{\rho}$.

Assumption 1. *(a) Either (i) ϵ_{ni}'s are i.i.d. $(0, \sigma_0^2)$ and $\mathrm{E}(|\epsilon_{ni}|^{4+\iota}) < \infty$ for some $\iota > 0$, or (ii) ϵ_{ni}'s are independent with mean zero and variances σ_{ni}^2's, and $\sup_n \sup_{1\leq i\leq n} \mathrm{E}(|\epsilon_{ni}|^{4+\iota}) < \infty$ for some $\iota > 0$. (b) W_n and M_n have zero diagonals, S_n and R_n are invertible, and $\{W_n\}$, $\{M_n\}$, $\{S_n^{-1}\}$ and $\{R_n^{-1}\}$ are bounded in both row and column sum norms. (c) Elements of X_n are uniformly bounded constants, and $\lim_{n\to\infty} \frac{1}{n} X_n' X_n$ is nonsingular.*

Assumption 2. *(a) Elements of Q_n are uniformly bounded constants, $\lim_{n\to\infty} \frac{1}{n} Q_n' Q_n$ and $\lim_{n\to\infty} \frac{1}{n} Q_n' Z_n$ have full column ranks. (b) $\{P_n^*\}$ and $\{P_n^{**}\}$ are bounded in both row and column sum norms. (c) For $P_n =$*

P_n^* or P_n^{**}, $\lim_{n\to\infty} \frac{1}{n} \operatorname{tr}(R_n'^{-1} M_n' P_n M_n R_n^{-1} \Sigma_n) \neq 0$. (d) $\lim_{n\to\infty} \frac{\operatorname{tr}(R_n'^{-1} P_n^{*(s)} M_n R_n^{-1} \Sigma_n)}{\operatorname{tr}(R_n'^{-1} M_n' P_n^* M_n R_n^{-1} \Sigma_n)} \neq \lim_{n\to\infty} \frac{\operatorname{tr}(R_n'^{-1} P_n^{**(s)} M_n R_n^{-1} \Sigma_n)}{\operatorname{tr}(R_n'^{-1} M_n' P_n^{**} M_n R_n^{-1} \Sigma_n)}$. (e) $\lim_{n\to\infty} \frac{1}{n} \operatorname{tr}(P_n^{(s)} M_n R_n^{-1} \Sigma_n) \neq 0$ for $P_n = P_n^*$ or P_n^{**}.

Assumption 1 contains typical regularity conditions in spatial econometrics in Kelejian and Prucha (1998, 1999, 2001, 2010), Lee (2004) and Lin and Lee (2010). Assumption 2(a) is a familiar condition on 2SLSEs. As P_n^* and P_n^{**} are constructed from M_n, it is reasonable to impose Assumption 2(b). Under Assumption 2(c), the moment equations involving P_n^* and P_n^{**} are quadratic in ρ. Assumption 2(d) guarantees that a consistent root estimator $\tilde{\rho}$ of ρ_0 can be located. Assumption 2(e) is needed for the $\sqrt{n}$-rate of convergence of $\tilde{\rho}$.

Theorem 1. *Under Assumptions 1–2, $\tilde{\eta} = \eta_0 + O_p(n^{-1/2})$ and $\tilde{\rho} = \rho_0 + O_p(n^{-1/2})$.*

3 Root estimator

Our root estimator for $\phi_0 = [\lambda_0, \rho_0, \beta_0']'$ is based on the scores. The quasi log likelihood function of (2.1), as if ϵ_{ni}'s were i.i.d. normal, is

$$\ln L_n(\theta) = -\frac{n}{2} \ln(2\pi\sigma^2) + \ln|S_n(\lambda)| + \ln|R_n(\rho)| - \frac{1}{2\sigma^2} \epsilon_n'(\phi) \epsilon_n(\phi),$$

where $\phi = [\lambda, \rho, \beta']'$, $\theta = [\phi', \sigma^2]'$, $R_n(\rho) = I_n - \rho M_n$, $S_n(\lambda) = I_n - \lambda W_n$, and $\epsilon_n(\phi) = R_n(\rho)[S_n(\lambda) y_n - X_n \beta]$. The first order derivatives of $\ln L_n(\theta)$ are

$$\frac{\partial \ln L_n(\theta)}{\partial \lambda} = -\operatorname{tr}[W_n S_n^{-1}(\lambda)] + \frac{1}{\sigma^2} \epsilon_n'(\phi) R_n(\rho) W_n y_n, \tag{3.1}$$

$$\frac{\partial \ln L_n(\theta)}{\partial \rho} = -\operatorname{tr}[M_n R_n^{-1}(\rho)] + \frac{1}{\sigma^2} \epsilon_n'(\phi) M_n [S_n(\lambda) y_n - X_n \beta], \tag{3.2}$$

$$\frac{\partial \ln L_n(\theta)}{\partial \beta} = \frac{1}{\sigma^2} X_n' R_n'(\rho) \epsilon_n(\phi), \tag{3.3}$$

$$\frac{\partial \ln L_n(\theta)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \epsilon_n'(\phi) \epsilon_n(\phi). \tag{3.4}$$

By (3.3), for given $\gamma = [\lambda, \rho]'$, the QMLE of β is

$$\hat{\beta}(\gamma) = [X_n' R_n'(\rho) R_n(\rho) X_n]^{-1} X_n' R_n'(\rho) R_n(\rho) S_n(\lambda) y_n; \tag{3.5}$$

by (3.4), for given ϕ, the QMLE of σ^2 is $\hat{\sigma}^2(\phi) = \frac{1}{n} \epsilon_n'(\phi) \epsilon_n(\phi)$. We shall derive a root estimator $\hat{\gamma}$ of $\gamma_0 = [\lambda_0, \rho_0]'$, then evaluate $\hat{\beta}(\gamma)$ at $\hat{\gamma}$, and finally evaluate $\hat{\sigma}^2(\phi)$.

Substituting $\hat{\sigma}^2(\phi)$ into (3.1)–(3.2), we derive the nonlinear moment vector

$$m_n(\phi) = [m_{1n}(\phi), m_{2n}(\phi), m_{3n}'(\phi)]', \tag{3.6}$$

where $m_{1n}(\phi) = \epsilon_n'(\phi) R_n(\rho) W_n y_n - \frac{1}{n} \epsilon_n'(\phi) \epsilon_n(\phi) \operatorname{tr}[W_n S_n^{-1}(\lambda)]$, $m_{2n}(\phi) = \epsilon_n'(\phi) M_n [S_n(\lambda) y_n - X_n \beta] - \frac{1}{n} \epsilon_n'(\phi) \epsilon_n(\phi) \operatorname{tr}[M_n R_n^{-1}(\rho)]$, and $m_{3n}(\phi) = X_n' R_n'(\rho) \epsilon_n(\phi)$. To derive a simple root estimator of ρ, we investigate $m_n(\phi)$ at ϕ_0, which suggests a modified moment vector. Since $y_n = S_n^{-1} X_n \beta_0 + S_n^{-1} R_n^{-1} \epsilon_n$, $m_n(\phi_0) = [\epsilon_n' G_n \epsilon_n + \epsilon_n' R_n W_n S_n^{-1} X_n \beta_0, \epsilon_n' T_n \epsilon_n, \epsilon_n' R_n X_n]'$, where $G_n = R_n W_n S_n^{-1} R_n^{-1} - \frac{1}{n} \operatorname{tr}(W_n S_n^{-1}) I_n$

and $T_n = M_n R_n^{-1} - \frac{1}{n}\operatorname{tr}(M_n R_n^{-1}) I_n$. Let $\tilde{\phi} = [\tilde{\lambda}, \tilde{\rho}, \tilde{\beta}']'$, $\tilde{R}_n = R_n(\tilde{\rho})$ and $\tilde{S}_n = S_n(\tilde{\lambda})$. Denote $\tilde{G}_{nd} = \tilde{R}_n W_n \tilde{S}_n^{-1} \tilde{R}_n^{-1} - \frac{1}{n}\operatorname{tr}(W_n \tilde{S}_n^{-1}) I_n$ and $\tilde{T}_{nd} = M_n \tilde{R}_n^{-1} - \frac{1}{n}\operatorname{tr}(M_n \tilde{R}_n^{-1}) I_n$ if ϵ_{ni}'s are homoskedastic; and $\tilde{G}_{nd} = \tilde{R}_n W_n \tilde{S}_n^{-1} \tilde{R}_n^{-1} - \operatorname{diag}(\tilde{R}_n W_n \tilde{S}_n^{-1} \tilde{R}_n^{-1})$ and $\tilde{T}_{nd} = M_n \tilde{R}_n^{-1} - \operatorname{diag}(M_n \tilde{R}_n^{-1})$ if ϵ_{ni}'s are heteroskedastic. We consider the modified moment vector:

$$\tilde{m}_n(\phi) = [\epsilon_n'(\phi)\tilde{G}_{nd}\epsilon_n(\phi) + \epsilon_n'(\phi)\tilde{R}_n W_n \tilde{S}_n^{-1} X_n \tilde{\beta},\ \epsilon_n'(\phi)\tilde{T}_{nd}\epsilon_n(\phi),\ \epsilon_n'(\phi)\tilde{R}_n X_n]'. \tag{3.7}$$

The $\operatorname{plim}_{n\to\infty} \frac{1}{n}\tilde{m}_n(\phi_0)$ is zero, because $\tilde{G}_{nd}$ and $\tilde{T}_{nd}$ have zero traces for the homoskedastic case, and zero diagonals for the heteroskedastic case. Since $\tilde{G}_{nd}$ and $\tilde{T}_{nd}$ play the role of quadratic matrices for moments in $\epsilon_n(\phi)$, the initial estimates $\tilde{\lambda}$ and $\tilde{\rho}$ in $\tilde{G}_{nd}$ and $\tilde{T}_{nd}$ will not affect the asymptotic distribution of an MME derived from $\tilde{m}_n(\phi)$. Neither will the estimated IVs $\tilde{R}_n W_n \tilde{S}_n^{-1} X_n \tilde{\beta}$ and $\tilde{R}_n X_n$. For a very large n, e.g., $n = 10^6$, computing $\tilde{S}_n^{-1}$ and $\tilde{T}_n^{-1}$ even once can be demanding. We suggest to use $\sum_{i=0}^{k} \tilde{\lambda} W_n^i$ and $\sum_{i=0}^{k} \tilde{\rho} M_n^i$ for some natural number k to approximate, respectively, $\tilde{S}_n^{-1}$ and $\tilde{T}_n^{-1}$. In general, we consider root estimators based on the moment vector:

$$g_n(\phi) = [\epsilon_n'(\phi)\tilde{G}_{nd,k}\epsilon_n(\phi) + \epsilon_n'(\phi)\tilde{R}_n W_n \tilde{S}_n^{-1} X_n \tilde{\beta},\ \epsilon_n'(\phi)\tilde{T}_{nd,k}\epsilon_n(\phi),\ \epsilon_n'(\phi)\tilde{R}_n X_n]', \tag{3.8}$$

where $\tilde{G}_{nd,k}$ and $\tilde{T}_{nd,k}$ are derived by replacing the involved $\tilde{S}_n^{-1}$ and $\tilde{T}_n^{-1}$ in $\tilde{G}_{nd}$ and $\tilde{T}_{nd}$ with, respectively, $\sum_{i=0}^{k} \tilde{\lambda} W_n^i$ and $\sum_{i=0}^{k} \tilde{\rho} M_n^i$.[3] When $k = \infty$ or is sufficiently large, $g_n(\phi) = \tilde{m}_n(\phi)$.

A root estimator of ρ_0

Let $g_{1n}(\rho, \eta) = \epsilon_n'(\phi)\tilde{T}_{nd,k}\epsilon_n(\phi)$ and $g_{2n}(\rho, \eta) = [\epsilon_n'(\phi)\tilde{G}_{nd,k}\epsilon_n(\phi) + \epsilon_n'(\phi)\tilde{R}_n W_n \tilde{S}_n^{-1} X_n \tilde{\beta},\ \epsilon_n'(\phi)\tilde{R}_n X_n]'$, where $\eta = [\lambda, \beta']'$. The moments $g_{1n}(\rho, \eta)$ and $g_{2n}(\rho, \eta)$ are subvectors composing $g_n(\phi)$, which are the derivatives of the quasi log likelihood function with respect to ρ and η. To focus on the estimation of ρ_0, we consider roots of the $C(\alpha)$-type moment:

$$\mathcal{G}_n(\rho, \tilde{\eta}) = g_{1n}(\rho, \tilde{\eta}) - \tilde{C}_{n\rho} g_{2n}(\rho, \tilde{\eta}) = a_{1n}\rho^2 + b_{1n}\rho + c_{1n} = 0, \tag{3.9}$$

where $\tilde{C}_{n\rho} = \frac{\partial g_{1n}(\tilde{\rho}, \tilde{\eta})}{\partial \eta'} (\frac{\partial g_{2n}(\tilde{\rho}, \tilde{\eta})}{\partial \eta'})^{-1}$, $a_{1n} = \tilde{u}_n' M_n' \tilde{T}_{nd,k} M_n \tilde{u}_n - \tilde{C}_{n\rho} \begin{pmatrix} \tilde{u}_n' M_n' \tilde{G}_{nd,k} M_n \tilde{u}_n \\ 0 \end{pmatrix}$ with $\tilde{u}_n = \tilde{S}_n y_n - X_n \tilde{\beta}$,

$$b_{1n} = -\tilde{u}_n' M_n' \tilde{T}_{nd,k}^{(s)} \tilde{u}_n + \tilde{C}_{n\rho} \begin{pmatrix} \tilde{u}_n' M_n' \tilde{G}_{nd,k}^{(s)} \tilde{u}_n + \tilde{u}_n' M_n' \tilde{R}_n W_n \tilde{S}_n^{-1} X_n \tilde{\beta} \\ X_n' \tilde{R}_n' M_n \tilde{u}_n \end{pmatrix},$$

and $c_{1n} = \tilde{u}_n' \tilde{T}_{nd,k} \tilde{u}_n - \tilde{C}_{n\rho} \begin{pmatrix} \tilde{u}_n' \tilde{G}_{nd,k} \tilde{u}_n + \tilde{u}_n' \tilde{R}_n W_n \tilde{S}_n^{-1} X_n \tilde{\beta} \\ X_n' \tilde{R}_n' \tilde{u}_n \end{pmatrix}$. The moment $\mathcal{G}_n(\rho, \tilde{\eta})$ eliminates possible asymptotic impact of $\tilde{\eta}$ on the GMM estimator of ρ_0 from (3.9), since the derivative of $\mathcal{G}_n(\rho, \eta)/n$ with respect to η at (ρ_0, η_0) converges to zero in probability. The consistent root estimator of ρ_0 derived from (3.9) will be asymptotically as efficient as the (joint) GMM estimator of ρ_0 derived from solving $[g_{1n}(\rho, \eta), g_{2n}'(\rho, \eta)]' = 0$, since the number of moments in $g_{2n}(\rho, \eta)$ is equal to the number of parameters in η (Jin and Lee, 2020). By an analysis similar to that in the last section, the consistent root estimator

[3] For $\tilde{S}_n^{-1}$ in $\tilde{R}_n W_n \tilde{S}_n^{-1} X_n \tilde{\beta}$, we do not replace it with $\sum_{i=0}^{k} \tilde{\lambda} W_n^k$, since the vector $\tilde{S}_n^{-1} X_n \tilde{\beta}$ can be efficiently computed using Gaussian elimination.

$\hat{\rho}$ of ρ_0 is

$$\hat{\rho} = \frac{-b_{1n} - \sqrt{b_{1n}^2 - 4a_{1n}c_{1n}}}{2a_{1n}} \text{ if } \operatorname*{plim}_{n\to\infty} \frac{1}{n} d_{1n} \geq 0; \text{ and } \hat{\rho} = \frac{-b_{1n} + \sqrt{b_{1n}^2 - 4a_{1n}c_{1n}}}{2a_{1n}} \text{ if } \operatorname*{plim}_{n\to\infty} \frac{1}{n} d_{1n} < 0, \tag{3.10}$$

where $d_{1n} = \tilde{u}_n' M_n' \tilde{T}_{nd,k}^{(s)} \tilde{\epsilon}_n - \tilde{C}_{n\rho} \begin{pmatrix} \tilde{u}_n' M_n' \tilde{G}_{nd,k}^{(s)} \tilde{\epsilon}_n \\ 0 \end{pmatrix}$ with $\tilde{\epsilon}_n = \tilde{R}_n \tilde{u}_n$.

A root estimator of λ_0

Denote $\tau = [\rho, \beta']'$. Let $g_{1n}(\lambda, \tau) = \epsilon_n'(\phi) \tilde{G}_{nd,k} \epsilon_n(\phi) + \epsilon_n'(\phi) \tilde{R}_n W_n \tilde{S}_n^{-1} X_n \tilde{\beta}$ and $g_{2n}(\lambda, \tau) = [\epsilon_n'(\phi) \tilde{T}_{nd,k} \epsilon_n(\phi), \epsilon_n'(\phi) \tilde{R}_n X_n]'$ be components of $g_n(\phi)$ in (3.8). We consider roots of the $C(\alpha)$-type moment

$$g_{1n}(\lambda, \tilde{\tau}) - \tilde{C}_{n\lambda} g_{2n}(\lambda, \tilde{\tau}) = a_{2n}\lambda^2 + b_{2n}\lambda + c_{2n} = 0, \tag{3.11}$$

where $\tilde{C}_{n\lambda} = \frac{\partial g_{1n}(\tilde{\lambda},\tilde{\tau})}{\partial \tau'} (\frac{\partial g_{2n}(\tilde{\lambda},\tilde{\tau})}{\partial \tau'})^{-1}$, $a_{2n} = (W_n y_n)' \tilde{R}_n' \tilde{G}_{nd,k} \tilde{R}_n W_n y_n - \tilde{C}_{n\lambda} \begin{pmatrix} (W_n y_n)' \tilde{R}_n' \tilde{T}_{nd,k} \tilde{R}_n W_n y_n \\ 0 \end{pmatrix}$, $b_{2n} = -(W_n y_n)' \tilde{R}_n' \tilde{G}_{nd,k}^{(s)} \tilde{R}_n (y_n - X_n \tilde{\beta}) - (W_n y_n)' \tilde{R}_n' \tilde{R}_n W_n \tilde{S}_n^{-1} X_n \tilde{\beta} + \tilde{C}_{n\lambda} \begin{pmatrix} (W_n y_n)' \tilde{R}_n' \tilde{T}_{nd,k}^{(s)} \tilde{R}_n (y_n - X_n \tilde{\beta}) \\ X_n' \tilde{R}_n' \tilde{R}_n W_n y_n \end{pmatrix}$, and $c_{2n} = (y_n - X_n \tilde{\beta})' \tilde{R}_n' \tilde{G}_{nd,k} \tilde{R}_n (y_n - X_n \tilde{\beta}) + (y_n - X_n \tilde{\beta})' \tilde{R}_n' \tilde{R}_n W_n \tilde{S}_n^{-1} X_n \tilde{\beta} - \tilde{C}_{n\lambda} \begin{pmatrix} (y_n - X_n \tilde{\beta})' \tilde{R}_n' \tilde{T}_{nd,k} \tilde{R}_n (y_n - X_n \tilde{\beta}) \\ X_n' \tilde{R}_n' \tilde{R}_n (y_n - X_n \tilde{\beta}) \end{pmatrix}$. The consistent root estimator $\hat{\lambda}$ of λ_0 is

$$\hat{\lambda} = \frac{-b_{2n} - \sqrt{b_{2n}^2 - 4a_{2n}c_{2n}}}{2a_{2n}} \text{ if } \operatorname*{plim}_{n\to\infty} \frac{1}{n} d_{2n} \geq 0; \text{ and } \hat{\lambda} = \frac{-b_{2n} + \sqrt{b_{2n}^2 - 4a_{2n}c_{2n}}}{2a_{2n}} \text{ if } \operatorname*{plim}_{n\to\infty} \frac{1}{n} d_{2n} < 0, \tag{3.12}$$

where $d_{2n} = (W_n y_n)' \tilde{R}_n' \tilde{G}_{nd,k}^{(s)} \tilde{\epsilon}_n + (W_n y_n)' \tilde{R}_n' \tilde{R}_n W_n \tilde{S}_n^{-1} X_n \tilde{\beta} - \tilde{C}_{n\lambda} \begin{pmatrix} (W_n y_n)' \tilde{R}_n' \tilde{T}_{nd,k}^{(s)} \tilde{\epsilon}_n \\ X_n' \tilde{R}_n' \tilde{R}_n W_n y_n \end{pmatrix}$. With the root estimator $\hat{\gamma} = [\hat{\lambda}, \hat{\rho}]'$, an estimator of β_0 can be $\hat{\beta} = \hat{\beta}(\hat{\gamma})$ in (3.5).

We provide regularity conditions for the consistency and asymptotic normality of $\hat{\phi} = [\hat{\lambda}, \hat{\rho}, \hat{\beta}']'$ below. Replacing the estimated parameters in $\tilde{G}_{nd,k}$ and $\tilde{T}_{nd,k}$ by their true values yields the matrices $G_{nd,k}$ and $T_{nd,k}$. Denote $D_n = R_n W_n S_n^{-1} R_n^{-1}$, $H_n = I_n - R_n X_n (X_n' R_n' R_n X_n)^{-1} X_n' R_n'$, $\Xi_n = \operatorname{tr}(G_{nd,k}^{(s)} G_{nd} \Sigma_n) + (R_n W_n S_n^{-1} X_n \beta_0)' H_n (R_n W_n S_n^{-1} X_n \beta_0)$, $\Psi_n = \Xi_n - \frac{\operatorname{tr}(G_{nd,k}^{(s)} T_{nd} \Sigma_n) \operatorname{tr}(T_{nd,k}^{(s)} G_{nd} \Sigma_n)}{\operatorname{tr}(T_{nd,k}^{(s)} T_{nd} \Sigma_n)}$,

$$a_{1n}^* = \operatorname{tr}(R_n'^{-1} M_n' T_{nd,k} M_n R_n^{-1} \Sigma_n) - \operatorname{tr}(T_{nd,k}^{(s)} G_{nd} \Sigma_n) \operatorname{tr}(R_n'^{-1} M_n' G_{nd,k} M_n R_n^{-1} \Sigma_n) \Xi_n^{-1}, \tag{3.13}$$

$$\begin{aligned} a_{2n}^* &= (R_n W_n S_n^{-1} X_n \beta_0)' G_{nd,k} (R_n W_n S_n^{-1} X_n \beta_0) + \operatorname{tr}(D_n' G_{nd,k} D_n \Sigma_n) \\ &\quad - \operatorname{tr}(G_{nd,k}^{(s)} T_{nd} \Sigma_n) \operatorname{tr}^{-1}(T_{nd,k}^{(s)} T_{nd} \Sigma_n) [(R_n W_n S_n^{-1} X_n \beta_0)' T_{nd,k} (R_n W_n S_n^{-1} X_n \beta_0) + \operatorname{tr}(D_n' T_{nd,k} D_n \Sigma_n)]. \end{aligned} \tag{3.14}$$

Assumption 3. *(a)* $\|\lambda_0 W_n\| < 1$ *and* $\|\rho_0 M_n\| < 1$. *(b)* $\lim_{n\to\infty} \frac{1}{n} X_n' R_n' R_n X_n$ *is nonsingular,* $\lim_{n\to\infty} \frac{1}{n} \Xi_n \neq 0$ *and* $\lim_{n\to\infty} \frac{1}{n} \operatorname{tr}(T_{nd,k}^{(s)} T_{nd} \Sigma_n) \neq 0$. *(c)* $\lim_{n\to\infty} \frac{1}{n} a_{1n}^* \neq 0$ *and* $\lim_{n\to\infty} \frac{1}{n} a_{2n}^* \neq 0$. *(d)* $\lim_{n\to\infty} \frac{1}{n} \Psi_n \neq 0$.

Under Assumption 3(a), $G_{nd,k} = G_{nd}$ and $T_{nd,k} = T_{nd}$ when $k = \infty$. Assumption 3(b) guarantees the existence of $\tilde{C}_{n\rho}$ and $\tilde{C}_{n\lambda}$ for large enough n. When ϵ_{ni}'s are homoskedastic and $k = \infty$, $\operatorname{tr}(T_{nd,k}^{(s)} T_{nd} \Sigma_n) = \frac{\sigma_0^2}{2} \operatorname{tr}(T_{nd}^{(s)} T_{nd}^{(s)}) \geq 0$ and $\operatorname{tr}(G_{nd,k}^{(s)} G_{nd} \Sigma_n) \geq 0$. As H_n is a projection matrix, $(R_n W_n S_n^{-1} X_n \beta_0)' H_n (R_n W_n S_n^{-1} X_n \beta_0) \geq 0$. Under Assumption 3$(c)$, the moment equations (3.9) and (3.11) are quadratic in their unknown parameters. Assumption 3(d) guarantees the nonsingularity of the gradient of $g_n(\phi)$ at ϕ_0 for the asymptotic distribution of the root estimator. When $k=\infty$ and ϵ_{ni}'s are homoskedastic, $\operatorname{tr}(G^{(s)}_{nd,k}G_{nd}\Sigma_n)-\frac{\operatorname{tr}(G^{(s)}_{nd,k}T_{nd}\Sigma_n)\operatorname{tr}(T^{(s)}_{nd,k}G_{nd}\Sigma_n)}{\operatorname{tr}(T^{(s)}_{nd,k}T_{nd}\Sigma_n)}$ in Ψ_n is non-negative by the Cauchy-Schwarz inequality.

We compare our root estimator of ϕ_0 with the QMLE in the homoskedastic case, and with the MME derived by solving $h_n(\phi)=0$ in the heteroskedastic case, where

$$\begin{aligned} h_n(\phi) = [&\epsilon_n'(\phi)R_n(\rho)W_ny_n - \epsilon_n'(\phi)\operatorname{diag}(R_n(\rho)W_nS_n^{-1}(\lambda)R_n^{-1}(\rho))\epsilon_n(\phi), \\ &\epsilon_n'(\phi)M_n(S_n(\lambda)y_n - X_n\beta) - \epsilon_n'(\phi)\operatorname{diag}(M_nR_n^{-1}(\rho))\epsilon_n(\phi),\, \epsilon_n(\phi)R_n(\rho)X_n]', \end{aligned}$$

with $\mathrm{E}[h_n(\phi_0)]=0$. When ϵ_{ni}'s are homoskedastic, let $\mu_3=\mathrm{E}(\epsilon_{ni}^3)$, $\mu_4=\mathrm{E}(\epsilon_{ni}^4)$, and ϕ^* be the QMLE of ϕ_0; when ϵ_{ni}'s are heteroskedastic, let ϕ^* be the MME derived from solving $h_n(\phi)=0$.

Theorem 2. *Under Assumptions 1–3,* $\sqrt{n}(\hat{\phi}-\phi_0)\xrightarrow{d} N\big(0,\lim_{n\to\infty}\Gamma^{-1}_{nd,k}(\Omega_{nd,k}+\Delta_{nd,k})\Gamma'^{-1}_{nd,k}\big)$, *where*

$$\Gamma_{nd,k}=\frac{1}{n}\begin{pmatrix} \operatorname{tr}(G^{(s)}_{nd,k}G_{nd}\Sigma_n)+(R_nW_nS_n^{-1}X_n\beta_0)'(R_nW_nS_n^{-1}X_n\beta_0) & \operatorname{tr}(G^{(s)}_{nd,k}T_{nd}\Sigma_n) & (R_nW_nS_n^{-1}X_n\beta_0)'R_nX_n \\ \operatorname{tr}(T^{(s)}_{nd,k}G_{nd}\Sigma_n) & \operatorname{tr}(T^{(s)}_{nd,k}T_{nd}\Sigma_n) & 0 \\ X_n'R_n'R_nW_nS_n^{-1}X_n\beta_0 & 0 & X_n'R_n'R_nX_n \end{pmatrix},$$

$$\Omega_{nd,k}=\frac{1}{n}\begin{pmatrix} \Phi_n & \operatorname{tr}(G^{(s)}_{nd,k}\Sigma_nT_{nd,k}\Sigma_n) & (R_nW_nS_n^{-1}X_n\beta_0)'\Sigma_nR_nX_n \\ \operatorname{tr}(T^{(s)}_{nd,k}\Sigma_nG_{nd,k}\Sigma_n) & \operatorname{tr}(T^{(s)}_{nd,k}\Sigma_nT_{nd,k}\Sigma_n) & 0 \\ X_n'R_n'\Sigma_nR_nW_nS_n^{-1}X_n\beta_0 & 0 & X_n'R_n'\Sigma_nR_nX_n \end{pmatrix},$$

with $\Phi_n=\operatorname{tr}(G^{(s)}_{nd,k}\Sigma_nG_{nd,k}\Sigma_n)+(R_nW_nS_n^{-1}X_n\beta_0)'\Sigma_n(R_nW_nS_n^{-1}X_n\beta_0)$, $\Delta_{nd,k}=0$ *in the heteroskedastic case, but in the homoskedastic case,*

$$\Delta_{nd,k}=\frac{1}{n}\begin{pmatrix} (\mu_4-3\sigma_0^4)\sum_{i=1}^n g^2_{ndk,ii}+\mu_3(R_nW_nS_n^{-1}X_n\beta_0)'\operatorname{vec}_{\mathrm{D}}(G_{nd,k}) & * & * \\ (\mu_4-3\sigma_0^4)\sum_{i=1}^n g_{ndk,ii}t_{ndk,ii}+\mu_3(R_nW_nS_n^{-1}X_n\beta_0)'\operatorname{vec}_{\mathrm{D}}(T_{nd,k}) & (\mu_4-3\sigma_0^4)\sum_{i=1}^n t^2_{ndk,ii} & * \\ \mu_3X_n'R_n'\operatorname{vec}_{\mathrm{D}}(G_{nd,k}) & \mu_3X_n'R_n'\operatorname{vec}_{\mathrm{D}}(T_{nd,k}) & 0 \end{pmatrix}$$

where $G_{nd,k}=[g_{ndk,ij}]$, $T_{nd,k}=[t_{ndk,ij}]$, *and* $\operatorname{vec}_{\mathrm{D}}(A)$ *is the column vector formed by the diagonal elements of a square matrix A. When* $k=\infty$ *and Assumption S.1 for the QMLE and MME in the supplementary file also holds,* $\sqrt{n}(\hat{\phi}-\phi_0)=\sqrt{n}(\phi^*-\phi_0)+o_p(1)$.

This theorem shows that the root estimator $\hat{\phi}$ with $k=\infty$ is asymptotically as efficient as the QMLE of ϕ_0 in the homoskedastic case, and as the MME ϕ^* from solving $h_n(\phi)=0$ in the heteroskedastic case. As $\Gamma_{nd,k}$ involves R_n^{-1} and S_n^{-1}, when the sample size is large, it is computationally demanding to use $\Gamma_{nd,k}$ for inference purposes. Since $\lim_{n\to\infty}\Gamma_{nd,k}$ is the probability limit of $-\frac{1}{n}\frac{\partial g_n(\phi_0)}{\partial\phi'}$, $\Gamma_{nd,k}$ can be estimated by $-\frac{1}{n}\frac{\partial g_n(\hat{\phi})}{\partial\phi'}$. In the homoskedastic case, if ϵ_{ni}'s are normal, $\Delta_{nd,k}=0$ as $\mu_3=0$ and $\mu_4-3\sigma_0^4=0$. In the heteroskedastic case, $\Gamma_{nd,k}$ and $\Omega_{nd,k}$ can be estimated by replacing Σ_n in them with $\operatorname{diag}(\hat{\epsilon}^2_{n1},\ldots,\hat{\epsilon}^2_{nn})$ as in White (1980), and replacing other true parameters with their root estimates.

4 Monte Carlo

We implement some MC experiments on our root estimators. The settings are described in the supplementary file. Our root estimators have similar performance to the QMLE and MME in terms of biases

Table 1: Computational time of estimates

	Homoskedastic case						Heteroskedastic case					
	$\lambda_0 = 0.2,\ \rho_0 = 0.2$			$\lambda_0 = 0.5,\ \rho_0 = 0.5$			$\lambda_0 = 0.2,\ \rho_0 = 0.2$			$\lambda_0 = 0.5,\ \rho_0 = 0.5$		
	MLE	RE	RE5	MLE	RE	RE5	MME	RE	RE5	MME	RE	RE5
$n = 10^4$	1.4(22.7)	20.1	0.2	2.3(37.9)	25.6	0.2	1558.3	20.7	0.2	4497.4	24.1	0.2
$n = 10^5$	41.2	-	2.7	64.6	-	2.4	-	-	2.4	-	-	2.4
$n = 10^6$	455.7	-	25.1	772.5	-	25.3	-	-	25.0	-	-	25.0

The table reports the average time in seconds for computing each estimate once. Each estimate has been computed 5 times. For the QMLE with $n = 10^4$, numbers in parentheses are the time for computing both the QMLE and its standard error. The results are from Matlab 2019b on a desktop computer with an Intel Core (TM) I7-8700 CPU and 16 gigabyte memory.

and dispersions, as seen from the supplementary file. Here we report the computational time of estimates in Table 1. RE denotes the root estimator with the quadratic matrices $\tilde{G}_{nd}$ and $\tilde{T}_{nd}$, and RE5 denotes the root estimator with $\tilde{G}_{nd,5}$ and $\tilde{T}_{nd,5}$. Although the inverses $\tilde{S}_n^{-1}$ and $\tilde{R}_n^{-1}$ only need to be computed once for RE, it takes most time. As the moment function of MME involves matrix inverses with unknown parameters, MME takes much longer to compute than MLE, RE and RE5. With $n \geq 10^5$, matrix inverses would not be computable on a desktop computer with 16 gigabyte memory. By using algorithms for sparse matrices in Matlab, the MLE can be computed relatively efficiently without using derivatives that involve matrix inverses. But the standard error of the MLE involves matrix inverses, which need to be computed. RE5 can be computed very fast. With $n = 10^6$, RE5 takes about 25 seconds to compute, while the MLE takes more than 7 minutes and the standard error of MLE would not be computable on a desktop computer.

5 Conclusion

This paper proposes simple closed-form root estimators for the SARAR model in both the homoskedastic and unknown heteroskedastic cases. We derive a simple initial consistent closed-form estimator and investigate closed-form root estimators. Our root estimator can be asymptotically as efficient as the QMLE in the homoskedastic case. Our Monte Carlo results show that our root estimators perform well in finite samples. They are computationally much faster than the QMLE and MME.

Acknowledgement. Fei Jin gratefully acknowledges the financial support from the National Natural Science Foundation of China (No. 71973030).

References

Jin, F., Lee, L.F., 2012. Approximated likelihood and root estimators for spatial interaction in spatial autoregressive models. Regional Science and Urban Economics 42, 446–458.

Jin, F., Lee, L.F., 2013. Cox-type tests for competing spatial autoregressive models with spatial autoregressive disturbances. Regional Science and Urban Economics 43, 590–616.

Jin, F., Lee, L.F., 2020. Efficient two-step generalized empirical likelihood estimation and tests with martingale differences. Forthcoming in the *Econometric Theory*.

Kelejian, H.H., Prucha, I.R., 1998. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. Journal of Real Estate Finance and Economics 17, 99–121.

Kelejian, H.H., Prucha, I.R., 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. International Economic Review 40, 509–533.

Kelejian, H.H., Prucha, I.R., 2001. On the asymptotic distribution of the Moran I test statistic with applications. Journal of Econometrics 104, 219–257.

Kelejian, H.H., Prucha, I.R., 2010. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. Journal of Econometrics 157, 53–67.

Lee, L.F., 2004. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. Econometrica 72, 1899–1925.

Lee, L.F., 2007a. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. Journal of Econometrics 137, 489–514.

Lee, L.F., 2007b. The method of elimination and substitution in the GMM estimation of mixed regressive, spatial autoregressive models. Journal of Econometrics 140, 155–189.

Lin, X., Lee, L.F., 2010. GMM estimation of spatial autoregressive models with unknown heteroskedasticity. Journal of Econometrics 157, 34–52.

Liu, S.F., Yang, Z., 2015. Modified QML estimation of spatial autoregressive models with unknown heteroskedasticity and nonnormality. Regional Science and Urban Economics 52, 50–70.

Liu, X., Lee, L.F., Bollinger, C.R., 2010. An efficient GMM estimator of spatial autoregressive models. Journal of Econometrics 159, 303–319.

Ord, K., 1975. Estimation methods for models of spatial interaction. Journal of the American Statistical Association 70, 120–126.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.