空間共分散構造の乗法型近似による擬似最良推定量について

NEC　中央研究所　　平野　敏弘

空間統計学において線形回帰モデルは基本的な統計手法として広く使用されている．本論文では，誤差項が空間相関を持ち，格子点上でデータが観測される場合の線形回帰モデル

$$y_{\boldsymbol{t}} = X_{\boldsymbol{t}}'\boldsymbol{\beta} + \epsilon_{\boldsymbol{t}},$$

について考えている．ここで，$X_{\boldsymbol{t}} = (x_{\boldsymbol{t},1}, \ldots, x_{\boldsymbol{t},p})'$ は非確率的な説明変数の p 次元ベクトル，$(y_{\boldsymbol{t}}, X_{\boldsymbol{t}})$ は $P_N = \{\boldsymbol{t} = (t_1, t_2)' \in \mathbb{Z}^2 | 1 \le t_1 \le N, 1 \le t_2 \le N\}$ という格子点上で観測されているとする．また，誤差項 $\{\epsilon_{\boldsymbol{t}}\}$ は平均 0，スペクトル密度関数 $f(\boldsymbol{\lambda})$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)' \in [-\pi, \pi]^2$ を持つ定常確率場であると仮定する．誤差項の共分散構造が既知ならば最良線形不偏推定量である一般化最小二乗推定量

$$\hat{\boldsymbol{\beta}}_{GLSE} = \left(X'\Sigma^{-1}X\right)^{-1} X'\Sigma^{-1}\boldsymbol{y},$$

を使用することによって，線形回帰モデルの係数 $\boldsymbol{\beta}$ を推定することができる．ここで，$\Sigma = E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}']$ である．また，$k = (l-1)N + m$ $(l, m = 1, \ldots, N)$ としたとき，$\boldsymbol{y}$ と $\boldsymbol{\epsilon}$ の k 番目の要素はそれぞれ $y_{(l,m)}$，$\epsilon_{(l,m)}$ であり，X の第 k 行は $X'_{(l,m)}$ となる．しかし，大規模空間データに対する一般化最小二乗推定量は誤差項の共分散行列の逆行列 Σ^{-1} を含んでいるので，サンプル数が大きい時その計算時間は極めて大きくなる．

そこで，高速計算可能な代替推定量として，一般化最小二乗推定量の共分散行列における真の共分散関数を自己回帰過程の共分散関数の積で置換した推定量

$$\hat{\boldsymbol{\beta}}_{PBE} = \left(X'\tilde{\Sigma}^{-1}X\right)^{-1} X'\tilde{\Sigma}^{-1}\boldsymbol{y},$$

を提案する．ここで，$\tilde{\Sigma} = \tilde{\Sigma}_1 \bigotimes \tilde{\Sigma}_2$, $\tilde{\Sigma}_i$ は $\mathrm{AR}(P_i)$ $(i = 1, 2)$ の共分散行列である．このとき，クロネッカー積の性質から $\tilde{\Sigma}^{-1} = \left(\tilde{\Sigma}_1 \bigotimes \tilde{\Sigma}_2\right)^{-1} = \tilde{\Sigma}_1^{-1} \bigotimes \tilde{\Sigma}_2^{-1}$ が成立し，$\tilde{\Sigma}_i$ $(i = 1, 2)$ の逆行列の形はよく知られているので高速計算が可能となる．このような推定量は擬似最良推定量と呼ばれており，時系列の文脈においていくつかの先行研究がある．本発表はこの擬似最良推定量の漸近分散の導出を行う．

また，共分散行列を含んでいないため，一般化最小二乗推定量に対する高速計算可能な代替推定量として最小二乗推定量

$$\hat{\boldsymbol{\beta}}_{LSE} = \left(X'X\right)^{-1} X'\boldsymbol{y},$$

も考えられる．本論文ではシミュレーションを通じて，一般化最小二乗推定量，擬似最良推定量，最小二乗推定量の 3 つを比較する．数値シミュレーションにおいて，真の共分散関数が Matérn 共分散関数のような距離のみに依存する等方型共分散関数であったとしても，2 つの共分散関数の積 (分離型共分散関数) を使用している擬似最良推定量は，最小二乗推定量が一般化最小二乗推定量に対して漸近有効でないとき最小二乗推定量よりも良い推定精度を示し，計算時間の観点からは一般化最小二乗推定量を改善している．これは，擬似最良推定量が高速計算可能な形で近似された空間共分散構造を持っているからだと考えられる．