Assignment

Recall the SAR model $\boldsymbol{y}_n = \lambda \boldsymbol{W}_n \boldsymbol{y}_n + \boldsymbol{X}_n \boldsymbol{\beta} + \boldsymbol{\varepsilon}$. The regressor matrix $\boldsymbol{X}_n$ is an $(n \times k)$ matrix of constant regressors.

1. Define $\boldsymbol{S}_n(\lambda) = \boldsymbol{I}_n - \lambda \boldsymbol{W}_n$ and assume that $\boldsymbol{S}_n^{-1}(\lambda)$ exists. Prove that the log-likelihood under $\boldsymbol{\varepsilon}_n \sim \mathrm{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n)$ is equal to

$$\begin{aligned}\log L_n(\boldsymbol{\theta}) = -\tfrac{n}{2}\log(2\pi) - \tfrac{n}{2}\log(\sigma^2) + \log\big(\det(\boldsymbol{S}_n(\lambda))\big) \\ - \tfrac{1}{2\sigma^2}\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\boldsymbol{\beta}\Big)'\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\boldsymbol{\beta}\Big),\end{aligned}$$

where $\boldsymbol{\theta} = (\lambda, \boldsymbol{\beta}', \sigma^2)'$.

Solution: Solving for $\boldsymbol{y}_n$, we get $\boldsymbol{y}_n = \boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{X}_n\boldsymbol{\beta} + \boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{\varepsilon}$. Apparently, $\boldsymbol{y}_n$ is an affine transformation of the normally distributed random vector $\boldsymbol{\varepsilon}_n$ and thus normally distributed itself. Straightforward calculations provide

$$\mathbb{E}\big[\boldsymbol{y}_n\big] = \mathbb{E}\big[\boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{X}_n\boldsymbol{\beta} + \boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{\varepsilon}\big] = \boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{X}_n\boldsymbol{\beta} + \boldsymbol{S}_n^{-1}(\lambda)\mathbb{E}\big[\boldsymbol{\varepsilon}\big] = \boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{X}_n\boldsymbol{\beta},$$

and

$$\begin{aligned}\mathbb{V}\mathrm{ar}\big[\boldsymbol{y}_n\big] &= \mathbb{V}\mathrm{ar}\big[\boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{X}_n\boldsymbol{\beta} + \boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{\varepsilon}\big] = \mathbb{V}\mathrm{ar}\big[\boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{\varepsilon}\big] \\ &= \boldsymbol{S}_n^{-1}(\lambda)\,\mathbb{V}\mathrm{ar}\big[\boldsymbol{\varepsilon}\big]\boldsymbol{S}_n^{-1\prime}(\lambda) = \sigma^2\boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{S}_n^{-1\prime}(\lambda).\end{aligned}$$

Our first conclusion is

$$\boldsymbol{y}_n \sim \mathrm{N}\Big(\boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{X}_n\boldsymbol{\beta},\, \sigma^2\boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{S}_n^{-1\prime}(\lambda)\Big). \tag{1}$$

We subsequently compute the associated log-likelihood. Recall that a normally distributed random n-vector, say $\boldsymbol{Y} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, has pdf

$$f(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \det(2\pi\boldsymbol{\Sigma})^{-1/2}\exp\big(-\tfrac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\big)$$

and log-likelihood

$$\log L(\mu, \boldsymbol{\Sigma}) = -\tfrac{n}{2}\log(2\pi) - \tfrac{1}{2}\log\big(\det(\boldsymbol{\Sigma})\big) - \tfrac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}).$$

It remains to evaluate this log-likelihood using the distribution in (1). We provide the following intermediate results:

- Exploiting the properties of the determinant gives $\det\big(\sigma^2\boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{S}_n^{-1\prime}(\lambda)\big) = \sigma^{2n}\det\big(\boldsymbol{S}_n^{-1}(\lambda)\big)\det\big(\boldsymbol{S}_n^{-1\prime}(\lambda)\big) = \sigma^{2n}\det\big(\boldsymbol{S}_n(\lambda)\big)^{-2}$.
- $\big(\mathbb{V}\mathrm{ar}\big[\boldsymbol{y}_n\big]\big)^{-1} = \frac{1}{\sigma^2}\boldsymbol{S}_n'(\lambda)\boldsymbol{S}_n(\lambda)$.
- Finally,

$$\begin{aligned}&(\boldsymbol{y}_n - \mathbb{E}[\boldsymbol{y}_n])'\big(\mathbb{V}\mathrm{ar}\big[\boldsymbol{y}_n\big]\big)^{-1}(\boldsymbol{y}_n - \mathbb{E}[\boldsymbol{y}_n]) \\ &\quad= \frac{1}{\sigma^2}\Big(\boldsymbol{y}_n - \boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{X}_n\boldsymbol{\beta}\Big)'\boldsymbol{S}_n'(\lambda)\boldsymbol{S}_n(\lambda)\Big(\boldsymbol{y}_n - \boldsymbol{S}_n^{-1}(\lambda)\boldsymbol{X}_n\boldsymbol{\beta}\Big) \\ &\quad= \frac{1}{\sigma^2}\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\boldsymbol{\beta}\Big)'\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\boldsymbol{\beta}\Big).\end{aligned}$$

The claim follows after combining these intermediate results.

2. Concentrate the log-likelihood with respect to σ^2 and show that

$$\widehat{\sigma}^2(\lambda) = \frac{1}{n}\boldsymbol{y}_n'\boldsymbol{S}_n'(\lambda)\boldsymbol{M}_X\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n,$$

where $\boldsymbol{M}_X = \boldsymbol{I}_n - \boldsymbol{X}_n(\boldsymbol{X}_n'\boldsymbol{X}_n)^{-1}\boldsymbol{X}_n'$.

Solution: A straightforward calculation (σ^2 is only a scalar) gives

$$\begin{aligned}\frac{\partial}{\partial\sigma^2}\log L_n(\boldsymbol{\theta}) &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\boldsymbol{\beta}\Big)'\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\boldsymbol{\beta}\Big)\\ &= -\frac{n}{2\sigma^4}\left(\sigma^2 - \frac{1}{n}\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\boldsymbol{\beta}\Big)'\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\boldsymbol{\beta}\Big)\right).\end{aligned}$$

The optimal value for σ^2 (given λ and $\boldsymbol{\beta}$) is

$$\widehat{\sigma}^2(\lambda,\boldsymbol{\beta}) = \frac{1}{n}\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\boldsymbol{\beta}\Big)'\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\boldsymbol{\beta}\Big).$$

As we have determined the optimal choice of $\boldsymbol{\beta}$, $\widehat{\boldsymbol{\beta}}(\lambda) = \big(\boldsymbol{X}_n'\boldsymbol{X}_n\big)^{-1}\boldsymbol{X}_n'\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n$, we can substitute. The substitution is easier if one realises that

$$\begin{aligned}\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\widehat{\boldsymbol{\beta}}(\lambda) &= \boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\big(\boldsymbol{X}_n'\boldsymbol{X}_n\big)^{-1}\boldsymbol{X}_n'\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n\\ &= \Big[\boldsymbol{I}_n - \boldsymbol{X}_n\big(\boldsymbol{X}_n'\boldsymbol{X}_n\big)^{-1}\boldsymbol{X}_n'\Big]\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n = \boldsymbol{M}_X\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n.\end{aligned}$$

Noting that $\boldsymbol{M}_X$ is idempotent (i.e. $\boldsymbol{M}_X = \boldsymbol{M}_X'$ and $\boldsymbol{M}_X^2 = \boldsymbol{M}_X$), we recover

$$\begin{aligned}\widehat{\sigma}^2(\lambda) = \widehat{\sigma}^2(\lambda,\widehat{\boldsymbol{\beta}}(\lambda)) &= \frac{1}{n}\Big(\boldsymbol{M}_X\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n\Big)'\Big(\boldsymbol{M}_X\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n\Big)\\ &= \frac{1}{n}\boldsymbol{y}_n'\boldsymbol{S}_n'(\lambda)\boldsymbol{M}_X\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n.\end{aligned}$$

3. Derive the concentrated log-likelihood

$$\log L_n(\lambda) = -\tfrac{n}{2}\Big(\log(2\pi) + 1\Big) - \tfrac{n}{2}\log\big(\widehat{\sigma}^2(\lambda)\big) + \log\big(\det(\boldsymbol{S}_n(\lambda))\big).$$

Solution: We only need to substitute $\widehat{\boldsymbol{\beta}}(\lambda)$ and $\widehat{\sigma}^2(\lambda)$. The details are below:

$$\begin{aligned}\log L_n(\lambda) &= -\tfrac{n}{2}\log(2\pi) - \tfrac{n}{2}\log\big(\widehat{\sigma}^2(\lambda)\big) + \log\big(\det(\boldsymbol{S}_n(\lambda))\big)\\ &\qquad - \tfrac{1}{2\widehat{\sigma}^2(\lambda)}\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\widehat{\boldsymbol{\beta}}(\lambda)\Big)'\Big(\boldsymbol{S}_n(\lambda)\boldsymbol{y}_n - \boldsymbol{X}_n\widehat{\boldsymbol{\beta}}(\lambda)\Big)\\ &= -\tfrac{n}{2}\log(2\pi) - \tfrac{n}{2}\log\big(\widehat{\sigma}^2(\lambda)\big) + \log\big(\det(\boldsymbol{S}_n(\lambda))\big) - \tfrac{n}{2}\\ &= -\tfrac{n}{2}\Big(\log(2\pi) + 1\Big) - \tfrac{n}{2}\log\big(\widehat{\sigma}^2(\lambda)\big) + \log\big(\det(\boldsymbol{S}_n(\lambda))\big)\end{aligned}$$