WORKING PAPER

11/2012

Conditional posteriors for the reduced rank regression model

Sune Karlsson

Statistics

ISSN 1403-0586

http://www.oru.se/Institutioner/Handelshogskolan-vid-Orebro-universitet/Forskning/Publikationer/Working-papers/

Örebro University School of Business

701 82 Örebro

SWEDEN

Conditional posteriors for the reduced rank regression model

Sune Karlsson*
Örebro University

May 27, 2012

Abstract

The multivariate reduced rank regression model plays an important role in econometrics. Examples include co-integration analysis and models with a factor structure. Geweke (1996) provided the foundations for a Bayesian analysis of this model. Unfortunately several of the full conditional posterior distributions, which forms the basis for constructing a Gibbs sampler for the poster distribution, given by Geweke contains errors. This paper provides correct full conditional posteriors for the reduced rank regression model under the prior distributions considered by Geweke.

JEL-codes: C11, C30, C53

Keywords: Gibbs sampling; full conditional posterior

*Department of Statistics, Örebro University, 701 82 Örebro, Sweden. E-mail: Sune.Karlsson@oru.se

1 Introduction

Geweke (1996) studied the reduced rank regression model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\Theta} + \mathbf{Z}\mathbf{A} + \mathbf{E} \tag{1}$$

where $\mathbf{Y}$ is an $n \times L$ matrix of dependent variables, $\mathbf{X}$ and $\mathbf{Z}$ contains p and k explanatory variables and $\boldsymbol{\Theta}$ and $\mathbf{A}$ are parameter matrices where $\boldsymbol{\Theta}$ is assumed to have reduced rank $q < \min(L, p)$. The rows of $\mathbf{E}$ are assumed to be independent normal with mean zero and variance matrix $\boldsymbol{\Sigma}$, i.e. $\mathbf{e} = \mathrm{vec}(\mathbf{E}) \sim N(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I})$. Under the reduced rank assumption $\boldsymbol{\Theta}$ can be factored into $\boldsymbol{\Theta} = \boldsymbol{\Psi}\boldsymbol{\Phi}$ with $\boldsymbol{\Psi}$ a $p \times q$ matrix and $\boldsymbol{\Phi}$ a $q \times L$ matrix, both of rank q. To identify the model Geweke considers two normalizations, $\boldsymbol{\Phi} = (\mathbf{I}_q, \boldsymbol{\Phi}^*)$ with $\boldsymbol{\Psi}$ unrestricted (normalization 1) and $\boldsymbol{\Psi}' = (\mathbf{I}_q, \boldsymbol{\Psi}^{*\prime})$ with $\boldsymbol{\Phi}$ unrestricted (normalization 2). Geweke then proceeds to derive the full conditional posterior distributions for the joint prior distributions

$$\pi(\boldsymbol{\Sigma}, \mathbf{A}, \boldsymbol{\Psi}, \boldsymbol{\Phi}^*) \propto |\boldsymbol{\Psi}|^{-(L+\underline{v}+1)/2} \exp\left[-\frac{1}{2}\,\mathrm{tr}\,\underline{\mathbf{S}}\boldsymbol{\Sigma}^{-1}\right] \exp\left[-\frac{\tau^2}{2}\,\mathrm{tr}\left(\mathbf{A}'\mathbf{A} + \boldsymbol{\Psi}'\boldsymbol{\Psi} + \boldsymbol{\Phi}^{*\prime}\boldsymbol{\Phi}^*\right)\right]$$

and

$$\pi(\boldsymbol{\Sigma}, \mathbf{A}, \boldsymbol{\Psi}^*, \boldsymbol{\Phi}) \propto |\boldsymbol{\Psi}|^{-(L+\underline{v}+1)/2} \exp\left[-\frac{1}{2}\,\mathrm{tr}\,\underline{\mathbf{S}}\boldsymbol{\Sigma}^{-1}\right] \exp\left[-\frac{\tau^2}{2}\,\mathrm{tr}\left(\mathbf{A}'\mathbf{A} + \boldsymbol{\Psi}^{*\prime}\boldsymbol{\Psi}^* + \boldsymbol{\Phi}'\boldsymbol{\Phi}\right)\right]$$

under normalizations 1 and 2 respectively. The priors are standard, $\boldsymbol{\Sigma}$ is distributed as inverse Wishart with parameter matrix $\underline{\mathbf{S}}$ and $\underline{v}$ degrees of freedom, $iW(\underline{\mathbf{S}}, \underline{v})$, and the elements of $\mathbf{A}$, $(\boldsymbol{\Psi}, \boldsymbol{\Phi}^*)$ and $(\boldsymbol{\Phi}^*, \boldsymbol{\Psi})$ are independent normal with mean zero and variance $1/\tau^2$.

The conditional posteriors given in Geweke (1996) are unfortunately incorrect in several cases – the conditional posteriors for $\boldsymbol{\Phi}^*$ and $\boldsymbol{\Psi}$ with $\tau > 0$ in normalization 1 and the conditional posterior for $\boldsymbol{\Psi}^*$ in normalization 2 with $\tau > 0$. The errors are trivial in nature but the incorrect expressions for the parameters of the posterior distributions have been picked up in the literature, e.g. Carriero, Kapetanios and Marcellino (2011), and used to construct Gibbs samplers with incorrect stationary distributions. Geweke (2004) developed a method for checking the correctness of posterior simulators and detected problems with the Gibbs sampler coded up for the 1996 paper but failed to connect this with the incorrect expressions for the full conditional posteriors. It thus seems to be of some importance to provide correct posterior distributions.

Section 2 below provides correct expressions for the full conditional posteriors in the reduced rank regression model. For completeness, the conditional posteriors for $\boldsymbol{\Sigma}$, $\boldsymbol{\Phi}$ (in normalization 2) and $\mathbf{A}$ (which are correct in Geweke (1996)) are restated. Section 3 provides derivations of the full conditional posteriors for $\boldsymbol{\Phi}^*$ and $\boldsymbol{\Psi}$ in normalization 1 and $\boldsymbol{\Psi}^*$ in normalization 2.

2 Full conditional posteriors

Following Geweke (1996) we rewrite the model as

$$\mathbf{Y}^* = \mathbf{Y} - \mathbf{Z}\mathbf{A} = \mathbf{X}\boldsymbol{\Theta} + \mathbf{E} \tag{2}$$

when conditioning on $\mathbf{A}$. In addition, conditioning on $\mathbf{\Theta} = \mathbf{\Psi\Phi}$, we can write the model as

$$\mathbf{W} = \mathbf{Y} - \mathbf{X\Theta} = \mathbf{ZA} + \mathbf{E} \tag{3}$$

2.1 A and Σ

Given $\mathbf{\Theta}$ the model reduces to the standard multivariate regression model (3) with the familiar full conditional posterior distributions given in Geweke (1996)

$$\mathbf{\Sigma}|\mathbf{Y}, \mathbf{\Psi}, \mathbf{\Phi}, \mathbf{A} \sim iW\left(\overline{v}, \overline{\mathbf{S}}\right)$$

for $\overline{v} = n + \underline{v}$ and $\overline{\mathbf{S}} = \underline{\mathbf{S}} + (\mathbf{W} - \mathbf{ZA})'(\mathbf{W} - \mathbf{ZA}) = \underline{\mathbf{S}} + (\mathbf{Y}^* - \mathbf{X\Theta})'(\mathbf{Y}^* - \mathbf{X\Theta})$ and

$$\text{vec}(\mathbf{A})|\mathbf{Y}, \mathbf{\Psi}, \mathbf{\Phi}, \mathbf{\Sigma} \sim N\left(\text{vec}\left(\widehat{\mathbf{A}}\right), \mathbf{\Sigma} \otimes \mathbf{Z}'\mathbf{Z} + \tau^2\mathbf{I}_{kL}\right)$$

for $\widehat{\mathbf{A}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{W} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{Y} - \mathbf{X\Theta})$.

2.2 Φ* and Ψ in normalization 1

Let $\boldsymbol{\phi}^* = \text{vec}(\mathbf{\Phi}^*)$, the correct full conditional posterior is then

$$\boldsymbol{\phi}^*|\mathbf{Y}, \mathbf{A}, \mathbf{\Psi}, \mathbf{\Sigma} \sim N\left(\overline{\boldsymbol{\phi}}^*, \mathbf{V}_{\phi^*}\right) \tag{4}$$

for

$$\begin{aligned}\mathbf{V}_{\phi^*} &= \left[\tau^2\mathbf{I} + \mathbf{\Sigma}^{22} \otimes \mathbf{\Psi}'\mathbf{X}'\mathbf{X\Psi}\right]^{-1} \\ \overline{\boldsymbol{\phi}}^* &= \mathbf{V}_{\phi^*}\,\text{vec}\left[\mathbf{\Psi}'\mathbf{X}'\mathbf{Y}_1^*\mathbf{\Sigma}^{12} - \mathbf{\Psi}'\mathbf{X}'\mathbf{X\Psi\Sigma}^{12} + \mathbf{\Psi}'\mathbf{X}'\mathbf{Y}_2^*\mathbf{\Sigma}^{22}\right]\end{aligned}$$

where $(\mathbf{Y}_1^*, \mathbf{Y}_2^*) = \mathbf{Y}^*$ partitions $\mathbf{Y}^*$ into $n \times q$ and $n \times (L - q)$ matrices and $\mathbf{\Sigma}^{-1}$ is partitioned into

$$\mathbf{\Sigma}^{-1} = \begin{pmatrix} \mathbf{\Sigma}^{11} & \mathbf{\Sigma}^{12} \\ \mathbf{\Sigma}^{21} & \mathbf{\Sigma}^{22} \end{pmatrix} \tag{5}$$

with $\mathbf{\Sigma}^{11}$ and $\mathbf{\Sigma}^{22}$ $q \times q$ and $(L - q) \times (L - q)$ matrices.

Similarly for $\boldsymbol{\psi} = \text{vec}(\mathbf{\Psi})$, the correct full conditional posterior is.

$$\boldsymbol{\psi}|\mathbf{Y}, \mathbf{A}, \mathbf{\Phi}, \mathbf{\Sigma} \sim N\left(\overline{\boldsymbol{\psi}}, \mathbf{V}_{\psi}\right) \tag{6}$$

for

$$\begin{aligned}\mathbf{V}_{\psi} &= \left(\tau^2\mathbf{I} + \mathbf{\Phi\Sigma}^{-1}\mathbf{\Phi}' \otimes \mathbf{X}'\mathbf{X}\right)^{-1} \\ \overline{\boldsymbol{\psi}} &= \mathbf{V}_{\psi}\,\text{vec}\left(\mathbf{X}'\mathbf{Y}^*\mathbf{\Sigma}^{-1}\mathbf{\Phi}'\right).\end{aligned}$$

Geweke (1996) also consider the case with improper uniform priors on $\mathbf{\Phi}^*$ and $\mathbf{\Psi}$ and gives correct expressions for the full conditional posteriors in this case. These posteriors can be obtained as special cases of the results above by setting $\tau = 0$.

2.3 $\mathbf{\Phi}$ and $\mathbf{\Psi}^*$ in normalization 2

The full conditional posterior distribution of $\boldsymbol{\phi} = \text{vec}(\mathbf{\Phi})$ is given in Geweke (1996) as

$$\boldsymbol{\phi}|\mathbf{Y}, \mathbf{A}, \mathbf{\Psi}, \mathbf{\Sigma} \sim N\left(\overline{\boldsymbol{\phi}}, \mathbf{V}_\phi\right) \tag{7}$$

for

$$\begin{aligned}\mathbf{V}_\phi &= \left(\tau^2\mathbf{I} + \mathbf{\Sigma}^{-1} \otimes \mathbf{\Psi}'\mathbf{X}'\mathbf{X}\mathbf{\Psi}\right)^{-1} \\ \overline{\boldsymbol{\phi}} &= \mathbf{V}_\phi \, \text{vec}\left(\mathbf{\Psi}'\mathbf{X}'\mathbf{Y}^*\mathbf{\Sigma}^{-1}\right).\end{aligned}$$

For the correct full conditional poster distribution of $\mathbf{\Psi}^*$ let $\boldsymbol{\psi}^* = \text{vec}(\mathbf{\Psi}^*)$. We then have

$$\boldsymbol{\psi}^*|\mathbf{Y}, \mathbf{A}, \mathbf{\Phi}, \mathbf{\Sigma} \sim N\left(\overline{\boldsymbol{\psi}}^*, \mathbf{V}_{\psi^*}\right) \tag{8}$$

for

$$\begin{aligned}\mathbf{V}_{\psi^*} &= \left(\tau^2\mathbf{I} + \mathbf{\Phi}\mathbf{\Sigma}^{-1}\mathbf{\Phi}' \otimes \mathbf{X}_2'\mathbf{X}_2\right)^{-1} \\ \overline{\boldsymbol{\psi}}^* &= \mathbf{V}_{\psi^*} \, \text{vec}\left(\mathbf{X}_2'\left(\mathbf{Y}^* - \mathbf{X}_1\mathbf{\Phi}\right)\mathbf{\Sigma}^{-1}\mathbf{\Phi}'\right)\end{aligned}$$

where $\mathbf{X}_1$ contains the first q columns of $\mathbf{X}$ and $\mathbf{X}_2$ the remaining columns.

As for normalization 1 the expressions for the full conditional posteriors with improper uniform priors on $\mathbf{\Phi}$ and $\mathbf{\Psi}^*$ given in Geweke (1996) are correct and can be obtained as special cases by setting τ to zero above.

3 Derivations

3.1 Conditional posterior for $\mathbf{\Phi}^*$ in normalization 1

In normalization 1 the model can be rewritten as

$$\begin{aligned}\mathbf{Y}^* = \left(\mathbf{Y}_1^*, \mathbf{Y}_2^*\right) &= \left(\mathbf{X}\mathbf{\Psi}, \mathbf{X}\mathbf{\Psi}\mathbf{\Phi}^*\right) + \mathbf{E} \\ \left(\widetilde{\mathbf{Y}}_1^*, \mathbf{Y}_2^*\right) = \left(\mathbf{Y}_1^* - \mathbf{X}\mathbf{\Psi}, \mathbf{Y}_2^*\right) &= \left(\mathbf{0}, \mathbf{X}\mathbf{\Psi}\mathbf{\Phi}^*\right) + \mathbf{E}\end{aligned}$$

with $\mathbf{Y}_1^*$ and $\mathbf{Y}_2^*$ $n \times q$ and $n \times (L - q)$ matrices. Vectorizing yields

$$\widetilde{\mathbf{y}}^* = \begin{pmatrix} \widetilde{\mathbf{y}}_1^* \\ \mathbf{y}_2^* \end{pmatrix} = \text{vec}\left(\widetilde{\mathbf{Y}}_1^*, \mathbf{Y}_2^*\right) = \begin{pmatrix} \mathbf{0} \\ \mathbf{I} \otimes \mathbf{X}\mathbf{\Psi} \end{pmatrix} \boldsymbol{\phi}^* + \mathbf{e} = \mathbf{H}\boldsymbol{\phi}^* + \mathbf{e}.$$

Conditional on $\mathbf{A}$, $\mathbf{\Phi}$ and $\mathbf{\Sigma}$ this is a standard regression model with a normal prior on $\boldsymbol{\phi}^*$, $\boldsymbol{\phi}^* \sim N\left(\mathbf{0}, \tau^{-2}\mathbf{I}\right)$, and known error variance matrix. Routine calculations yield the conditional posterior

$$\boldsymbol{\phi}^*|\mathbf{Y}, \mathbf{A}, \mathbf{\Psi}, \mathbf{\Sigma} \sim N\left(\overline{\boldsymbol{\phi}}^*, \mathbf{V}_{\phi^*}\right)$$

with

$$\begin{aligned}\mathbf{V}_{\phi^*} &= \left[\tau^2\mathbf{I} + \mathbf{H}'\left(\mathbf{\Sigma}^{-1} \otimes \mathbf{I}\right)\mathbf{H}\right]^{-1} = \left[\tau^2\mathbf{I} + \begin{pmatrix} \mathbf{0} \\ \mathbf{I} \otimes \mathbf{X}\mathbf{\Psi} \end{pmatrix}'\left(\mathbf{\Sigma}^{-1} \otimes \mathbf{I}\right)\begin{pmatrix} \mathbf{0} \\ \mathbf{I} \otimes \mathbf{X}\mathbf{\Psi} \end{pmatrix}\right]^{-1} \\ &= \left[\tau^2\mathbf{I} + \left(\mathbf{I} \otimes \mathbf{X}\mathbf{\Psi}\right)'\left(\mathbf{\Sigma}^{22} \otimes \mathbf{I}\right)\left(\mathbf{I} \otimes \mathbf{X}\mathbf{\Psi}\right)\right]^{-1} = \left[\tau^2\mathbf{I} + \mathbf{\Sigma}^{22} \otimes \mathbf{\Psi}'\mathbf{X}'\mathbf{X}\mathbf{\Psi}\right]^{-1}\end{aligned}$$

for $\boldsymbol{\Sigma}^{22}$ the lower right $(L-q)\times(L-q)$ block of $\boldsymbol{\Sigma}^{-1}$ in (5) and

$$\begin{aligned}
\overline{\boldsymbol{\phi}}^{*} &= \mathbf{V}_{\phi^*}\left[\mathbf{H}'\left(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I}\right)\mathbf{H}\right]\widehat{\boldsymbol{\phi}}^{*} \\
&= \mathbf{V}_{\phi^*}\begin{pmatrix}\mathbf{0}\\ \mathbf{I}\otimes\mathbf{X}\boldsymbol{\Psi}\end{pmatrix}'\left(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I}\right)\begin{pmatrix}\widetilde{\mathbf{y}}_1^{*}\\ \mathbf{y}_2^{*}\end{pmatrix} \\
&= \mathbf{V}_{\phi^*}\left[\left(\boldsymbol{\Sigma}^{21}\otimes\boldsymbol{\Psi}'\mathbf{X}'\right)\widetilde{\mathbf{y}}_1^{*}+\left(\boldsymbol{\Sigma}^{22}\otimes\boldsymbol{\Psi}'\mathbf{X}'\right)\mathbf{y}_2^{*}\right] \\
&= \mathbf{V}_{\phi^*}\operatorname{vec}\left[\boldsymbol{\Psi}'\mathbf{X}'\mathbf{Y}_1^{*}\boldsymbol{\Sigma}^{12}-\boldsymbol{\Psi}'\mathbf{X}'\mathbf{X}\boldsymbol{\Psi}\boldsymbol{\Sigma}^{12}+\boldsymbol{\Psi}'\mathbf{X}'\mathbf{Y}_2^{*}\boldsymbol{\Sigma}^{22}\right]
\end{aligned}$$

where $\widehat{\boldsymbol{\phi}}^{*}$ is the generalized least squares estimate,

$$\begin{aligned}
\widehat{\boldsymbol{\phi}}^{*} &= \left[\mathbf{H}'\left(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I}\right)\mathbf{H}\right]^{-1}\mathbf{H}'\left(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I}\right)\widetilde{\mathbf{y}}^{*} \\
&= \left(\left(\boldsymbol{\Sigma}^{22}\right)^{-1}\boldsymbol{\Sigma}^{21}\otimes\left(\boldsymbol{\Psi}'\mathbf{X}'\mathbf{X}\boldsymbol{\Psi}\right)^{-1}\boldsymbol{\Psi}'\mathbf{X}'\right)\widetilde{\mathbf{y}}_1^{*}+\left(\mathbf{I}\otimes\left(\boldsymbol{\Psi}'\mathbf{X}'\mathbf{X}\boldsymbol{\Psi}\right)^{-1}\boldsymbol{\Psi}'\mathbf{X}'\right)\mathbf{y}_2^{*}
\end{aligned}$$

or in matrix form

$$\begin{aligned}
\widehat{\boldsymbol{\Phi}}^{*} &= \left(\boldsymbol{\Psi}'\mathbf{X}'\mathbf{X}\boldsymbol{\Psi}\right)^{-1}\boldsymbol{\Psi}'\mathbf{X}'\widetilde{\mathbf{Y}}_1^{*}\boldsymbol{\Sigma}^{12}\left(\boldsymbol{\Sigma}^{22}\right)^{-1}+\left(\boldsymbol{\Psi}'\mathbf{X}'\mathbf{X}\boldsymbol{\Psi}\right)^{-1}\boldsymbol{\Psi}'\mathbf{X}'\mathbf{Y}_2^{*} \\
&= \left(\boldsymbol{\Psi}'\mathbf{X}'\mathbf{X}\boldsymbol{\Psi}\right)^{-1}\boldsymbol{\Psi}'\mathbf{X}'\mathbf{Y}_1^{*}\boldsymbol{\Sigma}^{12}\left(\boldsymbol{\Sigma}^{22}\right)^{-1}-\boldsymbol{\Sigma}^{12}\left(\boldsymbol{\Sigma}^{22}\right)^{-1}+\left(\boldsymbol{\Psi}'\mathbf{X}'\mathbf{X}\boldsymbol{\Psi}\right)^{-1}\boldsymbol{\Psi}'\mathbf{X}'\mathbf{Y}_2^{*}.
\end{aligned}$$

Setting τ^2 to zero recovers the results of Geweke for this case whereas both the posterior mean and variance for the $\tau > 0$ case are incorrectly stated in equation (13) of Geweke (1996).

3.2 Conditional posterior for $\boldsymbol{\Psi}$ in normalization 1

To derive the full conditional posterior for $\boldsymbol{\Psi}$ we vectorize (2)

$$\begin{aligned}
\mathbf{y}^{*} &= \left(\boldsymbol{\Phi}'\otimes\mathbf{X}\right)\operatorname{vec}\left(\boldsymbol{\Psi}\right)+\mathbf{e}. \\
&= \left(\boldsymbol{\Phi}'\otimes\mathbf{X}\right)\boldsymbol{\psi}+\mathbf{e}
\end{aligned}$$

With the prior $\boldsymbol{\psi}\sim N\left(\mathbf{0},\tau^{-2}\mathbf{I}\right)$ and $\mathbf{e}\sim N\left(\mathbf{0},\boldsymbol{\Sigma}\otimes\mathbf{I}\right)$ standard results yield

$$\boldsymbol{\psi}|\mathbf{Y},\mathbf{A},\boldsymbol{\Psi},\boldsymbol{\Sigma}\sim N\left(\overline{\boldsymbol{\psi}},\mathbf{V}_{\psi}\right)$$

for

$$\begin{aligned}
\mathbf{V}_{\psi} &= \left(\tau^2\mathbf{I}+\boldsymbol{\Phi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}'\otimes\mathbf{X}'\mathbf{X}\right)^{-1} \\
\overline{\boldsymbol{\psi}} &= \mathbf{V}_{\psi}\left(\boldsymbol{\Phi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}'\otimes\mathbf{X}'\mathbf{X}\right)\widehat{\boldsymbol{\psi}}=\mathbf{V}_{\psi}\left(\boldsymbol{\Phi}\boldsymbol{\Sigma}^{-1}\otimes\mathbf{X}'\right)\mathbf{y}^{*} \\
&= \mathbf{V}_{\psi}\operatorname{vec}\left(\mathbf{X}'\mathbf{Y}^{*}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}'\right)
\end{aligned}$$

since $\widehat{\boldsymbol{\psi}}=\left(\boldsymbol{\Phi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}'\otimes\mathbf{X}'\mathbf{X}\right)^{-1}\left(\boldsymbol{\Phi}\boldsymbol{\Sigma}^{-1}\otimes\mathbf{X}'\right)\mathbf{y}$.

In Geweke (1996) the matrix $\widetilde{\boldsymbol{\Sigma}}^{11}$ appears in the posterior variance instead of $\boldsymbol{\Phi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}'$. $\widetilde{\boldsymbol{\Sigma}}^{11}$ is the upper left submatrix of $\widetilde{\boldsymbol{\Sigma}}^{-1}=\left(\mathbf{C}'\boldsymbol{\Sigma}\mathbf{C}\right)^{-1}=\mathbf{C}^{-1}\boldsymbol{\Sigma}^{-1}\mathbf{C}^{-\mathsf{T}}$ for $\mathbf{C}=\left(\boldsymbol{\Phi}^{+},\boldsymbol{\Phi}^{0}\right)$ where $\boldsymbol{\Phi}^{+}$ is the Moore-Penrose generalized inverse of $\boldsymbol{\Phi}$, i.e. $\boldsymbol{\Phi}\boldsymbol{\Phi}^{+}=\mathbf{I}$ and $\boldsymbol{\Phi}^{0}$ is orthogonal to $\boldsymbol{\Phi}^{+}$. It is easily verified that $\mathbf{C}^{-1}=\begin{pmatrix}\boldsymbol{\Phi}\\ \boldsymbol{\Phi}^{0\prime}\end{pmatrix}$ and

$$\widetilde{\boldsymbol{\Sigma}}^{-1}=\begin{pmatrix}\boldsymbol{\Phi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}' & \boldsymbol{\Phi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}^{0\prime}\\ \boldsymbol{\Phi}^{0\prime}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}' & \boldsymbol{\Phi}^{0\prime}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}^{0\prime}\end{pmatrix}. \tag{9}$$

The expression for the posterior variance $\mathbf{V}_\psi$ is thus in agreement with Geweke for both the $\tau = 0$ and $\tau > 0$ case.

Turning to the posterior mean, this simplifies to $\overline{\boldsymbol{\psi}} = \widehat{\boldsymbol{\psi}} = \left(\left(\boldsymbol{\Phi\Sigma}^{-1}\boldsymbol{\Phi}' \right)^{-1} \boldsymbol{\Phi\Sigma}^{-1} \otimes \left(\mathbf{X}'\mathbf{X} \right)^{-1} \mathbf{X}' \right) \mathbf{y}^*$ when $\tau = 0$. Reshaping this to a matrix we have

$$\widehat{\boldsymbol{\Psi}} = \left(\mathbf{X}'\mathbf{X} \right)^{-1} \mathbf{X}'\mathbf{Y}^*\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}' \left(\boldsymbol{\Phi\Sigma}^{-1}\boldsymbol{\Phi}' \right)^{-1}$$

and we can verify Geweke's result for $\tau = 0$ using that

$$\begin{aligned} \boldsymbol{\Phi}^+ + \boldsymbol{\Phi}^0 \widetilde{\boldsymbol{\Sigma}}^{21} \left(\widetilde{\boldsymbol{\Sigma}}^{11} \right)^{-1} &= \boldsymbol{\Phi}^+ + \boldsymbol{\Phi}^0\boldsymbol{\Phi}^{0\prime}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}' \left(\boldsymbol{\Phi\Sigma}^{-1}\boldsymbol{\Phi}' \right)^{-1} \\ &= \boldsymbol{\Phi}^+ + \left(\mathbf{I} - \boldsymbol{\Phi}^+\boldsymbol{\Phi} \right) \boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}' \left(\boldsymbol{\Phi\Sigma}^{-1}\boldsymbol{\Phi}' \right)^{-1} \\ &= \boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}' \left(\boldsymbol{\Phi\Sigma}^{-1}\boldsymbol{\Phi}' \right)^{-1}. \end{aligned} \tag{10}$$

For $\tau > 0$ the posterior mean in equation (11) of Geweke is incorrectly stated since the factor $\left(\boldsymbol{\Phi\Sigma}^{-1}\boldsymbol{\Phi}' \otimes \mathbf{X}'\mathbf{X} \right)$ is missing.

3.3 Conditional posterior for $\boldsymbol{\Psi}^*$ in normalization 2

With the normalization $\boldsymbol{\Psi}' = (\mathbf{I}, \boldsymbol{\Psi}^{*\prime})$ and $\boldsymbol{\Phi}$ unrestricted we have

$$\mathbf{X\Psi\Phi} = \mathbf{X} \begin{pmatrix} \mathbf{I}_q \\ \boldsymbol{\Psi}^* \end{pmatrix} \boldsymbol{\Phi} = \left(\mathbf{X}_1 + \mathbf{X}_2\boldsymbol{\Psi}^* \right) \boldsymbol{\Phi} = \mathbf{X}_1\boldsymbol{\Phi} + \mathbf{X}_2\boldsymbol{\Psi}^*\boldsymbol{\Phi}$$

where $\mathbf{X}_1$ contains the first q columns of $\mathbf{X}$ and $\mathbf{X}_2$ the remaining columns. The vectorized equation is thus

$$\mathbf{y}^* = \mathrm{vec}\left(\mathbf{X}_1\boldsymbol{\Phi} \right) + \left(\boldsymbol{\Phi}' \otimes \mathbf{X}_2 \right) \mathrm{vec}\left(\boldsymbol{\Psi}^* \right) + \mathbf{e}$$

or

$$\mathbf{y}^*_\phi = \mathbf{y}^* - \mathrm{vec}\left(\mathbf{X}_1\boldsymbol{\Phi} \right) = \left(\boldsymbol{\Phi}' \otimes \mathbf{X}_2 \right) \boldsymbol{\psi}^* + \mathbf{e}.$$

With the prior $\boldsymbol{\psi}^* \sim N\left(0, \tau^{-2}\mathbf{I}\right)$ the posterior is immediate

$$\boldsymbol{\psi}^* | \mathbf{Y}, \mathbf{A}, \boldsymbol{\Phi}, \boldsymbol{\Sigma} \sim N\left(\overline{\boldsymbol{\psi}}^*, \mathbf{V}_{\psi^*} \right)$$

with

$$\begin{aligned} \mathbf{V}_{\psi^*} &= \left(\tau^2\mathbf{I} + \boldsymbol{\Phi\Sigma}^{-1}\boldsymbol{\Phi}' \otimes \mathbf{X}_2'\mathbf{X}_2 \right)^{-1} \\ \overline{\boldsymbol{\psi}}^* &= \mathbf{V}_{\psi^*} \left(\boldsymbol{\Phi\Sigma}^{-1}\boldsymbol{\Phi}' \otimes \mathbf{X}_2'\mathbf{X}_2 \right) \widehat{\boldsymbol{\psi}}^*. \end{aligned}$$

Using that $\widehat{\boldsymbol{\psi}}^*$ is the GLS estimate,

$$\begin{aligned} \widehat{\boldsymbol{\psi}}^* &= \left(\boldsymbol{\Phi\Sigma}^{-1}\boldsymbol{\Phi}' \otimes \mathbf{X}_2'\mathbf{X}_2 \right)^{-1} \left(\boldsymbol{\Phi\Sigma}^{-1} \otimes \mathbf{X}_2' \right) \mathbf{y}^*_\phi \\ &= \left(\left(\boldsymbol{\Phi\Sigma}^{-1}\boldsymbol{\Phi}' \right)^{-1} \boldsymbol{\Phi\Sigma}^{-1} \otimes \left(\mathbf{X}_2'\mathbf{X}_2 \right)^{-1} \mathbf{X}_2' \right) \mathbf{y}^*_\phi, \end{aligned}$$

the expression for $\overline{\boldsymbol{\psi}}^*$ simplifies to

$$\begin{aligned} \overline{\boldsymbol{\psi}}^* &= \mathbf{V}_{\psi^*} \left(\boldsymbol{\Phi\Sigma}^{-1} \otimes \mathbf{X}_2' \right) \mathrm{vec}\left(\mathbf{Y}^* - \mathbf{X}_1\boldsymbol{\Phi} \right) \\ &= \mathbf{V}_{\psi^*}\, \mathrm{vec}\left[\mathbf{X}_2' \left(\mathbf{Y}^* - \mathbf{X}_1\boldsymbol{\Phi} \right) \boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}' \right]. \end{aligned}$$

In matrix form the GLS estimate (and posterior mean for the $\tau = 0$ case) is

$$
\begin{aligned}
\widehat{\mathbf{\Psi}}^* &= \left(\mathbf{X}_2'\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\left(\mathbf{Y}^* - \mathbf{X}_1\mathbf{\Phi}\right)\mathbf{\Sigma}^{-1}\mathbf{\Phi}'\left(\mathbf{\Phi}\mathbf{\Sigma}^{-1}\mathbf{\Phi}'\right)^{-1} \\
&= \left(\mathbf{X}_2'\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{Y}^*\mathbf{\Sigma}^{-1}\mathbf{\Phi}'\left(\mathbf{\Phi}\mathbf{\Sigma}^{-1}\mathbf{\Phi}'\right)^{-1} - \left(\mathbf{X}_2'\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{X}_1 \\
&= \left(\mathbf{X}_2'\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{Y}^*\left(\mathbf{\Phi}^+ + \mathbf{\Phi}^0\widetilde{\mathbf{\Sigma}}^{21}\left(\widetilde{\mathbf{\Sigma}}^{11}\right)^{-1}\right) - \left(\mathbf{X}_2'\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{X}_1
\end{aligned}
$$

where the last line follows from (10) and $\widetilde{\mathbf{\Sigma}}^{21}$ and $\widetilde{\mathbf{\Sigma}}^{11}$ are submatrices of $\widetilde{\mathbf{\Sigma}}^{-1}$ from (9). This reproduces the posterior mean given by Geweke in equation (14) for this case.

The posterior variance agrees with Geweke for both $\tau = 0$ and $\tau > 0$ but the posterior mean for the $\tau > 0$ case is incorrectly stated in equation (15) of Geweke (1996) since the factor $\left(\mathbf{\Phi}\mathbf{\Sigma}^{-1}\mathbf{\Phi}' \otimes \mathbf{X}_2'\mathbf{X}_2\right)$ is missing.

References

Carriero, A., Kapetanios, G. and Marcellino, M. (2011), 'Forecasting large datasets with bayesian reduced rank multivariate models', *Journal of Applied Econometrics* **26**, 735–761.

Geweke, J. (1996), 'Bayesian reduced rank regression in econometrics', *Journal of Econometrics* **75**, 121–146.

Geweke, J. (2004), 'Getting it right: Joint distribution tests of posterior simulators', *Journal of the American Statistical Association* **99**, 799–804.