# 1 Model and Likelihood

Consider a linear $K$-factor model for $D$ assets of the form

$$y_{it} = \alpha_d + \mathbf{f}_t'\boldsymbol{\beta}_d + \varepsilon_{it}$$

where $d = 1, \ldots, D$ and $t = 1, \ldots, T$ and $\mathbf{f}_t' = (f_{t1}, \ldots, f_{tK})$ is a $K \times 1$ vector. This is a special case of the seemingly unrelated regression (SUR) model in which the regressors are *identical* across equations. Stacking observations for a given time period across assets, define $\mathbf{y}_t' = (y_{1t}, \ldots, y_{Dt})$ and analogously $\boldsymbol{\varepsilon}_t' = (\varepsilon_{t1}, \ldots \varepsilon_{tD})$. Now let $\mathbf{x}_t' = (1, \mathbf{f}_t)$ and $\boldsymbol{\gamma}_d' = (\alpha_d, \boldsymbol{\beta}_d')$ so we have

$$\mathbf{y}_t = X_t\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_t$$

where $X_t = I_D \otimes \mathbf{x}_t'$ and $\boldsymbol{\gamma}' = (\boldsymbol{\gamma}_1', \ldots, \boldsymbol{\gamma}_D')$. Now, suppose that

$$\boldsymbol{\varepsilon}_t | \mathbf{x}_t \sim \text{ iid } \mathcal{N}_D(0, \Sigma)$$

It follows that

$$L(\beta, \Sigma) \propto |\Sigma|^{-T/2} \exp\left[-\frac{1}{2}\sum_{t=1}^{T}\left(\mathbf{y}_t - X_t\boldsymbol{\gamma}\right)' \Sigma^{-1}\left(\mathbf{y}_t - X_t\boldsymbol{\gamma}\right)\right]$$

We parameterize this problem in terms of the $D \times D$ precision matrix $\Sigma$ and the $D(K+1) \times 1$ vector of regression coefficients $\boldsymbol{\gamma}$.

# 2 Prior and Posterior Distribution

To complete the model we specify a prior distribution on the full collection of parameters $\boldsymbol{\theta} = (\boldsymbol{\gamma}, \Omega^{-1})$, namely

$$\pi(\boldsymbol{\theta}) = \mathcal{N}_{d+p}\left(\gamma|\gamma_0, G_0\right)\mathcal{W}_d\left(\Omega^{-1}|\rho_0, R_0\right)$$

This prior is conditionally conjugate with the normal likelihood yielding the posterior distribution $Y_n = (\{y_t\}, \{f_t\})$

$$\pi(\theta|Y_n) \propto \mathcal{N}_{d+p}\left(\gamma|\gamma_0, G_0\right)\mathcal{W}_d\left(\Omega^{-1}|\rho_0, R_0\right)\prod_{t=1}^{n}\mathcal{N}_d(y_t|X_t\gamma, \Omega)$$

## 3 MCMC

We can summarize this posterior distribution by the MCMC algorithm of Chib and Greenberg (1995). It requires the repeated recursive sampling of the following two conditional distributions:

$$\gamma|Y_n, \Omega^{-1} \sim \mathcal{N}_{d+p}\left(\hat{\gamma}, G_n\right)$$

where

$$\hat{\gamma} = G_n\left(G_0^{-1}\gamma_0 + \sum_{t=1}^{n} X_t'\Omega^{-1}y_t\right)$$
$$G_n = \left(G_0^{-1} + \sum_{t=1}^{n} X_t'\Omega^{-1}X_t\right)^{-1}$$

and

$$\Omega^{-1}|Y_n, \gamma \sim \mathcal{W}_d\left(\rho_0 + n, \left(R_0^{-1} + \sum_{t=1}^{n}\left(y_t - X_t\gamma\right)\left(y_t - X_t\gamma\right)'\right)^{-1}\right)$$

## 4 Marginal likelihood

The marginal likelihood is available by the method of Chib (1995). From the Chib (1995) identity, we have

$$\log m\left(\{y_t\}\right) = \log \pi\left(\gamma^*\right) + \log \pi\left(\Omega^{-1*}\right) + \sum_{t=1}^{n} \log p\left(y_t|X_t\gamma^*, \Omega^*\right) - \log \pi\left(\gamma^*, \Omega^{-1*}|Y_n\right)$$

where the last term is calculated as

$$\pi\left(\Omega^{-1*}|Y_n\right) \times \pi\left(\gamma^*|Y_n, \Omega^{-1*}\right)$$

in which the first term is estimated by averaging the full-conditional Wishart density over the draws $\left\{\gamma^{(g)}\right\}_{g=1}^{G}$ from the main MCMC run

$$\pi\left(\Omega^{-1*}|Y_n\right) = \frac{1}{G}\sum_{g=1}^{G} \mathcal{W}_d\left(\Omega^{-1*}|\rho_0 + n, \left(R_0^{-1} + \sum_{t=1}^{n}\left(y_t - X_t\gamma^{(g)}\right)\left(y_t - X_t\gamma^{(g)}\right)'\right)^{-1}\right)$$

and the second term $\pi\left(\gamma^*|Y_n, \Omega^{-1*}\right)$ is available directly as

$$\pi(\gamma^*|Y_n, \Omega^{-1*}) = \mathcal{N}_{d+p}\left(\gamma^*|\hat{\gamma}^*, G_n^*\right)$$

where

$$\hat{\gamma}^* = G_n^* \left( G_0^{-1}\gamma_0 + \sum_{t=1}^{n} X_t'\Omega^{-1*} y_t \right)$$
$$G_n^* = \left( G_0^{-1} + \sum_{t=1}^{n} X_t'\Omega^{-1*} X_t \right)^{-1}$$

# 5 Prediction

Suppose we are interested in predicting the cross-section of returns $y_{n+1}$ at time $(n+1)$. The Bayes prediction density of these returns, conditioned on the data $Y_{n+1}$ and the factors $f_{n+1}$, is given by

$$p(y_{n+1}|Y_n, f_{n+1}) = \int_{\gamma,\Omega^{-1}} \mathcal{N}_d\left(y_{n+1}|X_{n+1}\gamma, \Omega\right) d\pi\left(\gamma, \Omega^{-1}|Y_n\right)$$

which is estimated by the ergodic Monte Carlo average

$$p(y_{n+1}|Y_n, f_{n+1}) = \frac{1}{G}\sum_{g=1}^{G} \mathcal{N}_d\left(y_{n+1}|X_{n+1}\gamma^{(g)}, \Omega^{(g)}\right)$$

with the MCMC draws $\left\{\gamma^{(g)}, \Omega^{(g)}\right\}$ from the posterior distribution.