UNIVERSITY OF TEXAS ARLINGTON

Bayesian Hierarchical Dynamic Factor Models

Anthony M. Thomas, Jr.
anthony.thomas3@mavs.uta.edu

Department of Mathematics
The University of Texas at Arlington

August 16, 2020

Contents

Section 1

Factor Analysis

1. **Factor Analysis** is a method that uses the covariances between a set of observed variables to described them in terms of a smaller set of unobservable variables called factors.

Bayesian Inference

Bayesian Inference can be described by three parts:

1. Build a model based on data $\boldsymbol{X}$ and parameters $\boldsymbol{\Theta}$
   - Likelihood: $p\left(\boldsymbol{X}|\boldsymbol{\Theta}\right)$
   - Prior: $p\left(\boldsymbol{\Theta}\right)$
2. Compute the posterior pdf
   - Posterior:

$$p\left(\boldsymbol{\Theta}|\boldsymbol{X}\right) = \frac{p\left(\boldsymbol{X}|\boldsymbol{\Theta}\right)p\left(\boldsymbol{\Theta}\right)}{p\left(\boldsymbol{X}\right)}$$

3. Report summaries, e.g. posterior expectations

$$\mathrm{E}\left[h(\boldsymbol{\Theta})|\boldsymbol{X}\right]$$

Section 2

Why Variational Bayesian Inference?

- Consider a model with data $\boldsymbol{X}$ and unknowns $\boldsymbol{Z}$
- The goal is to compute the posterior

$$p\left(\boldsymbol{Z}|\boldsymbol{X}\right) = \frac{p\left(\boldsymbol{X}|\boldsymbol{Z}\right) p\left(\boldsymbol{Z}\right)}{p\left(\boldsymbol{X}\right)} \tag{1}$$

  and report posterior expectations

$$\mathrm{E}\left[h(\boldsymbol{Z})|\boldsymbol{X}\right] = \int h(\boldsymbol{Z}) p\left(\boldsymbol{Z}|\boldsymbol{X}\right)\, d\boldsymbol{Z} \tag{2}$$

- The term in the denominator is called the evidence

$$p\left(\boldsymbol{X}\right) = \int p\left(\boldsymbol{X}, \boldsymbol{Z}\right)\, d\boldsymbol{Z} \tag{3}$$

Why Variational Bayesian Inference?

- For most interesting models $\boldsymbol{Z}$ is high-dimensional
- (**??**) and (**??**) would require high-dimensional integration
- (**??**), (**??**), and (**??**) are typically analytically intractable

Why Variational Bayesian Inference?

1. The gold standard has been to use MCMC algorithms
2. MCMC methods allow sampling from intractable distributions
3.

Why Variational Bayes?

, but takes too long to run. VB methods are faster.

Variational Bayes

VB turns the inference problem into an

Section 3

Classical Factor Analysis Model

The Classical (orthogonal) FA model assumes assumes the form

$$\boldsymbol{X} = \boldsymbol{\Lambda} \boldsymbol{F} + \boldsymbol{e}$$

where

1. $\boldsymbol{X} = (X_1, \ldots, X_N)^\top$ denotes the vector of observations
2. $\boldsymbol{\Lambda} = [\lambda_{nk}]_{N \times K}$ denotes the matrix of factor loadings
3. $\boldsymbol{F} = (F_1, \ldots, F_K)^\top$ denotes the vector of latent factors
4. $\boldsymbol{e} = (e_1, \ldots, e_N)^\top$ denotes the vector of latent error terms

Normal Theory Assumptions

The Normal Theory Classical FA model assumes the form

$$\boldsymbol{X} = \boldsymbol{\Lambda F} + \boldsymbol{e}$$

and adds the assumptions that

1. $\boldsymbol{F} \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{I}_K\right)$
2. $\boldsymbol{e} \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}\right)$ where $\boldsymbol{\Sigma} = \text{diag}\left(\sigma_1^2, \ldots, \sigma_N^2\right)$
3. F_k and e_n are independent for every pair k, n

Section 4

Hierarchical Dynamic Factor Model

$$\mathbf{X}_{bst} = \mathbf{\Lambda}_{H.bs}(L)\mathbf{H}_{bst} + \mathbf{e}_{Xbst} \tag{4}$$

$$\mathbf{H}_{bst} = \mathbf{\Lambda}_{G.bs}(L)\mathbf{G}_{bt} + \mathbf{e}_{Hbst} \tag{5}$$

$$\mathbf{G}_{bt} = \mathbf{\Lambda}_{F.b}(L)\mathbf{F}_{t} + \mathbf{e}_{Gbt} \tag{6}$$

$$\mathbf{\Psi}_{F}(L)\mathbf{F}_{t} = \boldsymbol{\epsilon}_{Ft}, \tag{7}$$