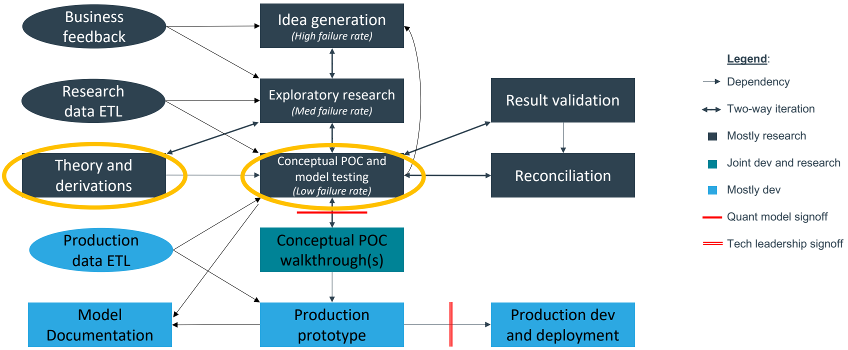
GMAM 3.0

iCapital Portfolio Analytics

December 9, 2022

The Research Process



Mean Field Variational Bayes: Motivation

- Frequentist methods fix the parameters in the data-generating process (DGP) and assume the data is generated randomly.
- Bayesian methods invert the DGP by fixing the data and inferring the likely distribution of parameters.
- Bayesian estimators are admissible in the statistical sense: given the data generating process and a loss function, no other estimator always performs better (weakly dominates) the Bayesian estimator.
- Drawbacks: priors (sometimes), tractability, and performance.

Mean Field Variational Bayes: Motivation

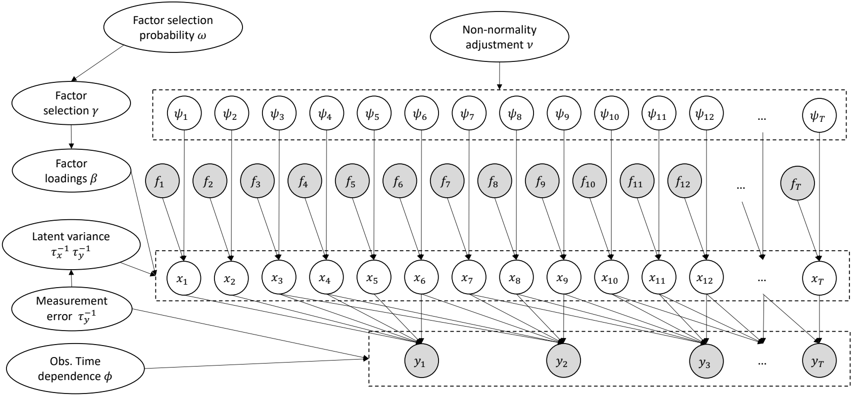
- Priors influence the results, but they also serve as a convenient and statistically rigorous approach for incorporating outside information (e.g. cross-sectional data) into the estimates.

- Tractability issues stem from what is also Bayesian model's greatest strength: every unknown parameter is viewed probabilistically with respect to the values it could take given fixed observations.

- Performance issues are specific to the solution methodology, but the most common (MCMC) requires extensive simulation in the estimation process

The Mean Field Variational Bayes (MFVB) likely provides a tractable high performance solution.

GMAM 3.0



The hierarchical structure of GMAM 3.0 provides a flexible and rigorous statistical framework for advanced portfolio analytics.

Mean Field Variational Bayes: MCMC

- Markov Chain Monte Carlo (MCMC) is a highly flexible procedure for computing the distribution of parameter estimates given the data.
- In infinite time and mild regularity conditions, MCMC provides the EXACT distribution of the posterior.
- But...it ranges from slow (Gibbs sampling with conjugate priors) to very slow (Metropolis-Hastings, other samplers).
    - A typical MCMC chain will draw from the conditional distributions well over a $100,000$ times.
    - For a model of the complexity of GMAM 3.0, that equates to tens of millions of random draws, each using different distributional parameters.
    - Smaller quantities of draws are possible, but would need careful testing of efficacy.

Mean Field Variational Bayes: MCMC

- Consider a simple MCMC:

$$\mu_i \sim p\left(\mu|\sigma^2_{i-1}, D\right) \quad (1)$$

$$\sigma^2_i \sim p\left(\sigma^2|\mu_i, D\right) \quad (2)$$

- Because the probability distributions are conditional, sequential numerical simulation is necessary to analyze the marginal

- Variational Bayesian (VB) techniques provide an alternative approach.

Mean Field Variational Bayes: Basic Idea

- The full posterior of the previously described problem is given by:

$$\mu, \sigma^2 \sim p(\mu, \sigma^2 | D)$$

- Suppose the posterior factored as follows:

$$p(\mu, \sigma^2 | D) \propto q_\mu(\mu) q_{\sigma^2}(\sigma^2)$$

- Then the econometrician could simply compute summary statistics on $q_\mu(\mu)$ and $q_{\sigma^2}(\sigma^2)$.

Mean Field Variational Bayes: Basic Idea

- Importantly, the posterior does NOT factor this way.

- VB is an approximation, the goal of which is to find the distribution $q(\cdot)$ which maximizes the accuracy of the approximation.

- In the variant employed in GMAM 3.0, $q(\cdot)$ is selected to minimize the KL divergence between the approximating distribution and the true (conditional) posterior.

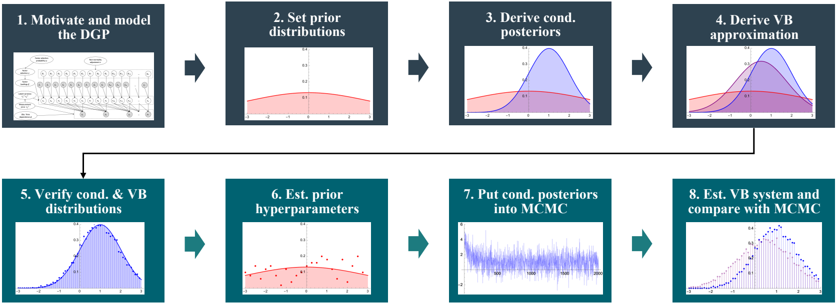
Mean Field Variational Bayes: Application to GMAM 3.0

The VB approximation to GMAM 3.0:

$$
\begin{aligned}
p\left(\Theta|D\right) \underset{\sim}{\propto} & \; q\left(\phi\right) \times q\left(\tau_y\right) \times q\left(x\right) \\
& \times q\left(\beta\right) \times q\left(\tau_x\right) \times \prod_{t \in 1:T} q\left(\psi_i\right) \\
& \times q\left(\nu\right) \times \prod_{k \in 1:K} q\left(\gamma_k\right) \times q\left(\omega\right)
\end{aligned}
$$

Variable	Type	Definition/Description
x	Local	x is a vector of latent economic returns
ϕ	Local	ϕ is the moving average window
τ_y	Global	Precision parameter for independent measurement error
β	Local	Regression coefficients of x on F
ψ	Local	ψ is a vector of precision weights for x
τ_x	Global	Precision parameter for the regression of x on F
γ	Local	Vector of variable selection indicators
ω	Global	Probability of variable selection
ν	Global	Non-normality parameter for x; DOF of posterior t distribution

GMAM 3.0: The plan



Analytical modeling

POC Development

Note: While the dependencies are strict (except for #6), the process is highly iterative: earlier steps will need rework based on findings in later steps.

GMAM 3.0: The plan

1. Model the DGP in terms of observed data and unobserved parameters
2. Assign prior distributions (usually conjugate)
3. Derive conditional posterior distributions
4. Derive approximating VB distributions
5. For each parameter, verify the conditional posterior and approximating VB distributions via numerical simulation
6. Determine a strategy for computing prior hyper parameters
7. Construct an MCMC for the conditional posterior distributions
8. Estimate the VB system and compare the accuracy to the MCMC results