Modern Bayesian Econometrics

By Hedibert Freitas Lopes
Professor of Statistics and Econometrics
INSPER Institute of Education and Research

PART 1: Modeling and computation (6h)

1. Likelihood function & prior, predictive and posterior densities
2. Model selection/averaging, posterior predictives, Bayes factors and model probabilities
3. Monte Carlo (MC) methods: MC integration & sampling importance resampling (SIR)
4. Markov chain Monte Carlo (MCMC) methods: Gibbs & Metropolis-Hastings-Green algorithms

We cover MC/MCMC computational tools to be able to talk about more recent developments. Also, we will use vector autoregressive (VAR) models as working examples, where out-of-sample forecasting and impulse response functions are natural by-products for model selection and vizualization. Most of the basic material appears in chapter 2, 3, 5 and 6 of Gamerman and Lopes (2006). More specifically, Sections 2.1-2.3, 3.1, 3.4, 3.5.1, 3.5.2, 5.1-5.4 and 6.1-6.4. The Oxford Handbook of Bayesian Econometrics chapter on "Introduction to Simulation and MCMC Methods" is also a general reference for MCMC methods. A more econometrics flavour can be found in Koop (2003), Lancaster (2004) ad Greenberg (2013).

PART 2: Time-varying parameter and sparsity modeling (6h)

Part 2a: State-space modeling

1. Introduction to normal dynamic linear modeling (NDLM)
2. Kalman filter, Kalman smoother and particle filters

Part 2b: TVP-VAR and DSGE models

1. Time-varying parameter VAR (TVP-VAR) models
2. VAR and dynamic stochastic general equilibrium (DSGE) applications

Part 2c: Static and dynamic sparsity

1. Regularizing shrinkage/selection priors: ridge, lasso, horseshoe, normal-inverse gamma
2. Dynamic sparsity

NDLM's are used as the backbone of several developments in the time-varying VAR literature and will be used to introduce well-known tools such as the *forward filtering and backward sampling* (FFBS) for full Bayesian inference. Particles filters become important when models are not linear nor Gaussian, such as in the non-linear VAR and DSGE contexts. We finish Part 2 by reviewing several of the existing sparsity strategies as well as the more recent literature on dynamic sparsity.

The main reference for Bayesian inference in NDLM and conditionally NDLM are Sections 5.5.2 and 6.5.2 of Gamerman and Lopes (2006). See also Prado and West's (2010) chapters on "Dynamic Linear

Models", "State-Space Time-Varying Autoregressive Models" and "Sequential Monte Carlo Methods for State-Space Models". For those more interested in Bayesian approaches to specific areas, I recommend reading (however superficially, at first) the following chapters of The Oxford Handbook of Bayesian Econometrics: Bayesian Methods in Microeconometrics (Li and Tobias), Bayesian Macroeconometrics (del Negro), and Bayesian Methods In Finance (Jacquier and Polson).

Particle filtering reviews can be found in several of my recent papers. In particular I recommend Lopes and Tsay (2011) Particle filters and Bayesian inference in financial econometrics, Lopes, Carvalho, Johannes and Polson (2011), Particle learning for sequential Bayesian computation (with discussion) and Carvalho, Johannes, Lopes and Polson (2010) Particle learning and smoothing. These papers can be downloaded from my professional page at www.hedibert.org.

Finally, for static sparsity I recommend these references: (i) Hans (2009) Bayesian lasso regression. Biometrika, 96, 835-845; (ii) Polson and Scott (2012) Local shrinkage rules, Lévy processes and regularized regression. JRSS-B, 74, 287-311; (iii) Tibshirani (2011) Regression shrinkage and selection via the lasso: a retrospective. JRSS-B, 73, 273-282; and (iv) Griffin and Brown (2013) Some priors for sparse regression modelling. Bayesian Analysis, 8, 691-702. Chapter 6 of James, Witten, Hastie and Tibshirani (2013) also reviews the non-Bayesian regularization schemes.

PART 3: Statistical and machine learning (3h)

1. Classification and regression trees (CART)
2. Ensemble methods: Random forests, bagging, boosting
3. Bayesian CART and Bayesian additive regression trees (BART)
4. Modern neural net (NN) models (deep learning)
5. Modeling document/text data via Latent Dirichlet Allocation (LDA)

Chapter 8 (Tree-Based Methods) of James, Witten, Hastie and Tibshirani (2013) reviews CART and random forests, as well as bagging and boosting techniques. Bayesian CART and BART main references are the following: (i) Chipman, George and McCulloch (1998) Bayesian CART model search. JASA, 93, 935-948; (ii) Denison, Mallick and Smith (1998) A Bayesian CART algorithm. Biometrika, 85, 363-377; (iii) Wu, Tjelmeland and West (2007) Bayesian CART: Prior specification and posterior simulation. JCGS, 16(1), 44-66; (iv) Chipman, George and McCulloch (2000) Hierarchical priors for Bayesian CART shrinkage. Statistics and Computing, 10, 17-24; (v) Chipman, George and McCulloch (2002) Bayesian treed models. Machine Learning, 48, 299-320; (vi) Chipman, George and McCulloch (2010) BART: Bayesian Additive Regression Trees. AOAS, 4, 266-298.

Chapter 11 of Hastie, Tibshirani and Friedman (2015) introduces a comprehensive overview neural network models, with a Bayesian perspective provided by Polson and Sokolov's (2017) Bayesian Analysis paper "Deep Learning: A Bayesian Perspective" (Volume 12, Number 4, 1275-1304). Additional references in asset pricing are Feng, Polson and Xu (2019) Deep Learning in Characteristics-Sorted Factor Models. See also, Gu, Kelly and Xiu (2019) Empirical Asset Pricing via Machine Learning.

For text modeling, one of the most important papers is Blei, Ng and Jordan (2003) Latent Dirichlet Allocation, *Journal of Machine Learning Research*, 3, 993-1022. Correlated, dynamic and probabilistic topic modeling are introduced, respectively, in Blei and Lafferty (2006) Correlated topic models. Advances in Neural Information Processing Systems (NIPS), Blei and Lafferty (2006) Dynamic topic models. In International Conference on Machine Learning (ICML), 2006, and Blei (2012) Probabilistic topic models. Communications of the Association for Computing Machinery (AMC), Vol. 55, No. 4. Two additional more recent references are Alghamdi and Alfalqi (2015) A survey of topic modeling in text mining. International Journal of Advanced Computer Science and Applications, Vol. 6, No. 1, 147-153, and Airoldi and Bischof (2016) Improving and evaluating topic models and other models of text. JASA, 111, 1381-1403. See also Gentzkow, Kelly and Taddy (2017) Text as Data. NBER Working Paper No. 23276 and Athey and Imbens (2019) Machine Learning Methods Economists Should Know About.

Basic bibliography

- Gamerman and Lopes (2006) *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference (2nd edition)*, Taylor & Francis CRC Press - http://www.dme.ufrj.br/mcmc
- Prado and West (2010) *Time Series: Modeling, Computation, and Inference*, Taylor & Francis CRC Press.
- Geweke, Koop and Van Dijk (2011) *The Oxford Handbook of Bayesian Econometrics*, Oxford University Press.
- James, Witten, Hastie and Tibshirani (2013) *An Introduction to Statistical Learning with Applications in R*, Springer - http://faculty.marshall.usc.edu/gareth-james/ISL/index.html

Additional basic bibliography

- Koop (2003) Bayesian Econometrics
- Greenberg (2013) Introduction to Bayesian Econometrics
- Lancaster (2004) Introduction to Modern Bayesian Econometrics
- Murphy (2012) Machine Learning: A Probabilistic Perspective
- Barber (2012) Bayesian Reasoning and Machine Learning
- Hastie, Tibshirani and Friedman (2001) The Elements of Statistical Learning
- Bishop (2006) Pattern Recognition and Machine Learning
- Hastie, Tibshirani and Wainwright (2015) Statistical Learning with Sparsity

Machine learning papers in economics

- Athey and Imbens (2019) Machine Learning Methods Economists Should Know About. *Annual Review of Economics*, Volume 11, pages 685-725.
- Bianchi, Büchner and Tamoni (2019) Bond Risk Premia with Machine Learning. Version of February 26, 2019.
- Chakraborty and Joseph (2017) Machine learning at central banks. Bank of England working papers 674.
- Feng, Polson and Xu (2019) Deep Learning in Characteristics-Sorted Factor Models.
- Gu, Kelly and Xiu (2019) Empirical Asset Pricing via Machine Learning. Version of September 13, 2019.
- Giannone, Lenza and Primiceri (2018) Economic Predictions With Big Data: The Illusion of Sparsity. Working Paper, Northwestern University.
- Gu, Kelly and Xiu (2019) Autoencoder Asset Pricing Models. Version of September 30, 2019.
- Hall (2018) Machine Learning Approaches to Macroeconomic Forecasting. Economic Review, Federal Reserve Bank of Kansas City, issue Q IV, pages 63-81.
- Heaton, Polson and Witte (2017) Deep learning for finance: deep portfolios. *ASMBI*, 33, 3-12.
- Korobilis and Pettenuzzo (2019) Machine learning econometrics: Bayesian algorithms and methods.
- Makridakis, Spiliotis and Assimakopoulos (2018) Statistical and Machine Learning Forecasting Methods: Concerns and Ways Forward. *PLoS One*, 13, e0194889.
- Medeiros *et al* (2019) Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods. *JBES*.
- Medeiros and Vasconcelos (2016) Forecasting Macroeconomic Variables in Data-Rich Environments. *Economics Letters*, 138, 50-52.
- Mullainathan and Spiess (2017) Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31, 87-106.
- Sirignano, Sadhwani and Giesecke (2018) Deep Learning for Mortgage Risk. Version of November 20, 2018.
- Wager and Athey (2018) Estimation and inference of heterogeneous treatment effects using random forests *JASA*, 113, 1228-1242.

A few recent conferences on machine learning in economics

- 10th ECB Workshop on Forecasting Techniques: Economic Forecasting with Large Datasets
  The European Central Bank (ECB), Frankfurt, June 18-19 2018.
- Machine Learning and Econometrics
  A cemmap Workshop jointly funded by The Alan Turing Institute, June 8-9 2018.
- Econometrics in the Castle: Machine Learning in Economics and Econometrics
  Munich, Germany May 29-30 2018.
- 2018 NBER-NSF Seminar on Bayesian Inference in Eonometrics and Statistics (SBIES)
  Department of Economics, Stanford University, May 25-26 2018.
- Big Data in Predictive Dynamic Econometric Modeling
  University of Pennsylvania, May 18-19 2017.
- Drawing Causal Inference from Big Data
  National Academy of Sciences , March 26-27 2015.

PART 1

Bayesian ingredients

1. Example i. Normal model and normal prior
2. Turning the Bayesian crank
   - Posterior and predictive distributions
   - Posterior predictive distribution
   - Sequential Bayes
3. Example ii. Simple linear regression
4. Example iii. Stochastic volatility model
5. Example iv. Multiple linear regression
6. Bayesian model criticism Posterior odds
   - Bayes factor
   - Bayesian Model Averaging
   - Posterior predictive criterion
   - Deviance Information Criterion

Monte Carlo methods

1. A bit of history
2. Monte Carlo methods
3. MC integration
4. MC via importance sampling
5. Rejection method
6. Sampling importance resampling (SIR) method
7. Examples
   - 3-component mixture of Gaussians
   - 2-component mixture of Gaussians

MCMC methods

1. Historical facts
2. MH algorithms
   - Special cases
   - Random walk Metropolis
3. Gibbs sampler

Bayesian ingredients and computation

- Bayesian paradigm
  - Example 1: Is Diego ill?
  - Example 2: Gaussian measurement error
- Bayesian computation: MC and MCMC methods
- Monte Carlo integration Monte Carlo simulation

- Gibbs sampler Metropolis-Hastings algorithm

More examples

- Bayesian AR(2)
- Bayesian CAPM
- Bayesian GARCH
- Bayesian VAR

PART 2

Multiple linear regression

- Multiple linear regression
  - Simplest linear regression model
  - `houseprice` dataset
  - R^2, R^2_{adj}, C_p, AIC and BIC
  - R package `regsubsets`
  - `Credit` dataset
- Shrinkage-L2, Ridge Regression
  - `Hitters`dataset
  - Constrained minimization
  - Karush Kuhn Tucker (KKT) conditions
- Shrinkage-L1: The LASSO
  - Soft threshholding function
  - Cyclic Coordinate Descent
  - R package `glmnet`
- More on regularization
  - Elastic net
  - Normal-gamma prior
  - Horseshoe prior
  - R package `bayeslm`
  - Simulation exercise

Bayesian regularization

1. Least absolute shrinkage and selection operator
2. Bridge regression and elastic net
3. Bayesian Lasso
4. Spike and slab variable selection
5. Horseshoe prior
6. Normal-gamma prior
7. Support vector machines

8. Sparse factor models
    - Case 1: Constructing Economically Justified Aggregates
    - Case 2: Exchange rates

Dynamic linear models

- 1st order DLM
  - n-variate normal
  - The Kalman filter
  - The Kalman smoother
  - Integrating out states x^n
  - MCMC scheme
  - Lessons
- Dynamic linear models (DLMs)
  - Linear growth model
  - Sequential inference
  - Smoothing
  - The FFBS
  - Individual sampling
  - Joint sampling
- Dynamic generalized linear model
  - Example: Advertising awareness

Stochastic volatility model

Particle filter: state learning

Particle filter: state and parameter learning

PART 3

Bayesian classification and regression trees (CART)

Bayesian additive regression trees (BART)

Random forest

Boosting

Deep neural nets (DNN)

Latent Dirichlet allocation (LDA)