Advanced Bayesian Econometrics 2025

Professor: Hedibert Freitas Lopes – www.hedibert.org

Objective: The end of the course goal is to allow the student to critically decide between a Bayesian, a frequentist or Bayesian-frequentist compromise when facing real world problems in the fields of micro- and macro-econometrics and finance, as well as in quantitative marketing, strategy and business administration. With this end in mind, we will visit well known Bayesian issues, such as prior specification and model comparison and model averaging, but also study regularization via Bayesian LASSO, Spike-and-Slab and related schemes, "small n, large p" issues, Bayesian statistical learning via additive regression trees, random forests, large-scale VAR and (dynamic) factor models.
Course description: Basic ingredients: prior, posterior, and predictive distributions, sequential Bayes, conjugate analysis, exchangeability, principles of data reduction and decision theory. Model criticism: Bayes factor, computing marginal likelihoods, Savage-Dickey ratio, reversible jump MCMC, Bayesian model averaging and deviance information criterion. Modern computation via (Markov chain) Monte Carlo methods: Monte Carlo integration, sampling-importance resampling, Gibbs sampler, Metropolis-Hastings algorithms. Mixture models, Hierarchical models, Bayesian regularization, Instrumental variables modeling, Large-scale (sparse) factor modeling, Bayesian additive regression trees (BART) and related topics, Dynamic models, Sequential Monte Carlo algorithms, Bayesian methods in microeconometrics, macroeconometrics, marketing and finance.
Part I Bayesian ingredients: i) Inference: likelihood, prior, predictive and posterior distributions; ii) Model criticism: Marginal likelihoods, Bayes factor, model averaging and decision theory; and iii) Computation: An introduction (Markov chain and sequencial) Monte Carlo methods.
Part II Multivariate models: i) Large-scale vector autoregressive models; ii) Factor models and other dimension reduction models; and iii) Time-varying high-dimensional covariance models.
Part III Modern Bayesian statistical learning: i) Mixture models and the Dirichlet process: handling non-Gaussian models; ii) Regularization: sparsity via shrinkage and variable selection; iii) Large vector-autoregressive and factor models: combining sparsity and parsimony; iv) Classification and support vector machines; v) Regression trees and random forests; and vi) Latent Dirichlet allocation: Text as data, text mining.

Paper presentations: i) *GP-VAR and Macroeconomic Uncertainty; ii) How Polarized are Citizens? Measuring Ideology from the Ground-Up; iii) Minnesota BART; iv) Text as Data; and v) What events matter for exchange rate volatility?*

Homework assignments
HW1: Repeat the AR(1) problem (Example 2 from Class 1), but this time with w(t) iid Student's t with nu=4 degrees of freedom. a) Plot both prior and posterior in the same figure. b) Which model is better, as far as Bayes Factor is concerned? c) Plot p(y(n+1)|y(1),…,y(n)) for w(t) ~ normal. d) Plot p(y(n+1)|y(1),…,y(n)) for w(t) ~ Student's t.
HW2: In my notes about the Monte Carlo (MC) methods, we illustrate the implementation of SIR where the target distribution is a three-component bivariate mixture of normals (page 26). Let us compare: a) Proposal A: bivariate normal proposal density presented in the notes (page 29); b) Proposal B: bivariate Student's t with the same location, 4 degrees of freedom and scale that replaces the multiplier 9 in the covariance matrix of Proposal A with the multiplier 4.5; c) Bonus: make the multiplier equal to 1 and play around with the number of degrees of freedom, say 1, 2, 3, 5, 10; d) Comment your findings.
HW4: Revisit the "Bayesian hierarchical linear regression (Class 4 – Example 6 below), by assuming now making beta[1]=beta[2]=…=beta[I]=beta, where the prior for beta is now Gaussian with mean 0.5 and variance 1.0. Repeat the change, but now let alpha[1]=…alpha[I]=alpha, where the prior for alpha is Gaussian with mean -4.0 and variance 25.0. Compare the models. In particular, compare the posterior distribution of the error variance, sigma^2 for the different hierarchical models.
HW3 + **HW5** + **HW6**

Material developed in class

- Example 1: iid Bernoulli trials & two alternative priors for proportion parameter
- Example 2: AR(1) model with normal errors (homework: Student's t errors)
- Example 3: Bivariate Student's t – learning correlation and degrees of freedom
- Example 4: Data augmentation/Gibbs sampler – Linkage example
- Example 5: Data augmentation/Gibbs sampler – Mean of Student's t data
- Example 6: Bayesian hierarchical linear regression – Gibbs Sampler
- Example 7: Bayesian hierarchical Beta-Binomial regression
- Example 8: Zero-inflated Poisson data
- HMC & STAN – Nonlinear regression illustration (R script + STAN script)
- HMC & STAN – AR(1) model with Student's t errors
- HMC & STAN – GARCH(1,1) model with Student's t errors
- HMC & STAN – SV-AR(1) model with Gaussian or Student's t errors
- Factor analysis (data)

LECTURE NOTES

PART I: Bayesian ingredients

- Basic Bayes
- Exchangeability
- Principles of data reduction
- More on estimators
- Decision theory (Nuisance parameters + travel insurance example)
- Introdução à Teoria da Decisão – by Victor Fossaluza (IME-USP)
- James-Stein estimator (from C.P.Robert's *The Bayesian Choice, 2nd edition*)
- Bayesian model criticism (pages 1-6 & 32-34)

Additional reading material:

- Gamerman and Lopes (2006) - Chapter 2 - Compact, but easy to read.
- Migon, Gamerman and Louzada (2014) - Chapters 2 to 4 - Classical and Bayesian inference.
- Gelman et al. (2013) - Chapters 1 and 2 - Application-oriented.
- Berger (1985) - Chapter 4 (Sections 4.1-4.4) - More technical.
- Parmigiani and Inoue (2009) *Decision Theory: Principles and Approaches* (Table of contentes)
- van de Schoot *et al.* Bayesian statistics and modelling. *Nat Rev Methods Primers*, 1 (2021).
- Discussion about p-values

PART II: Bayesian Computation

- Monte Carlo (MC) methods
- Markov chain: a brief review
- Markov chain Monte Carlo (MCMC) algorithms
- MC and MCMC: Key References
- More on Bayesian model criticism
- Hamiltonian Monte Carlo: A toy example
- Stan/rstan for posterior inference: Hamiltonian MC (HMC) methods
- Banana shaped bivariate target: MH vs HMC
- Bayesian hierarchical modeling – the Beta-Binomial case

PART III: Bayesian Learning

1. Modeling with mixtures of distributions
    - Finite mixture of distributions
    - Univariate mixture of normals: MCMC and EM algorithms
    - Multivariate mixture of distributions
    - AR(1) model with Markov Switching intercept
    - Linear regression with mixture of normal errors
    - How Many Data Clusters Are in the Galaxy Data Set? + The telescope sampling

2. Fundamentos de Aprendizagem Estatística + R code + MC exercise

3. Multiple linear regression: selection, shrinkage, sparsity
    - Motorcycle example
    - Slides from the 2015 School of Time Series and Econometrics tutorial
    - Hahn, He and Lopes (2018) Gaussian linear regression with arbitrary sparsity + slides of a talk
    - R package bayeslm
    - Fava and Lopes (2021) The illusion of the illusion of sparsity + slides of a talk + UFPE webinar
    - van Erp et al. (2019) Shrinkage priors for Bayesian penalized regression, JMP, 89, 31-50.
    - Michael Betancourt's Bayes Sparse Regression (stan/rstan example)

4. Classification: logistic regression and discriminant analysis
    - Sparse logistic regression for the spam/ham dataset (data)
    - Marketing campaigns of a Portuguese banking institution (data)
    - Sparse logistic regression: comparison of regularization and Bayesian implementations
    - Gelman et al. (2008) A weakly informative default prior distribution for logistic and other regression models, *AOAS*, 2(4), 1360-1383.
    - Polson et al. (2013) Bayesian Inference for logistic models using Pólya-Gamma latent variables, *JASA*, 108, 1339-1349.

5. Multivariate models and dimension reduction

- Bayesian factor analysis (BFA)
- BFA in R: a simple illustration (additional routines)
- Factor models: an annotated bibliography (Lopes, 2003) + review chapter (Lopes, 2014)
- Principal components analysis (PCA), PCA-based and FA-based regressions
- Probabilistic PCA + Rcode + data

6. Classification and regression trees (CART)
   - Machine Learning course by Paulo Orenstein from IMPA.
   - Chapter 12 and Chapter 13 are about *Tree-based methods*.
   - Chapter 15 and Chapter 16 are about *Deep learning*.

7. Bayesian CART

8. Bootstrap aggregating (bagging)
   - Bootstrap illustration
   - CART step-by-step for the Boston housing data
   - CART for the motorcycle data

9. Bayesian additive regression trees (BART)
   - Example 1: ICU data: CART, BART and random forest (R code)
   - Example 2: Stock and Watson's (2002) macro data (data)
   - More examples: Four BART applications & 2 reviews + cute CART trees and 3D plots
   - More recent references

10. Latent Dirichlet Allocation (LDA)
    - Twitter + BBC + Dickens, Wells, Verne or Austen? + The cat in the hat by Dr. Seuss

11. Neural Networks

Complementary material to PART III

- Boosting (weak/stronger learners)
- Random forests
- Bayesian instrumental variables
- General linear and hierarchical models
- Limited dependent variable models
- Spatial models
- P.Richard Hahn's top 25 books on Statistics, Causal Inference, Statistical Computing, Machine Learning and Data Science

Bibliography: Bayesian econometrics

1.Zellner (1971) An Introduction to Bayesian Inference in Econometrics
2.Goel and Iyngar (1992) Bayesian Analysis in Statistics and Econometrics
3.West and Harrison (1997) Bayesian Forecasting and Dynamic Models (2nd edition)
4.Dorfman (1997) Bayesian Economics Through Numerical Methods
5.Bauwens, Lubrano and Richard (2000) Bayesian Inference in Dynamic Econometric Models
6.Koop (2003) Bayesian Econometrics
7.Geweke (2005) Contemporary Bayesian Econometrics and Statistics
8.Lancaster (2004) Introduction to Modern Bayesian Econometrics
9.Rossi, Allenby and McCulloch (2005) Bayesian Statistics and Marketing
10.Prado and West (2010) Time Series: Modeling, Computation and Inference
11.Geweke, Koop and Van Dijk (2011) The Oxford Handbook of Bayesian Econometrics
12.Greenberg (2013) Introduction to Bayesian Econometrics
13.Herbst and Schorfheide (2015) Bayesian Estimation of DSGE Models
14.Chan, Koop, Poirier and Tobias (2019) Bayesian Econometric Methods (2nd edition)
15.Broemeling (2019) Bayesian Analysis of Time Series
16.Bernardi, Grassi and Ravazzolo (2020) Bayesian Econometrics

Bibliography: Bayesian statistics

1. Berger (1985) Statistical Decision Theory and Bayesian Analysis
2. Bernardo and Smith (2000) Bayesian Theory
3. Gelman and Hill (2006) Data Analysis Using Regression and Multilevel/Hierarchical Models
4. Robert (2007) The Bayesian Choice
5. Hoff (2009) A First Course in Bayesian Statistical Methods
6. Carlin and Louis (2009) Bayesian Methods for Data Analysis (3rd edition)
7. Gelman, Carlin, Stern, Dunson, Vehtari and Rubin (2016) Bayesian Data Analysis
8. Migon, Gamerman and Louzada (2015) Statistical Inference: An Integrated Approach (2nd edition)
9. Reich and Ghosh (2019) Bayesian Statistical Methods
10. Held and Sabanes-Bove (2020) Likelihood and Bayesian Inference

Bibliography: Bayesian computation

1. Gilks, Richardson and Spiegelhalter (1995) Markov Chain Monte Carlo in Practice
2. Doucet, de Freitas and Gordon (2001) Sequential Monte Carlo Methods in Practice
3. Robert and Casella (2004) Monte Carlo Statistical Methods (2nd edition)
4. Gamerman and Lopes (2006) MCMC: Stochastic Simulation for Bayesian Inference, Second Edition
5. Marin and Robert (2007) Bayesian Core: A Practical Approach to Computational Bayesian Statistics
6. Albert (2009) Bayesian Computation with R
7. Brooks, Gelman, Jones and Meng (2011) Handbook of Markov Chain Monte Carlo
8. Givens and Hoeting (2012) Computational Statistics (2nd edition)
9. Marin and Robert (2014) Bayesian Essentials with R (complete solution manual)
10. Turkman, Paulino and Mueller (2019) Computational Bayesian Statistics: An Introduction
11. McElreath (2020) Statistical Rethinking: A Bayesian course with Examples in R and STAN
12. Chopin and Papaspiliopoulos (2020) An Introduction to Sequential Monte Carlo

Bibliography: Bayesian statistical learning

1. Bishop (2006) Pattern Recognition and Machine Learning
2. Hastie, Tibshirani and Friedman (2008) The Elements of Statistical Learning, 2nd edition
3. Murphy (2012) Machine Learning: A Probabilistic Perspective
4. Barber (2012) Bayesian Reasoning and Machine Learning
5. James, Witten, Hastie and Tibshirani (2013) An Introduction to Statistical Learning
6. Hastie, Tibshirani and Wainwright (2015) Statistical Learning with Sparsity
7. Efron and Hastie (2016) Computer Age Statistical Inference: Algorithms, Evidence and Data Science
8. Fernandez and Marques (2018) Data Science, Marketing and Business
9. Izbicki & Santos (2020) Aprendizado de máquina: uma abordagem estatística

Bibliography: Monte Carlo integration, simulation & Markov chain Monte Carlo

1. Metropolis and Ulam (1949) The Monte Carlo method. JASA, 44, 335-341.
2. Metropolis, Rosenbluth, Rosenbluth, Teller and Teller (1953) Equation of state calculations by fast computing machines. Journal of Chemical Physics, 21, 2087-1092.
3. Hastings (1970) MC sampling methods using Markov chains and their applications. Biometrika, 57, 97-109.
4. Peskun (1973) Optimum Monte Carlo sampling using Markov chains. Biometrika, 60, 607-612.
5. Besag (1974) Spatial Interaction and the Statistical Analysis of Lattice Systems. JRSS-B, 36, 192-236.
6. Kirkpatrick, Gelatt and Vecchi (1983) Optimization by Simulated Annealing. Science, 220 (4598), 671-680.
7. Geman and Geman (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Trans. Pattern Analysis and Machine Intelligence, 6, 721-741.
8. Pearl (1987) Evidential reasoning using stochastic simulation of causal models. AI, 32, 245-257.
9. Tanner & Wong (1987) The Calculation of Posterior Distributions by Data Augmentation. JASA, 82, 528-540.
10. Geweke (1989) Bayesian Inference in Econometric Models Using MC Integration. Econometrica, 57, 1317-39.
11. Gelfand and Smith (1990) Sampling-Based Approaches to Calculating Marginal Densities. JASA, 85, 398-409.
12. Casella and George (1992) Explaining the Gibbs Sampler. The American Statistician, 46, 167-174.
13. Gilks and Wild (1992) Adaptive Rejection Sampling for Gibbs Sampling. Applied Statistics, 41, 337-348.
14. Smith&Gelfand (1992) Bayesian Statistics without Tears: A Sampling-Resampling Perspective.TAS,46,84-88.
15. Chib and Greenberg (1995) Understanding the Metropolis-Hastings algorithm. TAS, 49, 327-335.