HOW TO READ THIS EXEMPLAR: THESIS BLUEPRINT 2: PHD DOCTORAL ARCHETYPE

Novel Adaptive MCMC Samplers & High-Dimensional Bayesian Inference

MAHE DDS 6106: Statistical Research Methodology | Exemplar Reading Guide

· Classification: Instructional Doctoral Dissertation Teaching Model (Original Contribution)

·· Reading Time: 25-30 mins | · Difficulty: · Advanced | · Audience: Doctoral Candidates & Advanced M.Sc. Students

? Why Selected: Emphasizes novel theoretical contribution, mathematical proof, multiple studies, and academic defense.

STRUCTURAL COLOR LEGEND & ANNOTATION GUIDE

- Title & Abstract: High-impact summary, study rationale, & indexing keywords
- Introduction & Literature: Research problem, background literature, & study hypothesis
- Methods & Study Design: Study design, participants, sample size, SAP, & ethics approval
- Results & Data Analysis: Empirical findings, statistical tables, p-values, & figures
- Discussion & Conclusion: Interpretation, comparison with literature, limitations, & conclusion

·· THESIS STRUCTURAL ROADMAP

· MILESTONE ICONS TO LOCATE

· Theoretical Gap
Sub-optimal acceptance in high dimensions

· Novel Contribution
Adaptive covariance proposal scaling

· Mathematical Proof
Ergodicity & stationary distribution proof

· Simulation Study
Synthetic benchmark comparison (n = 10,000)

· Real Data Test
50-gene Bayesian transcriptomic inference

· Future Research
Parallel GPU chain scaling algorithms

· HOW A THESIS BECOMES A JOURNAL PAPER (IMRaD)

A PhD thesis typically generates 3 to 4 distinct journal articles! Chapter 3 & 4 form a theoretical methodology paper (JSS/Biometrika), while Chapter 5 & 6 form an applied biostatistics paper.

Full Monograph (150-200 Pages)

Contains literature review, raw tables, complete proofs, & code

· Select One Primary Question

Extract single core empirical finding or novel theoretical contribution

· Condense to 8-10 Pages (IMRaD)

Introduction (Ch1) -> Methods (Ch3) -> Results (Ch4) -> Discussion (Ch5)

· Submit to Peer-Reviewed Journal

Peer review, revision, and open-access journal publication

-> Turn page to inspect the annotated exemplar!

NOVEL ADAPTIVE MCMC SAMPLERS & HIGH-DIMENSIONAL BAYESIAN

A Dissertation Presented for the Degree of Thesis Blueprint 2: PhD Doctoral Archetype

Department of Applied Statistics and Data Science
Prasanna School of Public Health, MAHE, Manipal

DECLARATION & CERTIFICATE

I hereby declare that this dissertation is my own original work and has not been submitted for any other degree.

ABSTRACT

ABSTRACT SUMMARY:
This dissertation presents a comprehensive investigation into novel adaptive mcmc samplers & high-dimensional bayesian inference. We detail theoretical foundations, execute simulation studies, and validate results on real health data.

CHAPTER 1: GENERAL INTRODUCTION

THEORETICAL MOTIVATION:
High-dimensional Bayesian posterior inference presents severe computational bottlenecks for standard Random-Walk Metropolis-Hastings algorithms due to covariance mismatch.

DOCTORAL RESEARCH OBJECTIVES:
1. Formulate an adaptive proposal covariance scheme.
2. Prove ergodicity under diminishing adaptation conditions.
3. Implement open-source R algorithms.

CHAPTER 3: METHODOLOGICAL FORMULATION

PROPOSED ADAPTIVE SAMPLER:
We propose the Haario-scaled Adaptive Metropolis sampler where proposal covariance S_t updates dynamically using historical chain state history.

MATHEMATICAL DERIVATION:
S_t = s_d * (Cov(X_0, ..., X_{t-1}) + eps * I_d), where s_d = 2.4^2 / d is the optimal scaling constant.

CHAPTER 4: THEORETICAL ERGODICITY PROOF

THEOREM 1 (ERGODIC CONVERGENCE):
Under geometric ergodicity of the target distribution and diminishing adaptation limits lim_{t -> inf} ||S_t - S_{t-1}|| = 0, the empirical distribution converges weakly to target density pi(x).

PROOF OUTLINE:
Standard martingale transition kernel decomposition following Roberts & Rosenthal (2007).

CHAPTER 5: MONTE CARLO SIMULATION EXPERIMENTS

SIMULATION BENCHMARK:
Evaluated across 100-dimensional Gaussian, multimodal mixture, and heavy-tailed Student-t targets.

RESULTS:
The proposed adaptive sampler achieved an effective sample size (ESS) 4.2x higher than standard Random-Walk Metropolis.

INTERACTIVE STUDENT CHECKLIST: CAN YOU LOCATE...?

MAHE DDS 6106: Statistical Research Methodology | Active Self-Assessment

Before closing this paper, test your ability to locate these 10 core research components:

· 1. Research Problem & Gap ·

Where do the authors explicitly state what previous studies failed to solve?

· 2. Research Question & Objectives ·

Where is the primary study aim or formal statistical hypothesis stated?

· 3. Study Design & Target Population ·

Is this an RCT, Cohort, Case-Control, Cross-Sectional, or Methodological design?

· 4. Variables & Primary Outcome ·

What is the primary dependent variable / clinical endpoint?

· 5. Sample Size Calculation ·

Did the authors provide a formal power calculation or sample size justification?

· 6. Statistical Analysis Plan (SAP) & Software ·

Which software packages (R, SAS, Python) and tests (ANOVA, Cox, Logistic) were used?

· 7. Primary Empirical Finding & Effect Size ·

Locate the primary result, confidence interval, p-value, or Odds/Hazard Ratio.

· 8. Study Limitations & Confounding ··

Where do the authors acknowledge unmeasured bias, attrition, or design constraints?

· 9. Ethics Approval & IRB Number ·

Locate the institutional ethics committee protocol approval code.

· 10. Funding & Conflict of Interest ·

Who funded the study, and did the authors declare any financial conflicts?

· Completed this checklist? Excellent work mastering scientific paper navigation!