OUKA
The University of Osaka
Institutional Knowledge Archive

Title	ベイズ計量経済分析の手法とその応用
Author(s)	和合, 肇
Citation	大阪大学, 1999, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3155736
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

①

ベイズ計量経済分析の手法とその応用

新潟大学経済学部
和合　肇

序　　文

本書は，最近多くの実証研究で用いられるようになってきたベイズ統計学の計量経済分析への応用に関して，その手法における発展と応用分析例をいくつかまとめたものである．ベイズ統計学の実証分析への有効性についてはすでに認められているが，経済問題に関する古典的な方法による実証研究に比べてまだそれほど利用されていない．これは，実際に実証研究に携わっている研究者のベイズ統計は難しいという誤解に基づくものである．古典的な方法を利用して実証分析を行っている研究者にとっては，多くの使いやすいソフトウェアが存在するのに対し，ベイズ分析ではまだこのようなソフトウェアは少ない．しかし，最近の計算技術の発展によって，実用性の問題は除かれつつあるといえよう．さらに，研究者間でのコミュニケーションという面では，古典的な方法を利用する世界と比べてかなり進んでいるといえるかもしれない．現在，計算部分を共通の資産としてライブラリー化し，それぞれの研究者の専門的な知識をもとにいろいろな実験を可能にし，その結果を表示，検討，改訂するシステムが構築され，運用されている．

1960年代のはじめには，ベイズの方法を利用した計量経済分析が急激に進展した．この時代の研究は計量経済的な問題を解決するのにベイズの方法を開発し，適用することが中心であった．これらには科学上の一般的な問題が含まれる：すなわち，経済現象を記述し，理解し，モデル化し，経済理論によって示唆される仮説を検定し，予測に対する計量経済モデル化と手法を用いてさまざまな実際的な経済問題を解決する．これ以前には，ほとんどの計量分析は非ベイズ的な方法を用いて行われていた．1960年代にはベイズ的な計量経済学的研究が急速に発展し，1970年にはベイズ計量経済学のNBER-NSFセミナーの最初の会合がシカゴ大学で開催されている．初期の研究段階では，これらのベイズ的な方法は実際の計量経済学上の実証研究に適用する際には，計算問題，特に多重積分の計算が不可能であったこともあり，ごく小規模の研究で行われていたにすぎない．

ごく最近までベイズ計量経済学の応用は，主として理論面への貢献に関心がある人々によって行われており，真にベイズ的である応用実証研究は多くはない．この理由として，ベイズ計量経済分析は尤度関数と事前分布が必要であり，特に

主観的な事前分布ということに対して多少抵抗があり, この事前分布を用いることが, 客観性を重要視する古典的な研究者から同意されなかった. またほとんどの事後モーメントは, 結果を得るのに必要な多重積分が評価できないために, 求めるのが困難な場合が多いことが考えられる. このように応用研究が少ないということは, 最も単純な場合を除いて, ベイズ手法は実際に計算を行うには難しすぎるという認識によるものである. しかしながら, ハード面での計算力の最近の急速な発達は, 計算能力の発達を可能にした理論的な発展と相俟って, ベイズ手法が適用可能な問題の幅を急激に広げている. しかしながら, ベイズ的な計算手法を用いている多くの論文は, 統計的な性質や特別な数学的な知識を利用しているので, 応用計量経済分析を行っている人々にはまだあまり馴染みがない.

最近では多くの分野での実証研究問題にベイズ法を利用した例が増えてきている. それらの応用例, 研究動向などは1992年に設立された国際ベイズ分析学会(ISBA)や, 1979年以来4年ごとに開催され, 1998年には第6回目会合を行ったベイズ統計学に関する国際バレンシア研究集会, それに1996年に設立された米国統計学会ベイズ統計部会の年次大会などで発表された数多くの論文やProceedingsに見られる.

本書はベイズ分析が実際に役立つ方法を示すことを目的としている. そこで, なぜベイズ的な方法を採用するかという理由については, ベイズ的な方法と古典的な方法を比較して議論した優れたサーベイが他にあるので, あまり深くは立ち入らない. むしろ実証研究にベイズ手法をどのように適用するかに議論を集中する. 全体の目的は, ほとんどあるいは特別な知識を持っていない研究者に, 応用実証研究におけるベイズ手法の有用性を認識してもらうことにある.

本書の構成は次の通りである. 本書は第I部と第II部に分けられ, 第I部では, ベイズ計量経済分析と計算問題を中心として述べる. この部分は, 日本統計学会誌, Vol.28に「ベイズ計量経済分析における最近の発展」として掲載されている. ここではベイズ流統計学を計量経済分析に適用する方法と, 実証分析を行う場合に必要とされる計算問題について, 特に最近多く用いられ多くの研究で有効性が認められている, マルコフ連鎖モンテカルロ法と呼ばれる方法について述べる.

第1章では, ベイズ計量経済分析について簡単にサーベイする. ベイズ的な手

法を実証研究に応用することに焦点を合わせ, ベイズ実証分析の優れた点と問題点について述べ, ベイズ応用実証分析の例として, 単位根モデルとEBA(極値境界分析) を中心として解説する.

第2章では, ベイズ流の計量研究を行う際に用いられるいくつかの計算方法を中心として, その最近までの発展と事後分布のマルコフ連鎖モンテカルロ・シミュレーションをいくつかの実証モデルに適用する方法について述べる. まずベイズ計量経済分析で用いる基本的な計算方法である正規近似, 数値積分, ラプラス近似についてその概略を説明した後, 最近多く用いられるようになったモンテカルロ積分について, 単純な場合とインポータンス関数を用いた場合について, その使用例とともに解説する. さらに, もっと複雑な問題に用いることができるマルコフ連鎖モンテカルロ・シミュレーションについて多少詳細に述べる. 一般的なベイズ推定, マルコフ連鎖についてふれた後, 代表的な方法であるMetropolis-Hastingアルゴリズムと, よく使用されるGibbsサンプリング法, それにデータ拡大アルゴリズムについて解説する. 最後に, MCMCシミュレーションを計量経済分析にどのように適用するかが応用例とともに述べられる. また, 付録でベイズ実証分析を行う際に役立つコンピュータ・パッケージについて述べる.

第II部では, 筆者が今までに行ったベイズ計量経済分析の中から, いくつかの応用例を紹介する. 第3章は, 1992年に米国エール大学で開催されたYale-NSF共催の「ベイズ法と単位根」に関するシンポジウムで報告した論文であり, 時系列データを用いた最近の計量経済分析で大きなトピックスの一つである単位根問題をベイズの観点から分析した論文である. 本論文では, 誘導型方程式に基づいた単位根に対する事後オッズ比を導き, 次に単位根仮説をモデル選択問題として考えて, 予測の平均平方誤差の予測密度を導く. DSP対TSPを検定する事後オッズ比を導いた後, 単位根とトレンドパラメータに対する事後pdfを導く. 退化を避けるために, これらのパラメータに正則な事前分布を用い, 事前密度のパラメータを計算するのにブートストラップ法を用いている. そして, 誤差項が未知の次数のARMA(p,q)プロセスにしたがう場合に, 単位根をテストするための平方予測誤差の予測密度と事後密度を導く. さらに, このベイズ検定の方法を円の外国為替レートの月次データに応用した論文である.

第4章は, 1996年に出版された *Advances in Econometrics,* に収められている論文で, 為替レートモデルに用いられる変数を用いてベイズ流の単位根の分析と共和分のテストの方法を提案し, それを実際に応用したものである. 本論文は, 第3章で述べられた方法に基づいて単位根に関するベイズ分析を導き, さらに最高事後密度区間を用いた共和分の方法を導くものである. 帰無仮説の値の方向に事後分布を引っ張る自然共役事前分布を用いる方法を用いて, 事前パラメータの値をブートストラップ法によって経験ベイズの考え方で求めた論文である. 特に第5節でベイズ的な共和分分析の方法を示している. まず散漫な事前分布に基づく最高事後密度区間を導き, これがJohansenの尤度比検定と同等であることを示し, 共和分関係を含む誘導型係数行列のランクを示す特異値の事後分布を導いている. 特異値を用いたのは, VARモデルで方程式の順序に関して不変であるためである. そして第6節では, これらの方法を用いて為替レートモデルのベイズ分析を行う例が示されている.

第5章は, 1986年に東京大学出版会から発行された林　周二・中村隆英編『日本経済と経済統計』第7章に収められている論文である. トランスログ型生産関数を用いて石油危機前後における生産構造の変化を, 変化時点とともにある期間の生産構造から別の期間のそれへなめらかに変化するとし, その変化の調整スピードを同時に推定するベイズモデルを用いて分析した論文である. 特に, エネルギー多消費産業の中から鉄鋼業と紙・パルプ産業を取り上げ, エネルギー価格の変動に対する両産業の調整の違いと生産構造の変化について分析したものである.

第6章は, 1989年に東京大学出版会から発行された鈴木雪夫・国友直人編『ベイズ統計学とその応用』第6章に収められている論文である. 予測を行う場合に利用可能な情報源は, 大ざっぱにいえば, 過去の歴史的な出来事と, 将来のことに関する知識とに分けられる. 過去の歴史は単にGNPや物価の動きとかある製品の動きとかある製品の販売高や価格のような記録されたデータだけでなく, すべての過去の経験が含まれる. これらはすべて一緒になって予測システムを形成する. 1つは将来起きるかもしれない出来事について, その専門知識に基づいた見通しに関するものである. たとえば, 半導体の輸出規制の問題とか, 突然の円高や, 株暴落の可能性, 第3次オイルショックによるエネルギー危機の可能性,

等々についての予想に関しては, 分析者がそのインパクトをどのように考えるかによっている. もう1つは, フィードバック情報に関するもので, 予測パフォーマンスが不適切であるために生ずる. これは通常, 統計的モニタリング, すなわち最近時点でのモデルの予測能力を連続的に評価することによって発見することができる. 完全な予測システムは, このように統計モデルによって処理した情報と, 主観的な情報を結びつけたものでなくてはならない. ここで述べるベイズ予測の方法は, 動的線形モデルとその要素についてのモデル・ビルディングに基づいて行われるもので, システムを状態空間 (state space) 表現を用いて表し, 状態空間変数を未知の係数と考えたとき, この係数は時間で変化するので, このモデルのことをDLM (Dynamic Linear Model) という. このベイズ予測モデルの紹介とDLM分析を用いた応用例を一般的な読者を対象に解説した論文である.

各論文は別々に掲載されたため, 内容に一部重なる部分がある。また、論文の掲載時期も以前のものがあり, 現在の技術の発展から見ればもっと簡単に分析可能なものもあることをお断りしておく.

各章の論文の出典

第1章 ベイズ計量経済分析の手法

「ベイズ計量経済分析における最近の発展」第2,3章, 日本統計学会誌, 28–3, 253–305, 1998.

第2章 ベイズ計量経済分析の計算方法

「ベイズ計量経済分析における最近の発展」第4,5,6,7章, 日本統計学会誌, 28–3, 253–305, 1998.

第3章 単位根のベイズ分析

"A Bayesian Analysis of Unit Root under Unknown Order of an ARMA(p, q) Error and Application to the Foreign Exachange Rate of Yen" paper presented at Yale–NSF Symposium on 'Bayes Methods and Unit Root' held at April 24–25, 1992, Yale University. Working Paper No. 135, 1992.

第4章 為替レートモデルにおける時系列分析

"A Bayesian Analysis of Unit Root and Cointegraion with an Application to a Yen-Dollar Exchange Rate Model," Fomby, T.B. and Hill, R.C. eds. *Advances in Econometrics,* Vol.11, Part B, JAI Press, 51–86, 1996.

第5章 エネルギー価格の変動と産業調整

林　周二・中村隆英編『日本経済と経済統計』第7章, 東京大学出版会, 1986年, 119–145

第6章 ベイズ予測

鈴木雪夫・国友直人編『ベイズ統計学とその応用』第6章, 東京大学出版会, 1989年, pp.139-177

目 次

序 文 i

I ベイズ計量経済分析と計算問題 vx

1 ベイズ計量経済分析の手法 1

1.1 はじめに 1

1.2 ベイズ計量経済分析 3

1.2.1 事前分布 5

1.2.2 回帰モデルの標準的な分布 7

1.3 ベイズ的な方法の実証分析への応用 8

1.3.1 ベイズ的な方法の利点 8

1.3.2 ベイズ的な方法の問題点 10

事前分布の問題 10

1.3.3 ベイズ法を用いた応用分析のサーベイ 13

単位根問題 14

EBA(Extreme Bounds Analysis : 極値境界値分析) . . . 19

2 ベイズ計量分析における計算方法 23

2.1 ベイズ法における計算 23

2.1.1 正規近似 24

2.1.2 数値積分 25

2.1.3 ラプラス近似 25

例：ロジットモデル 27

2.2 モンテカルロ積分 28

2.2.1 単純なモンテカルロ積分 ・・・・・・・・・・・・・・・・・・・・ 29
対称反復法 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 29
例：単位根検定 ・・・・・・・・・・・・・・・・・・・・・・・・・・ 30
2.2.2 インポータンス・サンプリングを用いるモンテカルロ積分 31
インポータンス・サンプリングの計算 ・・・・・・・・・・・・・・・ 34
計算ソフトウェア問題 ・・・・・・・・・・・・・・・・・・・・・・・ 36
受容/棄却法 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 37
SIR:Sampling/Importance Resampling 法・・・・・・・・・・・・・・ 37
例：ロジットモデル(続き)・・・・・・・・・・・・・・・・・・・・・ 38
2.2.3 Poly-t 密度 ・・・・・・・・・・・・・・・・・・・・・・・・・ 39
2.3 マルコフ連鎖モンテカルロ法 ・・・・・・・・・・・・・・・・・・・ 39
2.3.1 ベイズ推定 ・・・・・・・・・・・・・・・・・・・・・・・・・・ 42
2.3.2 マルコフ連鎖 ・・・・・・・・・・・・・・・・・・・・・・・・・ 43
2.3.3 Metropolis-Hastings アルゴリズム・・・・・・・・・・・・・・ 45
2.3.4 Gibbsサンプリング法 ・・・・・・・・・・・・・・・・・・・・・ 47
例：マルコフ型スウィッチング回帰モデル ・・・・・・・・・・・・・ 52
例：打ち切りデータのある線形回帰モデル ・・・・・・・・・・・・・ 53
2.3.5 データ拡大(Data Augumentation)アルゴリズム ・・・・・ 54
例：センサー化自己回帰モデル ・・・・・・・・・・・・・・・・・・ 55
2.4 計量経済分析におけるMCMCシミュレーション法 ・・・・・・・・ 57
2.4.1 応用例 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 58
SUR:Seemingly Unrelated Regression Model ・・・・・・・・ 58
Tobit と Probit 回帰モデル ・・・・・・・・・・・・・・・・・・・・ 59
$AR(p)$ 誤差の回帰モデル・・・・・・・・・・・・・・・・・・・・・ 61
多重時点における構造変化・・・・・・・・・・・・・・・・・・・・・ 63
その他のモデル ・・・・・・・・・・・・・・・・・・・・・・・・・・ 67
2.5 おわりに・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 68
ベイズコンピュータパッケージについて ・・・・・・・・・・・・・・ 71
参考文献・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 76

II ベイズ計量経済分析の応用 91

3 単位根のベイズ分析 93

3.1 はじめに 93
3.2 ベイズ単位根モデルのレビュー 94
3.3 線形トレンドをもつAR(1)モデル 96
3.4 ARMA(p,q)誤差の次数が未知の場合の線形トレンドモデルにおける単位根 103
3.5 おわりに 111
参考文献 112

4 為替レートモデルにおける時系列分析 117

4.1 はじめに 117
4.2 為替レートモデルにおける実証分析 118
4.3 購買力平価モデルにおける伝統的な共和分分析 122
4.4 単位根のベイズ分析 126
4.5 共和分のベイズテスト 131
4.6 為替レートモデルの単位根検定と共和分分析 138
4.7 おわりに 143
参考文献 147

5 エネルギー価格の変動と産業調整 155

5.1 はじめに 155
5.2 エネルギー多消費産業とデータ 157
5.2.1 エネルギー多消費産業における省エネルギー 157
5.2.2 鉄鋼業と紙パルプ産業における省エネルギー 158
5.2.3 データの作成方法 159
5.3 モデルの定式化と推定 161
5.3.1 トランスログ型生産・コスト関数 161
5.3.2 確率制約下における構造変化モデル 166
5.4 計測結果の評価 171

5.5 おわりに ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 177
参考文献 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 178
付録1: 資本ストックデータの作成について ・・・・・・・・・・・・・・・・ 181
付録2: データの出所等 ・・・・・・・・・・・・・・・・・・・・・・・・・・ 183

6 ベイズ予測 185

6.1 はじめに ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 185
6.2 ベイズ予測と動的線形モデル(DLM) ・・・・・・・・・・・・・・・・・ 187
6.3 トレンド、季節性、回帰成分のDLM ・・・・・・・・・・・・・・・・・ 191
A. 線形成長モデル(多項式トレンドモデル)成分 ・・・・・・・・・・ 192
B. 加法的季節モデル成分 ・・・・・・・・・・・・・・・・・・・・・・ 193
C. 回帰モデル成分 ・・・・・・・・・・・・・・・・・・・・・・・・・ 194
6.4 DLMの予測と更新 ・・・・・・・・・・・・・・・・・・・・・・・・・・ 195
A. 主要結果 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 195
B. 情報割引 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 198
6.4.1 インターベンション分析 ・・・・・・・・・・・・・・・・・・ 199
6.5 DLM分析の応用例 ・・・・・・・・・・・・・・・・・・・・・・・・・・ 201
6.6 多変量モデルへの一般化 ・・・・・・・・・・・・・・・・・・・・・・・ 213
6.7 おわりに ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 220
参考文献 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 221

表目次

3.1 期待事後オッズ比, 平方予測誤差の平均の事後確率と差 ····· 102
3.2 事後オッズ比と単位根の標本検定の経験パワーの比較 ····· 104
3.3 事後確率と平方予測誤差の平均の差の計算した期待値 ····· 107
3.4 H_0' と平方予測誤差の平均の差の事後確率のパワーの計算値 ··· 108
3.5 ADF-t と F 検定の事後オッズ比, 事後平均と標準偏差: ····· 110

4.1 単位根検定 ····· 124
4.2 **Engle and Granger の共和分検定** ····· 124
4.3 **Johansen の検定** ····· 126
4.4 Dickey-Fuller F-検定と最高事後密度区間 (HPDI) ····· 141
4.5 共和分検定 ····· 143

5.1 製造業における生産額に占める燃料使用額(%) ····· 157
5.2 鉄鋼業における Allen の偏代替弾力性 (σ_{ij}) と自己価格弾力性 (η_{ii}) 174
5.3 紙・パルプ産業における Allen の偏代替弾力性 (σ_{ij}) と自己価格弾力性 (η_{ii}) ····· 176

6.1 事後分布の要約表(散漫な事前分布) ····· 205
6.2 事後分布の要約表(報知的な事前分布) ····· 205

図目次

4.1 η と α^* の事後密度 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 132
4.2 特異値 $\theta_1 = -.8$ の事後密度 ・・・・・・・・・・・・・・・・・・・・・・・・ 139
4.3 特異値(為替レートモデル)の事後密度 ・・・・・・・・・・・・・・・・・・・・ 142
4.4 使用したデータのプロット・・・・・・・・・・・・・・・・・・・・・・・・・・・ 144

5.1 鉄鋼産業のデータ：価格指数と数量指数 ・・・・・・・・・・・・・・・・・・・ 162
5.2 紙・パルプ産業のデータ：価格指数と数量指数・・・・・・・・・・・・・・・・ 163
5.3 鉄鋼業(コスト関数)における周辺事後分布 ・・・・・・・・・・・・・・・・・ 171
5.4 鉄鋼業(生産関数)における周辺事後分布 ・・・・・・・・・・・・・・・・・・ 171
5.5 紙パルプ産業(コスト関数)における周辺事後分布 ・・・・・・・・・・・・・・ 172
5.6 紙パルプ産業(生産関数)における周辺事後分布 ・・・・・・・・・・・・・・・ 172

6.1 原型列のプロット・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 204
6.2 散 布 図 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 205
6.3 1期先予測(散漫な事前分布) ・・・・・・・・・・・・・・・・・・・・・・・・ 206
6.4 1期先予測(報知的事前分布) ・・・・・・・・・・・・・・・・・・・・・・・・ 206
6.5 推定されたオンライン・レベル ・・・・・・・・・・・・・・・・・・・・・・・ 206
6.6 推定されたオンライン・回帰項とレベル ・・・・・・・・・・・・・・・・・・・ 207
6.7 推定されたオンライン・季節項 ・・・・・・・・・・・・・・・・・・・・・・・ 207
6.8 推定された季節ファクター・・・・・・・・・・・・・・・・・・・・・・・・・・ 207
6.9 推定された価格変数の係数プロット・・・・・・・・・・・・・・・・・・・・・・ 208
6.10 回顧的な食料品の1期先予測と90%信頼区間 ・・・・・・・・・・・・・・・・・ 208
6.11 回顧的な平滑化レベル ・・・・・・・・・・・・・・・・・・・・・・・・・・・ 208
6.12 回顧的な平滑化回帰項とレベル ・・・・・・・・・・・・・・・・・・・・・・・ 209
6.13 回顧的な平滑化残差 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 209

6.14 食料品のデータの将来予測 ・・・・・・・・・・・・・・・・・・・・・・・・ 209
6.15 販売データのプロット ・・・・・・・・・・・・・・・・・・・・・・・・・・ 211
6.16 販売データの1期先予測と90%信頼区間 ・・・・・・・・・・・・・・・・ 214
6.17 回顧的な1期先予測と90%信頼区間 ・・・・・・・・・・・・・・・・・・ 214
6.18 回顧的な平滑化レベル ・・・・・・・・・・・・・・・・・・・・・・・・・ 214
6.19 回顧的な平滑化成長率 ・・・・・・・・・・・・・・・・・・・・・・・・・ 215
6.20 回顧的な平滑化標準化残差 ・・・・・・・・・・・・・・・・・・・・・・ 215
6.21 自動的予測監視による1期先予測 ・・・・・・・・・・・・・・・・・・・ 215
6.22 回顧的な販売データの1期先予測(自動的) ・・・・・・・・・・・・・ 216
6.23 回顧的な平滑化レベル(自動的) ・・・・・・・・・・・・・・・・・・・ 216
6.24 回顧的な平滑化成長率(自動的) ・・・・・・・・・・・・・・・・・・・ 216
6.25 回顧的な平滑化標準化残差(自動的) ・・・・・・・・・・・・・・・・・ 217
6.26 販売データの将来予測(自動的) ・・・・・・・・・・・・・・・・・・・ 217

Part I

ベイズ計量経済分析と計算問題

第1章　ベイズ計量経済分析の手法

1.1　はじめに

計量経済学は, データを用いて経済的な現象についての確信 (経済理論) を改訂する分野の学問である. ベイズ流計量経済学では, 観測された事実をもとに確率法則にしたがってこの改訂が行われる. ベイズ流計量経済学で述べていることは, 経済理論における主要な問題である期待効用最大化とベイズ学習過程と同じであるといえる. 計量経済学が究極的に表明する問題は, 経済モデルの中で経済主体が直面しているものと同じである. すなわち, 新しく観測データが与えられると, 次にどのような決定をすべきであろうか? 2つの対立モデルを生成するデータを用いると, その相対的な妥当性はどのように変化するのか?

ベイズ統計学で採用されている手法は, 理論的な計量経済分析や統計分析においてますます多くの注目を集めるようになっているが, ベイズ手法を用いて経済現象の実証分析を行うことに対する関心は, 理論における関心の高さに比べて釣り合った増加を示していない. ごく最近までベイズ計量経済学の応用は, 主として理論面への貢献に関心がある人々によって行われており, 真にベイズ的である応用研究の部分は少ない. この理由はいくつかある. 第1に, ベイズ計量経済分析は尤度関数と事前分布が必要である. 他方, 非ベイズ分析では必要ない. 第2に, 主観的な事前分布の正しさを弁護しなければならず, もし同意が得られなければその研究は無視されるかも知れない. 第3に, ほとんどの事後モーメントは, 必要な積分が評価できないために, 求めるのが困難な場合が多い. このように応用研究が少ないということは, 最も単純な場合を除いて, ベイズ手法は実際に計算を行うには難しすぎるという認識によるものである. しかしながら, 計算力の最近の急速な発達は, 計算能力の発達を可能にした理論的な発展と相俟って, ベイズ手法が適用可能な問題の幅を急激に広げている. ベイズ的な計

算手法を用いている多くの論文は, 統計的な性質や特別な数学的な知識を利用しているので, 応用計量経済分析を行っている人々にはまだあまり馴染みがない.

本論の目的は実際的な面にある. ベイズ的な方法を採用する理由については, ベイズ的な方法と古典的な方法を比較して議論した優れたサーベイ (たとえば, Poirier, 1988, 1995; Berger, 1985 など) もあり, あまり深くは立ち入らない. むしろ実証研究にベイズ手法をどのように適用するかいうことに集中する. 全体の目的は, ほとんどあるいは特別な知識を持っていない研究者に, 応用実証研究におけるベイズ手法の有用性を認識してもらうことにある.

第 I 部では, ベイズ計量経済分析の紹介と実証分析を行う場合の利点と問題点について述べる. 第 II 部では, ベイズ分析にとって重要な計算問題に焦点を合わせ, 最近の発展について整理する. 第 I 部の構成は次の通りである. 第 1.1 節は, 全体的な構成を述べるはじめにの部分である. 第 1.2 節で, ベイズ計量経済分析について簡単にサーベイする. 第 1.3 節では, ベイズ的な手法を実証研究に応用することに焦点を合わせ, ベイズ実証分析の優れた点 (1.3.1 節) と問題点 (1.3.2 節) について述べ, 1.3.3 節はベイズ応用実証分析の例として, 単位根モデルと EBA(極値境界分析) を中心としてサーベイを行う. 第 II 部の構成は次の通りである. 第 2 章以降では, ベイズ流の実証研究を行うのに用いられている, いくつかの計算方法を中心として述べる. まずベイズ計量経済分析で用いる基本的な計算方法である正規近似 (2.1.1 節), 数値積分 (2.1.2 節), ラプラス近似 (2.1.3 節) についてその概略を説明した後, 第 2.2 節では, 最近多く用いられるようになったモンテカルロ積分について, 単純な場合 (2.2.1 節) とインポータンス関数を用いた場合 (2.2.2 節) について, その使用例とともに解説する. Poly-t 密度について簡単に触れた後 (2.2.3 節), 第 2.3 節ではもっと複雑な問題に用いることができるマルコフ連鎖モンテカルロ・シミュレーションについて述べる. 一般的なベイズ推定 (2.3.1 節) とマルコフ連鎖 (2.3.2 節) についてふれた後, 代表的な方法である Metropolis-Hasting アルゴリズム (2.3.3 節) とよく使用される Gibbs サンプリング法 (2.3.4 節), それにデータ拡大アルゴリズム (2.3.5 節) について述べる. 次に, 第 2.4 節では MCMC シミュレーションを計量経済分析にどのように適用するかを応用例とともに述べる. ここでは, 事後分布からシミュレーションを行う方法の計量経済分析への応用について強調し, 重要な問題である収束

の診断方法や数学的詳細は省略する. 最後に, 第2.5節ではいくつかの結論を述べ, 付録でベイズ実証分析を行う際に役立つコンピュータ・パッケージについて述べる.

1.2 ベイズ計量経済分析

最初に, ベイズ計量経済理論の背後にある基本的な理論的概念について簡単に整理しておこう. 最近では, 英語で書かれた多くの計量経済学のテキストではいくつかの章でベイズの方法が扱われているが, 残念ながら日本語で書かれた計量経済学の教科書ではほとんどふれられていないのが現状である. ベイズ計量経済理論についてより詳しく知りたい場合は, このトピックスについての多くの参考文献があるので, それらを参考にすると良い (たとえば, Zellner, 1971; Poirier, 1988, 1995).

ベイズ計量経済学は, 次のベイズの法則に基づいている.

$$p(\theta|y) = \frac{p(\theta)L(y|\theta)}{p(y)} \propto p(\theta)L(y|\theta)$$

ここで θ はパラメータベクトル, y はデータベクトルあるいは行列, $p(\cdot)$ は確率密度関数を表す. $p(\theta)$ はデータを分析する前にパラメータについての主観的な確信を表す**事前密度** (*prior pdf*), $L(y|\theta)$ は θ に特定の値が与えられたときの y の条件付き密度で**尤度関数** (*likelihood function*) と呼ばれる. そして $p(\theta|y)$ は θ の**事後密度** (*posterior pdf*) であり, データに含まれる新しい情報を考慮して事前の確信を更新することを表している. 事後密度は, 観測されたデータ y の条件付で θ について得られるすべての情報を要約したもので, θ に関する推測は事後密度から得られる. たとえば, θ の事後平均 $E(\theta|y) = \int \theta p(\theta|y)d\theta$ は, θ の点推定値としてしばしば用いられる.

予測も簡単に得ることができ, y^* を1組の将来の観測値とすると, 予測は予測密度 $p(y^*|y)$ をもとにして行われる. 確率法則を用いて, $p(y^*|y) = \int p(y^*|\theta, y)p(\theta|y)d\theta$ となり, ここで $p(y^*|\theta, y)$ は将来観測値に関する尤度関数, そして $p(\theta|y)$ は事後密度である.

ベイズの方法を用いて仮説を比較するいくつかの方法がある. その中で最もよく使われるのは**事後オッズ比** (*posterior odds ratio*) である. たとえば, もし2

つの異なる仮説 H_1 と H_2 (入れ子型あるいは非入れ子型) を比較しようとすれば, 事後オッズ比 K_{12} は $K_{12} = p(H_1|y)/p(H_2|y)$ となる. 言い換えると, K_{12} は観測データから得られる2つの仮説の確率の比である. もし θ_1 と θ_2 が, H_1 と H_2 それぞれの仮説のもとで妥当なパラメータベクトルであるなら, 事後オッズ比は次のように計算される.

$$K_{12} = \frac{\pi(H_1) \int p(\theta_1|H_1) L(\theta_1|y, H_1) d\theta_1}{\pi(H_2) \int p(\theta_2|H_2) L(\theta_2|y, H_2) d\theta_2}$$

K_{12} の計算には, 仮説が成り立つ事前確率 $(p(H_i), i = 1, 2)$ と同様に, 各仮説のもとでの事前情報と尤度が必要である. 1つの決定は, $K_{12} > 1$ なら H_1 を受容し, その他の場合は H_2 を受容することである. もちろん, もっと手の込んだ方法も利用できる. たとえば, 非線形制約の検定に関心がある場合, 制約のないモデルと制約付のモデルの両方の下で事前分布を作成するのは大変であるので実際的ではない. しかし, McCulloch and Rossi (1992) は, この種の問題を克服できるある優れた方法を開発している.

応用研究ではベイズファクターも通常報告される. これは2つの仮説に対する尤度比のサポートになる. 言い換えると, ベイズファクターは2つの競合する仮説に対して同じ事前分布を与えて計算した事後オッズ比である.

計量経済学で通常行われる代表的な統計的分析であるパラメータの推定, 仮説検定, そして予測に対するベイズ的な方法を簡単に要約する. ベイズ的な方法と従来からの古典的な計量経済分析の方法との重要な違いは, 前者が未知の変量 (たとえば θ, y^* や仮説の真実性など) を確率変数として扱い, 既知の変量 (たとえば, データ) に条件づけるのに対し, 後者はデータを確率変数として扱い, パラメータや仮説の状態を非確率的なものとして考えることにある. ベイズ分析を行うときには, 情報のない事前分布が用いられる場合が多いけれども, 尤度関数と事前分布の両方を定式化する必要がある[1]. 定式化を行った後, パラメータの推定値を得たり, 仮説を検定したり, あるいは予測を行ったりするために, 簡単に上で述べた方法を使うことができる. したがって, 概念的には少なくともベイズ的な方法は, 古典的な方法に比べればずっと簡単に使うことができる.

ベイズ法を用いる場合には, 積分の計算が必要になる. 事後平均はしばしば

[1] ベイズ意思決定理論的な分析では, 損失関数も定式化する必要があるが, この段階は実際にはよく省略される.

推定量になり, $E(\theta|y) = \int \theta p(\theta|y)d\theta$ を計算する. 事後オッズを得るためには, $\int p(\theta|H_1)L(\theta|y, H_1)d\theta$ を計算する必要がある. **自然共役事前分布**[2]を持つ正規線形回帰モデルのような標準的な場合では, これらの式の解析的な解は存在する. しかし, 一般的な場合には存在せず, ベイズ分析を行う場合はコンピュータ集約的な手法に頼らざるを得ない.

1.2.1 事前分布

事前分布は, 分析者が関心をもつパラメータ θ についての確信の度合いを記述したもので, 現在のデータセットからすべての情報が得られる以前の情報である. 事前分布はそれが確信の度合いを表すことから, 厳密な意味では不正確ではあり得ないが, 確信が合理的であることの検定が必要であるという意味で主観的である. しかし, 事前分布はデータセットからのどんな情報も拒否するほど独善的であるべきではない.

事前分布は次のようないくつかの理由によって分類される. 最も簡単な事前分布を作成する方法は, とりうるパラメータの値を集めて全体が 1 になるように基準化することによって, 事前分布をできるだけ事前の確信を反映させるようにする. θ が離散的な値をとる場合は自然な方法であり, 連続的な値をとる場合は実線上の区間をとって θ に対する事前のヒストグラムを作成する. あるいは θ に対する事前分布が, あるパラメトリックな分布族 $\pi(\theta|y)$ に属すると仮定して, y を結果ができるだけ真の事前の確信に近くなるように選ぶ. たとえば, y が 2 次元の場合, 2 つのモーメント (たとえば平均と分散) か, 2 つの分位値 (たとえば 50%と 95%) を与えれば, その厳密な値を決めることができる. しかしこの方法の限界は, 事前の確信が標準的なパラメトリック型に収まらない場合がある点である. さらに, 一見すると同じように見える 2 つの分布が, 全く異なる性質を持っている場合がある. たとえば, Berger(1985, p.79) で指摘されているように, Cauchy(0, 1) と Normal(0, 2.19) 分布は同じ 25%, 50%, 75%点 $(-1, 0, 1)$ を持ち, プロットすると非常に似たように見える密度関数が全く異なる事後分布になることが知られている.

[2] 事前密度は, 事後密度と事前密度が同じクラスの確率密度関数になるようにしばしば選ばれる. このような事前密度は, 自然共役事前分布と呼ばれる.

共役事前分布 (*conjugate prior*) は, 尤度関数 (とそのパラメトリック型) と結合したとき解析的に調べることができる形 (すなわち, 多変量 t 分布のような既知の認識できる形) を持つ事後分布になる. 分布を解析し, 積分を評価する数値的な方法における最近の理論的発展がある前は, 共役事前分布はベイズ分析で非常に重要であった. その事前分布が仮定した尤度と共役な特定の分布のクラスの制約の中で, よく当てはまるようにする必要があった. 事後分布を形成する数値手法に新しい発展があった後は, 共役事前分布は必要性がなくなり, 自由に自らの事前の確信を真に表しているどんな事前分布も選べるようになった. 事実, 最近ではリーズナブルと思われる自由な事前分布を描けるようになり, コンピュータを使って様々な分布 (正規分布, t 分布, ベータ分布等) を描き, 分布の形をコントロールするパラメータを適当に変えることによって, 自分の確信と合うように分布を作成することができる. これは事前情報の数学的表現を与え, 少なくとも共同研究者の 1 人が様々な統計分布を描くことができれば, かなり容易に行うことができる.

事前分布は, 少なくとも θ のいくつかの要素に関して事前に持っている知識を表す**報知的** (*informative*) であるか, あるいは無知の状態を表す**非報知的** (*non-informative*) に分けられる. 最もよく知られた非報知事前分布, すなわち無知である事前分布は *Jeffreys* 型事前分布で, 情報行列 (すなわち, パラメータ θ に関して対数尤度関数の 2 階の偏導関数の負の期待値) の行列式の平方根に比例してとられる. 誤差項が正規分布にしたがう古典的な線形回帰モデルでは, $\theta = (\beta', \sigma)'$ であり, *Jeffreys* 型事前分布は $p(\theta) \propto \sigma^{-1}$ で与えられる. 簡単に言えば, Jeffreys は, もし事前の無知を表したければ, 無限の許容可能な範囲を持つパラメータ (すなわち, 一様分布) に対しては定数に比例するような事前分布を使うべきであり, $[0, \infty)$ の許容できる範囲を持つ (標準偏差のような) パラメータに対してはパラメータの自然対数上で一様分布である事前分布を用いるべきであることを提案した. 古典的な回帰モデルに対する事前分布につながるのはこれらの法則である.

事前分布は**正則** (*proper*) か, あるいは**非正則** (*improper*) である. **正則な事前分布**は, θ の許容可能な範囲に関して積分すると 1 になる. 報知的である分布は, 適切に基準化すればほとんど常に正則である. たとえば, もし切断された正規分

布を限定された範囲のパラメータに関する事前分布として用いると, この分布は切断を考慮して基準化しなければならず, 事前分布全体が 1 になるように分布を基準化し直す必要がある. 非報知事前分布は, しばしば非正則である. 無限の範囲の一様分布は積分すると無限になり, 1 にはならないので**非正則**である. 非正則事前分布は仮説検定の分野で問題が生じる. これは, 事後オッズ比を任意に基準化することができるために, 競合モデルに対する相対的な支持の評価が不可能になるためである. このために, 非正則事前分布をできるだけ使用しないようにする. 多くの応用分析では, データの尤度関数によって事後分布をほとんど完全に決めることができ, 他方, 仮説検定に対してはオッズ比が任意に基準化されないような弱い報知事前分布を作ることができる.

1.2.2 回帰モデルの標準的な分布

計量経済分析では線形回帰モデルがよく使われるが, 通常用いられている散漫な事前分布の下で, ベイズ分析で用いるいくつかの標準的な分布を示すことにする. 必ずしもそのような非報知事前分布を常に用いる必要はないが, ベンチマークとしては意味があり, 回帰パラメータの事後分布を解析的に調べるのに役立つ. さらに, 本節で示される分布は, 4 章以降で述べるいくつかの数値積分法で確率抽出する乱数に関して役に立つ情報を与える.

まず, すべての古典的な仮定の下での標準回帰モデルを考える.

$$y = X\beta + \varepsilon$$

ここで y は $(n \times 1)$ 従属変数の観測値ベクトルでその変動をモデルで説明しようとし, X は $(n \times k)$ 説明変数行列, β は $(k \times 1)$ の推定されるベクトルすなわち回帰パラメータ, そして ε は分散が σ^2 のホワイトノイズ誤差項の同じ次元のベクトルである. もし, 回帰パラメータに非正則な一様事前分布 (これも *Jeffreys* 型事前分布), そして σ に標準的な *Jeffreys* 型事前分布を選択すると, (β, σ) に対する同時事前分布は次のように書ける.

$$p(\beta, \sigma) \propto \frac{1}{\sigma}$$

正規性の仮定の下で, 従属変数の観測値に対する尤度関数は

$$p(y|\beta, \sigma, X) = (2\pi\sigma^2)^{-n/2} \exp[-\frac{1}{2}(y - X\beta)'(y - X\beta)/\sigma^2]$$

であり, 同時事後分布は次のように表せる.

$$p(\boldsymbol{\beta}, \sigma|\boldsymbol{y}, \boldsymbol{X}) \propto (2\pi)^{-n/2}\sigma^{-(n+1)} \exp[-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})/\sigma^2]$$
$$\propto (2\pi)^{-n/2}\sigma^{-(n+1)} \exp[-\frac{1}{2}(n-k)s^2/\sigma^2 + (\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})'\sigma^{-2}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})]$$

ここで, 上の式は二次形式の項に分けられ, $\boldsymbol{\beta}$ と σ^2 の最小 2 乗推定量 $\widehat{\boldsymbol{\beta}}$ と s^2 を用いて表されている. パラメータの周辺事後分布は

$$p(\boldsymbol{\beta}|\boldsymbol{y}, \boldsymbol{X}) \propto [(n-k)s^2 + (\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})'\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})]^{n/2}$$

となり, これは多変量 t 分布にしたがう. そして

$$p(\sigma|\boldsymbol{y}, \boldsymbol{X}) \propto \sigma^{-(n-k+1)} \exp[-\frac{1}{2}(n-k)s^2/\sigma^2]$$

は逆ガンマ分布にしたがい, σ^2 の周辺事後分布は χ^2 分布を基準化したものであることを意味している. 分散-共分散行列を Σ と表す一般化誤差分布の場合は, この分布は簡単に一般化でき, Σ の周辺事後分布は逆ウィッシャート分布にしたがう. この分布は, 逆ガンマ分布を一般化したものである.

1.3 ベイズ的な方法の実証分析への応用

1.3.1 ベイズ的な方法の利点

本稿での主要な関心は, ベイズ的な方法を経済問題に対する実証研究に応用することである. ベイズ的な見方と古典的な見方との違いをはっきりさせるために, 両者の考え方と方法論に主に焦点を合わせた分析は Poirier(1988) あるいは Berger(1985), 118–126 ページを参照すると良い. 実証的な問題の分析に関しては, ベイズ的な計量経済分析の方法は古典的な方法と比べて, その単純性 (*simplicity*) と整合性 (*coherency*) という点において優れている. 古典的な手法に必要とされる多くのテクニックと基準とは異なり, ベイズ的な方法では (非報知的な) 事前密度と尤度関数だけを必要とする. 対照的に, 古典的な実証分析では適切な手法を選択し, その統計的な性質を知らなければならない. しかも多くの性質は漸近的にのみ既知であり, 小標本でのパフォーマンスにともなう問題は残されている.

ベイズ法は観測されたデータの条件付の事後密度に基づいている. これは, 古典的な手法が推定量の性質が繰り返し標本で評価される不偏性のような基準に基づいているのと対照的である. このような基準はベイジアンにとってはほとんど関心がない. さらに, この条件付けを行うことは, 特定の応用に対して合理的な計量経済理論を単純化するだけでなく, しばしば応用そのものを単純化する. たとえば, Geweke (1988) の不等号制約にしたがうパラメータをもつモデルのベイズ分析は, 古典的な方法に比べて非常に簡単になるのはよい例である (Poirier, 1988). 同様に, McCulloch and Rossi (1992) は, ランダム係数構造をもつ多対プロビットモデルを用いて計算的に可能な方法を開発した. パネルデータを用いた研究では, 消費者の異質性を認めることが必要であり, そのような異質性をモデル化する標準的な方法は, ランダム係数モデルである. しかし, 古典的な方法を用いて多対プロビットモデルの枠組みでランダム係数モデルを用いると, 多くの計算上の厄介な問題が生じる. しかし Gibbs サンプリングを用いたベイズ法では, 階層モデルに対してと同様な利点がある.

ベイズの方法による条件付けを行った場合の性質は, 観測された標本に対して厳密な結果が得られることを意味している. しかし, 古典的な計量経済学では, 特定の検定や推定量を用いるのに小標本での性質が漸近的な結果に対応していると考えて, しばしば漸近的な性質を用いて正当化する. 小標本と大標本の性質が大きく異なる場合がたくさんあるが, 推定量の小標本の性質を得るのは難しい場合が多い.

ベイズ法のもう 1 つの利点は, **局外母数** (*nuisance parameter*) の取り扱いにある. これは古典的な計量経済分析を行うには特に厄介な問題である. 通常は, それらを有限標本の性質を失うというコストを払って一致推定量で置き換える必要がある. しかしベイズ法では局外母数に関して簡単に積分できる. たとえば, 線形回帰モデルで回帰係数 β と誤差分散 σ^2 とすると, 事前分布に尤度関数を掛けて事後分布 $p(\beta, \sigma^2|y)$ を得る. もし β だけが必要なら, ベイズ分析では $p(\beta|y) = \int p(\beta, \sigma^2|y)d\sigma^2$ とするだけである[3)].

[3)] 実際には, 局外母数に関して積分すると問題をずっと簡単にできる. しかし, 次の 2 つの理由から結果に影響を与えることがある. それは i) 古典的な計量経済分析では, 局外母数を一致推定値でしばしば置き換える方法があるが, 小標本では一致推定量によって結果が異なることがあり得る. しかし, そのような類似の問題はベイジアンでは生じない. ii) 局外母数のベイジアンでの扱いは, 周辺密度 ($p(\beta|y)$) で考えることになる. 古典的な手法に似たベイズ的な方法では

1.3.2 ベイズ的な方法の問題点

事前分布の問題

ベイズ法で一般に認識されている問題点は, 事前情報を含めることを認めることである. この事前情報を用いることに対して多くの論争が行われたが, その詳細についてここでサーベイを行わない. しかし, もし事前情報が得られればそれを含めるべきであり, ベイズ法はそれを行う最良の方法を提供する.

実証分析では, 事前分布を変化させたことによる事後分布の感度の問題は重要である. しかし, 事前分布の生成 (*elicitation*) の問題はより重要である. 計量経済分析では自分だけが満足する結果が得られればよいわけではない. このために, 自分の事前の確信に対応するただ1つだけの事前分布を選択すべきでなく, むしろ, 様々な専門家の意見を取り込んだ一定の範囲の事前分布を選択すべきである. ほとんどの応用問題では, それまでの研究や経済理論から研究者間にかなりの合意がある. たとえば, β を生産モデルでの規模に対する収穫を測るパラメータとすると, 研究者によって様々な事前分布を β に持つ. ほとんどは β が1に近く, 人によっては β が0.75に近いと思っているが, だれも β が100に近いと思っていない. そこで, 合理的な事前密度は主要な確率が区間 $[0.5, 1.5]$ 内にあるとする. しかし, $\mathcal{N}(1, 0.25), \mathcal{N}(0.75, 1)$, あるいは $\mathcal{N}(1.25, 1)$ のように, この性質を持つ多くの事前分布があるので, 応用研究ではこれらすべての様々な事前分布を試すことができる. 尤もらしい一定の範囲の事前分布に対して結果が頑健であることは, ただ1つの事前分布に対して示される結果よりもずっと受け入れやすい. 多くの実証分析では, 推定は事前分布の選択に関してそれほど敏感でない. もし異なる合理的な事前分布によってまったく異なる推定値が得られたとすると, 重要な情報が尤度関数の性質について明らかになる. すなわち, 尤度関数は典型的にはかなりフラットで, したがって非報知的であることになる.

もし事前分布の生成が期待できないか非常に困難である場合には, 非報知事前情報のクラスをそのかわりに用いることができる[4].

条件付密度 ($p(\beta|\sigma^2 = s^2$, ここで, s^2 は σ^2 の一致推定値) で考えることである. 実際には, 古典的計量経済学では滅多に $p(\beta|\sigma^2 = s^2)$ を使わない. いうまでもなく, 局外母数に一致推定値を当てることは小標本効果を持つ.

[4] 事前分布の感度分析についてのベイジアンの文献は急速に増加している. この分野への役

多くの場合, パラメータについてたぶん曖昧な事前情報しか持っていない場合が多い. 再び, β を規模に関する収穫を測るパラメータとすると, β は 1 に近いと感じており, 1.5 より大きいか 0.5 より小さくはないというはっきりとした確信を持っている. この性質を持つ非常に多い事前分布がある (たとえば $\mathcal{N}(1, .0625)$). 実際には, 多くのベイズ的な応用研究では扱いやすい事前密度の関数形が選択される. 次に, 事前情報に基づいて事前密度のパラメータ (ハイパーパラメータと呼ばれる) を選ぶ. したがって, 以前の例での β に対する事前分布は正規分布であると決める. それはたんに正規分布が扱いやすいからである. そして, 事前の情報に基づいて事前平均が 1 で事前分散が 0.0625 とする. β と σ^2 によって標準的な方法で定式化された正規線形回帰モデルに対しては, β と σ^2 に対する正規・ガンマ事前分布は自然共役になり, 解析的な結果が存在し, この事前分布がこのモデルに対して広く用いられるようになった.

ある関数形が事前分布に対して選ばれたら, 事前分布の導出問題は事前のハイパーパラメータの選択に要約される. 実際は, ベイジアンは, パラメータの経済学的な解釈を通じて非公式に推論することによって, 事前のハイパーパラメータをしばしば選択する. たとえば, 上の例で β は規模に対する収穫の弾力性であるから, それはたぶん 1 に近いと推論する. したがって, 1 の事前平均と 0.0625 の事前分散を選択する.

モデルのパラメータに単純な経済学的解釈ができない場合が数多く存在するが, 合理的な事前分布を導出することはそれでも可能である. たとえば, Koop and Poirier (1993) はロジットモデルに対する自然共役事前分布を導出している. 彼らは最初に, このモデルのパラメータを直接に解釈するのが難しく, したがって事前分布の導出が難しいと述べている. それでも, それらの事前分布に対するハイパーパラメータを, パラメータ自体ではなく観測可能性に関して推論することによって, 選択する方法を述べている. すなわち, 事前分布の導出を次のような問いをもとにして行う : 「ある特性をもつ個人は, その特性をもたない人に関連する別の特性を選択することがどのくらいあり得るか?」

Kadane et al. (1980) は, 正規線形モデルに対してこの種の接近方法をとるソフトウェアを開発している. この対話型ソフトウェアは使用者に, X のある値

に立つ入門書として Berger (1985), 195–247 ページ, あるいは Kadane (1984) がある. 線形回帰モデルに対する事前分布の感度についての議論は, Leamer (1978, 1983, 1985) にある.

に対して y のとる値を予測するように促す. すると想定した事前分布が, いくつかの質問に対する反応を基に計算できる.

これまでは, 事前分布の選択に関するいくつかの問題に対する簡単な紹介を行った. その他の事前分布の作成方法は Berger (1985) の第 3 章にある. この段階でいくつかの点を強調しておく必要がある. i) 多くのモデルのクラスでは, 通常用いられる事前分布の標準的なファミリーが存在する. たとえば正規線形回帰モデルに対する正規・ガンマ事前分布である. これらのクラスに対しては, 事前分布の作成問題は適切な事前のハイパーパラメータを選択することに帰着する. ii) 多くの人の納得を得るために, 分析者は事前分布の感度分析を行うべきである. iii) そのような事前情報を含めることは, 合理的な非報知事前分布がほとんどのモデルに対して存在するので, どうしても必要なわけではない.

後者の場合, そのようなクラスの**非報知事前分布**は非常に役に立つ. ほとんどのモデルに対して, 様々な定義の非報知性に対応する非報知事前分布があるが, 実証的な結果は通常そのような事前分布の選択に対して敏感ではない. 最もよく知られた非報知事前分布は, $p(\theta) =$ 定数 である. そのようなフラットな事前分布を用いることは, 尤度関数に比例した事後分布を得ることになり, したがって, これを用いるベイズ分析は尤度関数だけを含むことになる. このようなフラットな事前分布を用いた結果は, 古典的な研究で得られたものとしばしば似通った結果になる[5]. たとえば, 事後密度のモードは最尤推定値と同じである.

フラット事前分布は, パラメータ変換に対して不変ではないことで批判されている. すなわち, θ がある区間に入る事前確率は, $f(\theta)$ を任意の関数としたときに $f(\theta)$ がその区間に入る最大確率と同じでない. 非報知事前分布を導出するもう 1 つの方法はジェフリーズのルールとして知られており, 事前分布を情報行列の行列式の平方根に比例する様に選択することを示唆した. このようにして作成された非報知事前分布は不変であり, 多くの場合結果として得られた事前分布はフラットなものであることに注意する. たとえば, 正規線形回帰モデルではジェフリーズのルールの応用は, β と $\ln(\sigma)$ に対するフラット事前分布と同

[5] しかし, これは常に成り立つわけではない (たとえば, 以下の単位根計量経済学についての議論を参照). 大雑把な経験則では, フラットな事前分布を用いたパラメータ推定値は最尤推定値と非常に似た数値になる傾向にあるが, 仮説検定はしばしば非常に異なる結果を生み出す. この論文の第 1 節では, ベイジアン仮説検定に含まれる方法を簡単に述べている. 概観した方法は古典的なものとは全く異なり, このためにかなり異なった結果になることを強調する価値がある.

じである $p(\beta, \sigma^2) \propto \sigma^{-1}$ の形の事前分布になる[6].

しかし, 別のもっと深刻な非報知事前密度の問題点は, それらがしばしば非正則になる, すなわち積分して1にならないことである. 実線上での一様密度が通常用いられているが, 非正則な非報知事前分布である. 非正則な事前分布は推定や予測に対しては問題を起こさないが, 同じことは仮説検定に対しては当てはまらない. 特に, 制約のない対立仮説に対するシャープな帰無仮説の検定に対する事後オッズは, しばしばデータに拘わらず帰無仮説が好まれる (Leamer (1978) 参照)[7]. 結果として, 非正則な事前分布は, もし研究者が事後オッズ比で一点帰無仮説を検定したいならば, 用いるべきではない[8].

1.3.3 ベイズ法を用いた応用分析のサーベイ

ベイズ流の分析は応用経済分析への研究には僅かしか利用されてこなかったけれども, それが受け入れら始めたのは90年代になってからである. 初期の応用研究は, 最近のいくつかの *Journal of Econometrics* のベイズ応用分析の特集号[9]や, *Advances in Econometrics*[10] に見られる. 本稿ではベイズ分析を採用したすべての応用分野の詳細なサーベイは行わないが, 計量経済分析についての包括的な情報は Zellner(1983, 1985) を参照すると良い. そのかわり, ここではベイズ法が広く用いられており, 古典的な方法と比較可能な2つの分野である単位根検定と EBA(極値境界値分析) を取り上げることにする. 前者は, 多くの論

[6]非報知事前分布についてのより詳細な議論については, Zellner (1971), 41–53 ページ, あるいは Berger (1985), 82–89 ページを参照.

[7]事後オッズは, 帰無仮説と対立仮説の下での尤度関数の加重平均の比であり, ここでウェイトは事前密度によって与えられる. もし対立仮説の下での事前密度がフラットになったら, 尤度関数の裾にもっと大きなウェイトを与える. したがって, 対立仮説の下で加重平均は小さくなり, 帰無仮説は事後オッズ比では好まれるようになる.

[8]もっと正確には, 事前密度は, 帰無仮説の下では制約付であるこのようなパラメータに対しては非正則にならない. しかし非正則な事前密度は, 帰無仮説と対立仮説の両方の下で制約のない, このようなパラメータに対しては問題を生じない. たとえば, 線形回帰モデルでの関心が回帰係数 β_i がゼロかどうかを検定するのに事後オッズを用いるときに, β_i に対しては正則な事前密度を使用しなければならない. しかし, $\beta_j (j \neq i)$ に対してや誤差分散に対して, 非正則な事前密度を使用することは完全に受け入れることができる.

[9]Bayesian Empirical Studies in Economics and Finance, ed. by D.J.Poirier, 1991; Bayesian and Classical Econometric Modelling of Time Series, ed. by L.Bauwens and M. Lubrano, 1995; Bayes, Bernoullis, and Basel, ed. by L.Bauwens, W. Polasek and H. van Dijk, 1996 の各特集号に多くのベイズ流の応用分析がある.

[10]T.B.Fomby and R.C.Hill ed. *Advances in Econometrics, Vol.11 A,B* に多くのベイズ応用分析がまとめられており, Part A は Bayesian Computational Methods and Applications, Part B は Bayesian Methods Applied to Time Series Data のそれぞれ特集号である.

争と有意な差がベイズ法と古典的な方法との間に見られるので, 多少詳しく説明する. 一方, 後者はベイジアンと古典的な計量経済学者が多くの一致する立場に立てるケースの良い例である. このほかにも, 制約付回帰分析の応用でベイズ的な方法が成功した例としてシラー・ラグがあるが, 前に別の論文で取り上げたので省略する.

単位根問題

ある系列が階差定常過程 (DSP) かトレンド定常過程 (TSP) かという問題は, 1980 年代に多くの計量経済学者の注目を集めたトピックスであるが, 経済学と統計学の両方の意味を持っている. 系列が単位根を含むかどうかはマクロ経済学にとって非常に重要であり, 技術パラメータが単位根にしたがうと考えて, リアル・ビジネスサイクルの理論の検証に多くの経済学者が参加した. 統計的な問題としては, トレンド除去法と単位根検定の問題がある. 古典的な方法による単位根問題の分析はすでに数多く行われており, それらの結果は Hatanaka (1996) にまとめられている.

経済理論はいくつかの時系列は単位根を必ず含むはずであることを意味しているが, 一方統計理論は定常な系列と単位根系列の性質の違い, そしてそれに基づく推定量と検定統計量を強調している. Dickey and Fuller (1979) と Nelson and Plosser (1982) を始めその他のよく知られた研究以来, 古典的な計量経済学者は経済時系列における単位根の検定問題にかなりの努力を注いできた. 単位根計量経済学をめぐる多くの問題は, 次の AR(p) を再定式化した単純なモデルを考えることによって示すことができる.

$$y_t = \rho y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + u_t$$

ここで y_t $(t = 1, \ldots, T)$ は観測されたベクトルの t 番目の要素であり, u_t は観測されない誤差項である. 実証が可能な単位根は $\rho = 1$ ならば存在する. しかし, 誤差項の性質について異なった仮定をすることが多い. たとえば, Dickey and Fuller (1979) は u_t が *iid.* 正規を仮定し, t 分布に従わない $\rho = 1$ を検定するのに, t 統計量に対する臨界値をシミュレートした. その後 Phillips and Perron (1988) では, 誤差項についてもっと一般的な仮定を考え, t 統計量の関数

である検定統計量を導いた. 実証研究で単位根を検定するときには, しばしば Dickey-Fuller テストか Phillips-Perron の拡張が用いられる. しかし, いくつかのモンテカルロ研究からの事実 (たとえば Schwert, 1989 や DeJong, Nankervis, Savin and Whiteman, 1992 参照) は, これらの古典的な単位根検定は極めて低い検出力しかもたず, 特にトレンド定常な対立仮説に対しては検出力は低いことを示している.

Sims (1988) と Sims and Uhlig (1991) による議論の発端となった論文は, 単位根仮説を調べるのに古典的な計量経済学的な方法を用いることを批判し, フラットな事前分布を用いることを支持しており, 最近の多くの実証分析では, フラットな事前分布を用いたベイズ法を採用していることを示した. これらの論文は, ベイズ的な方法による単位根計量経済分析に大きな関心を呼び起こした. 最近の実証分析を見ると, フラットな事前分布を用いるベイズ法では, 経済時系列の単位根仮説に関する推測が古典的な検定による結果とはしばしば異なることを示している. 特に, DeJong and Whiteman (1991a,b,c) によるフラットな事前分布を用いたベイズ分析では, Nelson and Plosser (1982) によって研究されたマクロ時系列や, Shiller (1981) と Marsh and Merton (1986) で分析されたストック価格や配当系列に単位根はほとんど見い出されていない. 同様に, Schotman and van Dijk (1991a) によるフラットな事前分布を用いたベイズ分析では, 実質為替レートや実質 GNP データに単位根があるという事実は, 古典的な単位根検定の結果によって支持されよりもずっと弱いことを見い出した.

ベイズ単位根の分析は, ベイズ分析を行っている人々の内から批判されている. たとえば DeJong and Whiteman (1991a,b) で用いられたフラット事前分布の型は非報知ではないということについての論争から, Peter Phillips は時系列モデルにフラット事前分布を用いるのに強く反対した. それは, フラット事前分布はある種の不変性を欠いており, 自己回帰係数がコレログラムの形状に影響を及ぼし, データやその標本積率に含まれる期待情報量に影響を与えるという経路で, 一般的なモデルに基づいた重要な情報を無視しているからである. Phillips (1991) は, 非報知事前分布の作成に対する Jeffreys のルールを用いて, **無知の事前分布** (*ignorance prior*) と呼ばれる事前分布を開発した. Phillips の事前分布は, 拡散的な動きを意味するパラメータ空間に様々な事前ウェイトをお

くという点で DeJong and Whiteman のものとは違う. この反直感的な性質は, Phillips の事前分布作成法の性質から生じるものであり, これは事前分布がモデルの標本理論的性質をまねていることを保証するものである. すなわち, もし ρ が実際に 1 より大きかったとすれば, モデルから生成されたデータはこの系列がすぐに発散する性質があることは明らかになる. したがって, 一般的に 1 より大きい ρ の値はより多くの情報を含んでおり, Phillips の無知の事前分布が表している特性である. これらのモデルに基づいた事前分布を用いて, ベイズ法による事後解析は, 単位根の有無について古典的な有意性検定と同じくらい不確実であることを示した. 加えて, Phillips は Jeffreys 型事前分布を利用した方法を Nelson-Plosser データに応用し, もっと一般的な Jeffreys 型の事前分布を採用し, 同様な結果を得ている DeJong and Whiteman (1989a) や Zivot and Phillips (1991) で報告されたよりは, 単位根モデルに対して決定的な証拠がないことを見い出している.

Phillips の無知の事前分布は, ベイジアンからも非ベイジアンからもともに批判された (Phillips, 1991 に対する討論を参照). この事前分布はパラメータ空間の非定常区域に非常に多くの事前ウェイトをおくので, 非定常である事後確率はそれに応じて増加する. 結果的として, 多くのマクロ経済時系列に単位根があるという事実は, その他のベイズ流の分析で行った研究よりも多くなることを Phillips は見い出した[11]. ある意味では, Phillips の研究はベイジアンと古典的な方法との架け橋と考えることができる. これは公式的にはベイジアンであるが, その基礎にある考え方の多くは古典的な単位根とベイジアンの結果との乖離を説明することに捧げられている. 時系列モデルで合理的な事前分布を導出するものは何かという点については, まだ回答が得られていない. 非常に多くの研究が現在もこの分野で行われつつある.

DeJong and Whiteman (1991a,b), Phillips (1991) それに Zivot and Phillips (1991) によって行われた単位根仮説のベイズ分析では, モデルのパラメータに関して線形である推定された定数項とトレンドをもつ共通の AR(k) モデルを用いている. Schotman and van Dijk (1991a) は Phillips への返答で, 定数項とト

[11] しかし, Phillips の考えた U.S. マクロ経済時系列における非定常性の確率は, 2 分の 1 よりもかなり小さいものである. この特性は, 何人かの研究者にベイジアン単位根の結果が事前密度にかなり変化があってもかなり頑健であると結論づけることになる.

レンドのある AR(k) モデルの非線形の観測されない要素モデルを用いて, モデルのパラメータに関してフラット事前分布と Jeffreys 型事前分布の動きと, 主要な自己回帰係数 ρ の結果として得られた事後分布を調べた. 彼らは, Schmidt and Phillips (1989,1992) におけるように, この要素モデルによる表現は単位根仮説の分析に対してより適切であるとした. それは要素モデルにおける確定項パラメータは, 単位根帰無仮説とトレンド定常対立仮説の下では同じ解釈ができるが, 線形モデルでのパラメータでは同じ解釈ができないからである. しかし, 彼らの見方は DeJong and Whiteman や Phillips によるものとはかなり異なる. 特に, 要素モデルの中ではフラット事前分布と比較した場合, Jeffreys 型の事前分布は単位根モデルのウェイトを小さくすることを示した. そして, DeJong and Whiteman (1991a,b) の結果と対照的に, フラット事前分布から導かれた事後分布の推定は確率 1 で単位根仮説になることも示した. さらに, 単位根モデルを受け入れるという事後オッズ比は, レベルパラメータに非正則な事前分布を用いると発散することを示した. レベルパラメータに対する Jeffreys 型の事前分布は非正則であるから, Schotman and van Dijk は Jeffreys 型の事前分布は「単位根仮説のベイズ分析には不適切である」と批判している.

ベイジアンは古典的な単位根の論文に対して, いくつかの面で批判している (Sims, 1988; Sims and Uhlig 1991; Koop 1992b,c; Koop and Steel, 1994; Schotman and van Dijk, 1991a,b; DeJong and Whiteman, 1991a,b). 単に検出力が低いためだけでなく, それらが単位根でのナイフの刃のような漸近性にかなり依存しているためである. すなわち, 古典的な漸近理論は, もし $|\rho| < 1$ なら標準的な結果が成り立つが, $\rho = 1$ なら結果は大きく異なるという事実に悩まされている. さらに, もし $|\rho| > 1$ で系列が発散するなら, 漸近的な結果は単位根あるいは定常系列の場合とは異なる. もちろん, 小標本では, 定常であるが持続性のある系列 (たとえば, $\rho = 0.95$) から生成された系列は単位根とは実際上区別することができない. これらの論文が示しているように, ベイズテストは厳密な有限標本の結果を与える. この方法は ρ について連続な尤度関数に基づいているから, 古典的な漸近性における不連続性はない[12].

[12] 多分, この問題に対するベイジアンと古典派の間の大きな差を最も分かりやすく示したのは, Sims and Uhlig(1991) である. ここでは, ベイジアンと非ベイジアンそれぞれに関心がある $\rho|\widehat{\rho}$ と $\widehat{\rho}|\rho$ のグラフを提示し, いかにベイジアンのグラフが両者のうちでより合理的かを示して

ベイズ流単位根検定を行う代表的な2つの方法がある. 1つの方法は, ρ に対する報知情報を用いて, 単位根モデルを非単位根モデルと比較する事後オッズを計算する (Koop, 1992b,c; Koop and Steel, 1994; など). もう1つの方法は報知事前分布を用いず, ρ が1の近傍の一定の区域 (たとえば, $P(0.95 < \rho < 1.05|y)$ あるいは $P(|\rho| \geq 1|y)$ に入る事後確率を計算する (DeJong and Whiteman, 1991a, b). これらの考え方と異なる方法が, Tsurumi and Wago (1996) で提案されている. この方法は, DeJong and Whiteman の用いたフラットな事前分布による結果を利用して, 事後密度を帰無仮説の値の方向に引っ張る自然共役事前分布を用いて分析している. ここで, 次の2つの点を強調できる. 1) ベイズ単位根検定は, 必要なのは尤度関数と事前分布であるから導出が容易であり, 厳密な有限標本の結果を与える. 古典的な単位根検定を導出するのに必要な漸近理論は極めて複雑であり, モンテカルロ研究によれば, これらの漸近的な結果を小標本に適応するのは疑問がある. 2) ベイズ単位根検定は, 古典的な検定よりも単位根の存在をずっと小さく見る傾向にある.

古典的な単位根検定とベイズ単位根検定との差は, 用いられるそれぞれ仮説検定の方法の違いに大きく依存している. 古典的な検定方法は単位根を帰無仮説の状態で検定し, 標本情報が圧倒的に否定する場合だけ単位根を棄却する. 逆に, ベイズ事後オッズ比は単位根が存在する確率を計算し, 次にそれを, たとえば系列がトレンド定常である確率と比較する. 小標本では, 単位根モデルから生成されたデータは強い永続性のあるトレンド定常なモデルから生成されたデータと事実上区別できないであろう. 強い永続性を示し, 時間がたつにつれて上方トレンドがある, あるいはその両方がある有限の標本の系列の場合には, ベイズ法は単位根とトレンド定常モデルが似た事後確率をもつことを明らかにする. しかし, (圧倒的な標本情報に直面する場合だけ棄却する) 古典的な検定方法では, 単位根という帰無仮説を棄却できないだろう. 別の方法で示すと, ほとんどのマクロ経済時系列が単位根を含んでいるという事実は, 古典的な検定が棄却できないことを確実に保証する. ベイジアンはトレンド定常というようなもう1つの仮説よりも単位根 (あるいは非定常性) のほうがどれほど起こりうるかと尋ね, 否定的に答えがちである. 簡単に言うと, ベイズ法は古典的な受容/棄却

いる.

という検定方法よりも, 標本情報をより適切に要約している.

EBA(Extreme Bounds Analysis : 極値境界値分析)

多くのベイズ分析を用いている人たちはモデル選択の古典的方法を批判しており, たぶん最も強いものは Ed Leamer の研究 (1978,1983,1985) から生じたものであろう. Leamer が行っている批判は, 経済学は実験科学ではないのだから, **真のデータ生成プロセス**を見つけることができると考えることは間違いである, というものである. 結果として, 主要な推定がモデルの定式化に敏感なのか, あるいは他の仮定に対して敏感なのかを確かめることは重要になる. 線形回帰モデルでは, 感度分析のパフォーマンスに対するこの一般的な名称は, EBA として知られている一連のテクニックによって適応される.

EBA は明確な実証分析での成功例であるが, これは一部は疑いもなくそれがベイズ分析としても古典的分析としても解釈できるという事実のためである. このことを理解するために, 正規誤差をもつ次のような線形回帰モデルを考える.

$$y_t = \boldsymbol{x}_t'\boldsymbol{\beta} + \varepsilon_t \qquad (t = 1, \ldots, T)$$

ここで $\boldsymbol{x}_t$ は $k \times 1$ の説明変数ベクトル, そして ε_t は $iid.\ \mathcal{N}(0, \sigma^2)$ である. 事前分布は自然共役で, $\boldsymbol{R}$ と $\boldsymbol{r}$ はそれぞれ既知の $J \times k$ と $k \times 1$ 行列で $\mathrm{rank}(\boldsymbol{R}) = J$ とする. アプリオリに $E(\boldsymbol{R}\boldsymbol{\beta}) = \boldsymbol{r}$ を仮定するが, 事前分散 V_0 は定式化しないままである. もし, $\boldsymbol{\beta}$ の線形結合 ($\boldsymbol{\phi}'\boldsymbol{\beta}$, ここで $\boldsymbol{\phi}$ は $k \times 1$) に関心があれば, この線形結合の論理的推定値は $E(\boldsymbol{\phi}'\boldsymbol{\beta}|y)$ となる. しかし, この例では事前分散を定式化できないし, しようとしない. Leamer は $E(\boldsymbol{\phi}'\boldsymbol{\beta}|y)$ がとりうる極値を計算することを提唱している. すなわち $\min_{V_0} E(\boldsymbol{\phi}'\boldsymbol{\beta}|y)$ と $\max_{V_0} E(\boldsymbol{\phi}'\boldsymbol{\beta}|y)$ を解くことを考える. 2 つの解は, $E(\boldsymbol{\phi}'\boldsymbol{\beta}|y)$ の事後平均がとる極値になる. この差 δ を計算し, パラメータの推定値の曖昧さの尺度としようとするものである. δ が大きければ, 推定値のとりうる範囲が広がることを意味している. 言い換えると, EBA は研究者はパラメータの事前の位置についてある情報をもっているが, あえて事前分散行列を定式化しようとしない, 共通の場合に対する事前分布の感度分析として解釈できる.

EBA の古典的な解釈も直感に訴え, 納得できる. 説明変数は 2 つのグループ, **自由**変数と**候補**変数, に分割される. 自由変数は回帰に必ず入る変数であり, こ

れらの変数の係数にどんな制約もおかないことに確信を持っている. それに対して, 候補変数は推定に必須ではないが方程式の適合度に寄与する変数である. どんな候補変数も $\boldsymbol{R}\beta = \boldsymbol{r}$ の形の任意の線形制約に従うとさせることができる. たとえば, 貨幣需要関数では利子率を自由変数として選び, その他のたとえば, 金融イノベーションに対する代理変数, ラグ付き変数を, ゼロ制約に従う候補変数とする. Leamer は, これは基本的には多くの応用研究者が通常行っていることであり, 通常 2,3 のキー変数に関心があり, その他の実験をおこなうと述べている. しかし, この実験から唯 1 つの最終的な回帰式を示すだけよりも, Leamer は, **重要な変数と考える**キーパラメータのどれかが起こる極値を示すべきであると述べている. 簡単には, Leamer は候補変数に課される線形制約に対する重要パラメータ推定値の感度を調べ, これらの極値境界が上のベイズ事前感度問題におけるものと正確に同じになるかどうかを調べる.

EBA はいくつかの実証分析で行われている. たぶん最も重要なものは Cooley and LeRoy (1981) であろう. この研究は McAleer, Pagan and Volker (1985) によって反論され, ふたたび Cooley and LeRoy (1986) と Leamer(1985) によって弁護されている. 彼らの研究では, Cooley and LeRoy (1981) は EBA を貨幣需要関数に適用し, 利子率弾力性に対する EBA は非常に広く, ゼロを含んでいるので, いかなる推測も不安定なものであるとの結論を得た. たとえば, 利子率の符号が負で推定されるのは, 候補変数を都合よく組み合わせて得られるものであり, データに含まれている情報だけでは利子率が負である結果は得られない. したがって, 推定されている関数の多くは, 分析者の強い先験的情報でゆがめられた結果である可能性が大きいことを明らかにしている. McAleer, Pagan and Volker (1985) はこの見解に様々な理由を挙げて反論した. その多くはここで関心はない. 今の目的のためには, 彼らが EBA が候補変数の有意性に依存しているだけでなく, 異なる分類が異なる EBA を意味するという意味で, 変数を自由と候補のカテゴリーに分けることに依存している範囲の大きさを強調していることを注意するだけで十分である. これらは Leamer(1985) や Cooley and LeRoy (1986) が厄介な問題とは見なさなかった性質である. 一般的に言うと, これらの性質は十分考慮して遂行された古典的分析とベイズ的な方法を用いた分

析との間の本質的な類似性を強調するものと考えることができる[13].

Leamer は SEARCH と呼ばれるコンピュータ・パッケージを作成し, 彼が提唱した多くの手法が取り入れられている.

[13]EBA の分析例については伴 (1991), 163-165 を参照のこと.

第2章　ベイズ計量分析における計算方法

2.1　ベイズ法における計算

第2章で強調したように, 事後平均とか事後オッズのような統計量を解析的に求める方法は滅多に存在しない. そのような場合には, ベイズ流の応用研究では数値的な手法に頼らなければならない. そこで, そのいくつかの方法について述べる. まず最も単純な近似から始め, 次にもっと手の込んだ方法を考えることにする.

事後分布を求めるには, 複雑で, しばしば高次元の積分を評価することが必要になる. 加えて, 事後分布の要約統計量の計算にはモーメントや四分位値が含まれており, そのためにはさらに積分計算が必要になる. この問題に対する初期の取り組みは事後密度への解析的な近似解を得るために漸近的な方法を用いることであり, 最も単純な結果は事後分布への正規近似で, 基本的には中心極限定理のベイズ版といえる. もっと複雑な漸近的な方法はラプラス法 (Tierney and Kadane, 1986) の様な手法で, もう少し正確で, かなり非対称な事後分布への近似ができる. このような近似法が適用できないかあるいは精度が不十分である場合には, 数値積分に頼ることになる. 何年もの間ガウス求積法のような伝統的な方法がこの目的のために使われてきた[1]. しかし, この方法を適用するのは, 対象とする計算式が大きなモデルでは扱えないので, (パラメータ数が10位までの) 低次元の計算に限られてきた. これらの方法には, 一定の精度を維持するために必要な関数を評価する数が, 次元数とともに指数的に増加するという次元の呪いという問題がある. もう一つの方法は, Dempster, Laird and Rubin (1977) のEMアルゴリズムである. 本来は欠測値問題における最尤推定値を見つける

[1] これらの方法についての一般的な議論については Thisted (1988, Ch.5) を, また事後分布の計算への応用については Naylor and Smith (1982) を参照.

道具として開発された方法であるが, 多くのベイズ分析での事後分布のモードを見つけるのに用いられている. EM アルゴリズムは多くの実際問題に適用されているが, かなり遅いアルゴリズムであり, 事後分布全体を推定するよりも主として事後モードを見つける目的だけに向けられている. 結果として, 最近の多くの応用分析家はモンテカルロ積分法に注目するようになった. この方法はより完全な情報を提供し, かなり高次元のモデルでも比較的プログラムが簡単であるという特徴がある. この方法には, インポータンス・サンプリング (Geweke, 1989a) や単純な受容/棄却サンプリングのような伝統的な非反復的方法が含まれる. もっと強力な方法は, 最近開発された反復的マルコフ連鎖モンテカルロ法で, データ拡大 (Tanner and Wong, 1987), メトロポリス-ヘイスティング・アルゴリズム (Hastings, 1970), そしてギブス・サンプラー (Geman and Geman, 1984; Gelfand and Smith, 1990) が含まれる. これらの方法はマルコフ連鎖を作成し, そのアウトプットは同時事後分布からの (相関のある) サンプルに対応する. これらの方法について, 以下順に紹介する.

2.1.1 正規近似

多くの場合, 事後密度は近似的に正規分布にしたがう. 正規分布の性質はよく知られているので, 事後密度から近似的な推測を行うのは容易である. 説明上, そのモードの回りで事後密度の対数をテーラー展開する.

$$\ln p(\theta|y) = p(\widehat{\theta}|y) + (\theta - \widehat{\theta})' G(\widehat{\theta}|y) + \frac{1}{2}(\theta - \widehat{\theta})' H(\widehat{\theta}|y)(\theta - \widehat{\theta}) + \text{残余}$$

ここで $G(\theta|y)$ は θ で評価した傾斜, $H(\theta|y)$ は θ で評価したヘッセ行列である. モードで評価した傾斜はゼロであることに注意し, 残余を無視すると, 事後密度は近似的に正規分布にしたがうことが分かる. 実用的には, これは近似的なベイズ分析はちょうど古典的な計量経済分析で最尤値を求めることに対応している. 標準的な最適化のパッケージであれば事後モードを得るのに使うことができ, 対数事後密度のヘッセ行列の逆行列にマイナスを付けたものを近似的な事後共分散行列として用いる. たとえば, Zellner and Rossi (1984) はロジットモデルでの研究にこの近似を採用している. しかし, 彼らは正規近似が極めて良くないことを見い出している.

正規近似に関する問題は, 近似誤差が非常に大きいことである. それはこの方法が事後平均に実際は事後モードで近似しているためである. もし, 事後分布がかなり歪んでいるか双峰ならば, 結果は惨めである. また, パラメータの非線形関数である量に関心がある場合は, 正規近似はそれほど役に立たない. さらに, この方法は近似の精度を評価するための診断が得られない. 可能なら, 次に述べる方法を用いてもっと正確な分析をするべきであろう.

2.1.2 数値積分

ベイズ実証分析に通常見られるタイプの積分を計算する最も直接的な方法は, 多くの数値積分公式の中の 1 つを用いることである. これらの手法の中で最も単純なものはシンプソンの公式を用いており, 多少手の込んだものはガウス・エルミート球積法を用いる. これらの手法は, 標準的な汎用数値積分プログラムが低次元の問題に対してだけ有効であるという重大な欠点がある. しかし, 計量経済分析では通常多次元の問題を取り扱う場合が多い. Naylor and Smith (1982, 1988) と Smith et al.(1985) は数値積分の全般的な議論をしており, 9 次元までの数値積分を行っているが, プログラムは複雑で, 個別の問題に対するものであるので, 経験のあるコンピュータ・プログラマーだけが使いこなせるものである. しかし, 1 つか 2 つのパラメータに関する積分が必要な応用問題に対しては, 数値積分は実行可能である. たとえば, BRAP[2] は, 1 次元の問題に対してはシンプソンの法則に基づいた簡単なサブルーチンがあり, IMSL や NAG ライブラリーは Fortran プログラム[3] から直接呼び出せるいくつかのサブルーチンを含んでいる. 多分最も手の込んだプログラムは Bayes Four であり, 数値積分を行うための広範囲のサブルーチンを含んでいる (Naylor and Shaw (1991) 参照).

2.1.3 ラプラス近似

ベイズ分析では積分を近似するためにラプラス法を用いることができる. Tierney and Kadane (1986), Tierney, Kass and Kadane (1989), そして Kass, Tierney

[2] BRAP はシカゴ大学の Arnold Zellner を中心としたグループによって考案された TSP に似たベイズ計量経済分析用のコンピュータパッケージである. このプログラムの詳細は付録を参照のこと.

[3] IMSL と NAG ライブラリーは, ほとんどの大学で利用できる Fortran サブルーティン・ライブラリーである.

and Kadane (1988) による研究によって, ベイズ計算に役立つ解析的な近似を得ることができるようになった. 特に, 実数値あるいはベクトル値パラメータや周辺事後密度の事後平均と分散を近似する積分にラプラス法を適用した. この方法を用いると, ベイズ分析は積分 $\int f(\theta)\exp[-Th(\theta)]d\theta$ (スカラーの θ に対して) を

$$f(\theta)\sqrt{\frac{2\pi}{T}}\sigma\exp[-Th(\widehat{\theta})]$$

によって近似する. ここで θ は $h(\theta)$ の最大値, そして σ^2 は θ で評価したヘッセ行列の逆行列にマイナスを付けたものである. ほとんどのベイズ積分はこの形を取る. たとえば Tierney and Kadane (1986) は

$$E(g(\theta)|y)=\frac{\int\exp(TL^*)d\theta}{\int\exp(TL)d\theta}$$

と表している. ここで $L=[\ln(p(\theta))+\ln(L(\theta|y))]/T$, $L^*=L+\ln[g(\theta)]/T$ そして T は標本のサイズである. ラプラス法を分子と分母に適用すると, 両者とも非常に似ているので近似誤差も非常に似ており, そこで互いに相殺することが分かる. 結果的に, Tierney–Kadane 法を用いて計算された積分は通常極めて正確である.

多変量の場合には, 次のように $E(g(\theta)|y)$ に対する 2 次近似値を得る.

$$\widehat{g}=\frac{|\Sigma^*(\theta^*)|^{1/2}\exp[TL^*(\widehat{\theta}^*)]}{|\Sigma(\theta)|^{1/2}\exp[TL(\widehat{\theta})]}$$

ここで θ と θ^* は最大値であり, それぞれ L と L^* に対応する Σ と Σ^* は, ヘッセ行列の逆行列にマイナスを付けたものである. Kass, Tierney, and Kadane (1989) で示されているように, この近似の相対誤差は T^{-2} のオーダーで上の正規近似よりもかなり良い. 対応する $g(\theta)$ の分散の 2 次近似は次のようになる.

$$\widehat{\mathrm{V}}\,(g(\theta)|y)=\widehat{E}(g(\theta)^2|y)-[E(g(\theta)|y)]^2$$

Tierney–Kadane 法を行うには, 関数を最適化することだけが必要である. 特に L と L^* は最大化され, そのヘッセ行列を得る必要があるが, 多くの最適化パッケージが存在するのでわざわざプログラムする必要はない. 最尤法をよく使う応用分析を行う人にとっては, 積分問題を 2 つの最適化問題に変えるので Tierney–Kadane 法は簡単に使える. そして L と L^* は似た関数であるから, も

し $\widehat{\theta}$ を $\widehat{\theta}^*$ の探索での初期値として使えば計算が容易になる. Tierney–Kadane 法は, たとえ $\widehat{\theta}^*$ が $\widehat{\theta}^*$ の方向で $\widehat{\theta}$ からの 1 ニュートン・ステップによって置き換えられても, その高い精度を漸近的に維持する. この方法は, Tierney–Kadane の 1 段階法として知られている. 簡単には, (いくらか精度を犠牲にしても)2 段階目の最適化を行う必要はない.

Tierney–Kadane 法に関しては, 4 つの欠点を指摘しておく必要がある. 第 1 に, この方法は計算したい特性値に対して繰り返さなければならない. したがって, もし 20 個のパラメータの事後平均と分散を計算する場合には, 41 回の通常時間がかかる最適化を行わなければならない. 第 2 に, この方法は $g(\theta|y)$ の対数をとるので正の関数に限定される. しかし, この制約は大きな定数を $g(\theta|y)$ に加えるか, (常に正である) 積率母関数を用いてそれを微分することによって回避することができる. 第 3 に, この方法は, 近似の精度を診断する方法がない. 第 4 に, この方法の近似の質を改良する方法がない. 事実, たとえ Tierney–Kadane 近似の精度が当面の応用問題にとってはあまり良くないということを気付いていたとしても, 全く別の方法を考えることができない.

まとめると, Tierney–Kadane 法は少数の重要な積分を速く近似する非常によい方法である. しかし, 10 個以上の多数の積分を計算しなければならないような大きな問題に対しては, モンテカルロ積分の方が好ましい方法である.

例：ロジットモデル

Koop and Poirier (1993) は, 1982 年の社会調査を用いて 337 人の雇用されている男性に対して, 5 つの職業カテゴリー間での選択が 3 つの特性に依存している多対ロジットモデルを考えた. 次の多対ロジット尤度が仮定される.

$$L(y|\theta) \propto \prod_{i=1}^{T} \prod_{j=1}^{5} P_{ij}^{y_{ij}}$$

ここで, 個人 i が職業 j を選択すると $y_{ij} = 1$ (その他の場合は $y_{ij} = 0$), そして $P_{ij} = \text{Prob}\,(y_{ij} = 1)$ である. P_{ij} は, 次のように定式化される.

$$P_{ij} = \frac{\exp(z_i'\theta_j)}{\sum_{j=1}^{5} \exp(z_i'\theta_j)}$$

ここで, z_i は個人 i についての 4×1 のデータベクトル (経験, 教育, 人種の 3 つの特性値と定数項), そして θ_j は 4×1 のパラメータのベクトルである[4]. この尤度関数をフラットな事前分布と結びつける. すなわち, 事後密度は尤度関数に比例する. 目的はいくつかの特性を持つ個人が職業につく確率を予測することである. 言い換えれば, f を特性 z_f をもつ個人の指標とした場合に, P_{fj} を計算しようとした. これらの予測された確率は, ロジットパラメータ θ の非線形関数である.

最尤推定量に対して closed form の解は存在しないことはよく知られているので, 最尤推定値を計算するためには最適化プログラムが必要になる. さらに, 一般にこの尤度関数を含む積分に対する解析的な解は存在しないことは容易に証明できる.

そこで, 予測された確率の事後密度を評価するのに, Tierney–Kadane の方法とその 1 段階法 (One-Step methods) の両方を用いている. 前者の方法では, ロジット事後密度の対数 (Tierney–Kadane の記号では L) を最適化し, 次に L^* のその後の最適化に対する初期値としてこの最適化の値を用いている. L^* は計算された各事後平均, すなわち $E(P_{ij}|y)$ で異なるから, 多くの最適値が必要になる. モンテカルロ積分の結果と比較すると, これらの両方の近似の精度はこの例では非常によく, 必要な計算は最小ですむ.

2.2 モンテカルロ積分

数値解析の最も簡単な方法は, モンテカルロ・サンプリングに基づく方法である. モンテカルロ積分は, ベイズ実証分析で多分最も広く用いられている手法である. 計量経済学への導入は, ほとんど Kloeck and van Dijk (1978) と Geweke(1989a) の研究を通じで行われている. モンテカルロ・サンプリングは積分を直接に数値的に評価するか, あるいは数式を用いて解析的に求めた解の代わりに妥当な標本空間における多数の点で構成される確率分布に依存している.

この方法は, 次のようないくつかの利点がある. それらは 1) 高次元問題を扱うのに適している, 2) 目標に近似する際の誤差のサイズをコントロールできる,

[4] θ_1 はゼロに正規化されている.

3) プログラムが簡単である. しかし, モンテカルロ法を用いた積分計算がコンピュータ時間の大半を占める場合が多い.

2.2.1 単純なモンテカルロ積分

単純なモンテカルロ積分は, もし n 個の確率変数 $(\theta^i, i = 1, \ldots, n)$ を事後分布から抽出すると

$$g^* = \frac{1}{n}\sum_{i=1}^{n} g(\theta^i) \to E(g(\theta)|y) \qquad n \to \infty$$

となるという事実に基づいている. この単純な定理に基づくコンピュータ・プログラムは非常に簡単に書くことができる. 必要なことは, 事後分布から繰り返し乱数を抽出し, これらの乱数それぞれに対して必要な統計量を計算し, その各統計量を抽出回数で平均することである. 近似の精度は n に依存し, それをコントロールすることができる. n についての漸近理論を用いると, 数値的な精度の尺度を導くことができる. すなわち, 弱い正則条件の下で, g^* は漸近的に, 平均 $E(g(\theta)|y)$ と分散 σ^{*2}/n の正規分布にしたがう. ここで, $\sigma^{*2} = E(g(\theta)^2|y) - [E(g(\theta)|y)]^2$ である. σ^{*2} は

$$s^{*2} = \frac{1}{n}\sum_{i=1}^{n} g^2(\theta^i) - g^{*2}$$

によって推定することができる. $E(g(\theta)|y)$ に対する 95%信頼区間は, $[g^* - 1.96s^*/\sqrt{n}, g^* + 1.96s^*/\sqrt{n}]$ となる. ここで $s^*/\sqrt{n}$ を**数値標準誤差** (NSE) と呼ぶ. 漸近理論によれば, $g(\theta)$ の標準誤差の NSE に対する比は $1/\sqrt{n}$ を意味する. したがって, もし $n = 10,000$ なら, 漸近理論から NSE はその関数の事後標準偏差の 1%になることを意味する.

対称反復法

数値近似の効率を増加するために, モンテカルロ・サンプリングを僅かに変形するのが対称反復法である. もし事後密度が近似的に対称であるなら, 一定の精度を達成するのに必要な抽出回数は, Geweke (1988) の提案した**対称加速法** (*antithetic acceleration*) を使えばかなり減らすことができる. この方法は確率分布のモーメント自身の数値評価に対する方法ではなく, 標準的なモンテカルロ・サンプリングに加えられる方法である. この手法は次のようにペアにし

てサンプルを抽出する. モードの一方でのそれぞれの抽出と比例して, 逆の方向で1つ抽出する. もっと正確には, もし事後分布のモードが μ で i 番目の抽出が θ^i であったとすると, 対称加速法は $(i+1)$ 番目の抽出を $2\mu - \theta^i$ とする. Geweke(1988) によれば, この方法は数値的な精度を大きく増加させ, したがって, 特に事後分布への漸近的な近似が正確ならば n を大きく減少させることができる. このように対称反復法は, 確率分布やそのモーメント自身の数値評価に対する一つの方法ではなく, 標準的なモンテカルロ・サンプリング図式に加えることができる方法である. 対称反復法を導入するには, 前述の基本プロセスに1ステップを加え, 事後分布内の点が分布の平均の回りに公平に分布する様にする. 正の分布から確率抽出された後で, 対称反復法は平均を通じて投影した直近の抽出のミラーイメージを作成する. このより早い率で真の値へ収束する数値近似によって数値近似の分散の減少は, サンプルサイズがかなり大きくなると計算効率がかなり向上する. 標準的なモンテカルロ・サンプリングで適切な数値精度を保証するのに $n = 5000$ と $n = 10,000$ が通常選択されるが, もしデータセットが $T = 100$ のサンプルサイズなら, 500 と 1000 の間の対称確率抽出 (250 から 500 組) で一般に同じ程度の数値精度が達成できる. 対称反復での数値精度の正確な増加は, 関数 $g(\theta)$ の非線形性の程度に依存し, 非線形性が増加するにつれて利点が減少する.

例：単位根検定

時系列モデルでは, 回帰パラメータの関数にしばしば注目する場合がある. 例を用いて, DeJong and Whiteman (1991a,b) で述べられているベイズ単位根検定の方法を考える. DeJong and Whiteman は次のモデルを使っている.

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \beta_4 t + \varepsilon_t$$

ε_t は平均がゼロで分散が σ^2 の *iid.* 正規, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)'$, そして非報知事前分布 $p(\boldsymbol{\beta}, \sigma^2) \propto \sigma^{-2}$ が仮定されると, $\boldsymbol{\beta}$ に対する周辺事後密度は t 分布に従う. したがって, $\boldsymbol{\beta}$ についての推測は多変量 t 密度の性質を用いて極めて簡単に行える. しかしながら, DeJong and Whiteman は $\boldsymbol{\beta}$ 自体には関心がなく, 自己回帰の dominate(支配的) な根 Λ に関心がある (Λ は自己回帰表現で絶対値で

の最大の根として定義される). DeJong and Whiteman は $p(\Lambda \geq 0.975|y)$ のような数値を計算した. 残念ながら, これらの数値は β の高度な非線形関数であり, したがって解析的な結果は存在しない. しかし, $p(\beta|y)$ からのサンプリングは簡単であるから, 簡単なモンテカルロ積分を用いることができる.

事実, t 事後分布は, 1) β の最小 2 乗推定値である平均, 2) 最小 2 乗推定から得られる統計量である分散, そして 3) 自由度, すなわち $T-1$, で決められる. これらの統計量はどんな計量経済分析用のコンピュータ・パッケージからでも簡単に得ることができ, 乱数は t 事後密度からいくつかの方法で得ることができる. BRAP のようなプログラムでは, もし上の 1), 2), 3) を与えれば自動的に t 密度から乱数を得ることができる. 別の方法として, 応用経済分析を行う人は, (たとえば, GAUSS のような) ほとんどのパッケージに組み込まれている正規乱数発生機能を使って t 乱数を簡単に作成することができ, 次に正規分布と t 分布との関係を利用して t 密度に変換する.

t 乱数の場合では, モンテカルロ積分は事後分布から乱数 β^i を発生し, Λ を計算する (1 回毎の抽出から Λ を計算する). Λ に関しての平均は $E(\Lambda|y)$ に収束し, 0.975 以上の Λ^i の割合は $p(\Lambda \geq 0.975|y)$ に収束する.

DeJong and Whiteman は, 単位根の存在を証明するのに $p(\Lambda \geq 0.975|y)$ のような値を使っている. しかし, モンテカルロ積分は自己回帰モデルやベクトル自己回帰 (VAR) モデルでのその他いろいろな統計量を計算するのに使うことができる. もし事前分布が自然共役あるいはフラットであれば, 事後密度は多変量 t あるいは行列変量 t になり, 事後密度から乱数を抽出することができる. VAR に対するこの種のテクニックは, Doan (1988) と Koop(1992c) を参照すると良い.

2.2.2 インポータンス・サンプリングを用いるモンテカルロ積分

乱数が事後分布から簡単に得られない場合は, インポータンス・サンプリングを用いたモンテカルロ積分が必要になる. 単純なモンテカルロ積分は事後分布から直接乱数を抽出し, 次に抽出した数で単純平均をとる方法であるが, 他方, インポータンス・サンプリングを用いるモンテカルロ積分は, インポータンス関数と呼ばれる代替密度から乱数を発生し, 次にその抽出された乱数の加重平均

をとる方法である.

たとえば, 事後密度の事後期待値を近似する場合を考える.

$$E(f(\theta)|y) = \frac{\int f(\theta)L(\theta)p(\theta)d\theta}{\int L(\theta)p(\theta)d\theta}$$

正規化された尤度に事前密度を掛けた $cL(\theta)p(\theta)$ を簡単にサンプルできる, たとえば, 多変量 t 密度やスプリット t のような密度[5] $g(\theta)$ によってラフな近似をすることができる (Geweke, 1989a を参照). ここで $g(\theta)$ は $cL(\theta)\pi(\theta)$ に近ければそれだけ近似がよい. 他方, $g(\theta)$ への近似が悪ければ多くのウェイトはゼロに近くなり, 僅かな θ^i によって合計が変動し, 不正確な近似になる.

$g(\theta)$ をインポータンス関数とし, それから簡単に乱数を発生することができる. θ^i を $g(\theta)$ からの乱数とし, ウェイト関数を $w_i = p(\theta^i)L(\theta^i|y)/g(\theta^i)$ とする. 非常に弱い条件の下で[6]

$$g^* = \frac{\sum_{i=1}^{n} w_i f(\theta^i)}{\sum_{i=1}^{n} w_i} \to E(f(\theta)|y) \qquad n \to \infty$$

を示すことができる. この定理は一見するとすべてのインポータンス関数に対して成り立つように見える. しかしながら, 現実にはもし適切なインポートタンス関数が選択されなければ, n はかなり大きくなる. この選択に対する良い経験則は, インポータンス関数を事後密度関数に近似するか, 少なくとも事後分布より裾を厚くとるべきであるということである. もし事後分布の裾がインポータンス関数のそれよりも厚ければ, 推定は非常に不安定である. それは, インポータンス関数の裾のずっと端の方でサンプリングされたときを考えると分かる. もし事後分布の裾がインポータンス関数のそれよりも厚かったら, このサンプリングに対する w_i は非常に大きくなり, したがって推定値に非常に大きなウエイトを与える. 言い換えると, 裾の方での滅多にないサンプリングが $E(f(\theta)|y)$ のモンテカルロ推定値の動きを支配する傾向になる.

ベイズ分析で共通によく使われるのは, 前述の t インポータンス関数である. t 分布は3つのパラメータをもっており, 平均, 分散, そして自由度である. 関数

[5] スプリット-t 密度は, 非対称な場合に認められる以外は通常の t と同じである. 定式上は, モードの一方での分散は反対側では異なる場合がある.

[6] 基本的には, これらの条件は $g(\theta)$ のサポートが $p(\theta|y)$ のサポートをカバーしていることが必要である.

の最適化は比較的易しいので, インポータンス関数のパラメータを選択するためには, 事後分布の対数を最適化すればよい. ほとんどの最適化パッケージは自動的に事後モードを計算し, モードでのヘッセ行列を計算する. インポータンス関数の平均と分散は, それぞれ事後モードと対数事後分布のマイナスのヘッセ行列の逆行列に等しいと置く. ほとんどの応用分析で, これはインポータンス関数の形状が事後分布のそれと似たものであることを保証する. 自由度のパラメータは, インポータンス関数の裾が事後分布のそれよりも厚いことを保証するために, 裾の動きをコントロールするのに使うことができる.

ところで, t 分布が事後分布の十分な近似にならない場合には, 解析的な方法かあるいは等高線プロットのような数値的な道具のどちらかを用いて事後分布を注意深く調べるしかない[7]. しかし, 多くの応用研究では事後分布は近似的に正規であり, 正規あるいは t インポータンス関数を用いると十分満足するものが得られる. さらに, 対称加速法を使うことによって必要なサンプルの数を減少させることもできる.

ひとたび1つのインポータンス関数が得られれば, その質を評価してモンテカルロ計算の実行毎の推定値の安定性を調べるための最も単純な**アドホックな方法**が必要になる. すなわち, 10,000 個のサンプリングを行う 2 つのモンテカルロ計算を行って, それらが非常に似た推定値を得たなら, これがサンプリングの数が高次の数値的な精度を達成するのに十分である良い証拠になる. 同様に, 事後分布に十分よく近似するインポータンス関数は, 大体コンスタントになるウェイト w_i になる. そこで, ベイズ分析でインポータンス関数が適切であるかどうかは, 1) ウェイトを調べる, 2) これらの変動係数を計算する, 3) これらのウェイトの分布のグラフをプロットする (Hop and van Dijk, 1990), のいずれかを行う.

あるいはもう 1 つの別の方法として, Geweke (1989) による漸近理論を用いたインポータンス関数の診断を行うことができる. 前節で紹介した Geweke が用いた最も単純な診断法は, 数値標準誤差 (NSE) である. この場合 NSE の推定

[7] たとえば van den Broeck, Koop, Osiewalski and Steel (1994) の研究では, 事後分布の詳しい分析によってガンマ・インポータンス関数が適切であることが示されている.

値は, 次式で与えられる.

$$\mathrm{NSE} = \sqrt{\frac{\sum_{i=1}^{n}[g(\theta^i) - g^*]^2 w_i^2}{\left[\sum_{i=1}^{n} w_i\right]^2}}$$

もう1つの有望な数値は相対的数値有効性 (RNE) であり, Geweke は次式で推定した.

$$\mathrm{RNE} = \frac{\mathrm{V}\,(g(\theta)|y)}{n^*\mathrm{NSE}^2}$$

$\mathrm{V}\,(g(\theta)|y)$ はモンテカルロ法によって通常は計算されるので, RNE は NSE を一度計算してしまえば簡単に得ることができる. RNE の大きい値が望ましい. すなわち, もし RNE = 1 なら, インポータンス関数は事後分布から直接サンプリングして得た結果と同程度の精度を持つ結果を得る. さらに, NSE は $(\mathrm{RNE}^* n)^{-1/2}$ と事後標準偏差の積である. したがって, RNE = 0.1 で $n = 1{,}000$ とすると, 数値標準誤差は事後標準誤差の 10%になる.

事後オッズの計算には, $E(f(\theta)|y)$ の形にはできない $\int \pi(\theta)L(\theta|y)d\theta$ の計算が必要である. しかし, Geweke (1989) は次式になることを示した.

$$\int \pi(\theta)L(\theta|y)\ d\theta \doteq \frac{1}{n}\sum_{i=1}^{n} w_i$$

すべての正規化定数は事後オッズを計算する場合には含める必要がある. あるいは, 少なくとも分析の最後で再び加える必要がある. すなわち, $\int p(\theta)d\theta = 1, \int \ell(\theta|y)d\theta = 1$, そして $\int g(\theta)d\theta = 1$ が必ず保証される.

インポータンス・サンプリングの計算

インポータンス・サンプリングは, ガウス型尤度関数を持つが共役でない報知事前分布をもつ共通の回帰モデルを分析するのに有効な方法である. そこで, 事後密度は非正則で, そこから確率抽出が困難か不可能である. インポータンス・サンプリングの計算アルゴリズムは, まとめると次の通りである. $f(\theta|y)$ を y が与えられたときの条件付き確率密度関数としたとき, もし $p(\theta)$ から確率変数を簡単に発生することができるなら, $h(\theta) = \int f(\theta|y)p(\theta)d\theta$ を求めるために

1. $p(\theta)$ から互いに独立な n 個の標本 $y_1, \cdots, y_m$ を発生させ

2. $\widehat{h(\theta)} = \frac{1}{m}\sum_{i=1}^{m} f(\theta|y_i)$ を計算する

$h(\theta)$ からの標本を発生させるには, $\widehat{h(\theta)}$ の代わりに $f(\theta|y_i)$ から θ_i を発生させる. しかし, $p(\theta)$ から標本を発生させることが難しい場合には, この方法は適用できないので Geweke(1989a) は $p(\theta)$ に近い, 標本を発生させやすいインポータンス分布と呼ばれる分布 $g(\theta)$ を用いて

(1) $g(\theta)$ から互いに独立な m 個の標本 $y_1^*, \cdots, y_m^*$ を発生させ
(2) $\widehat{h(\theta^*)} = \sum_{i=1}^{m} w_i f(\theta|y_i^*) / \sum_{i=1}^{m} w_i, w_i = p(\theta_i^*)/g(\theta_i^*)$ を計算する

とすると, $g(\theta)$ のサポートが $p(\theta)$ のサポートを含むとき, 確率 1 で $\widehat{h(\theta^*)}$ が $\widehat{h(\theta)}$ に収束することを証明した. $h(\theta)$ からの標本を発生させるには (2) の代わりに

(2)' $f(\theta|y_i^*)$ から θ_i を発生させる $(i = 1, \cdots, m)$
(3) $(\theta_1, \cdots, \theta_m)$ に確率 $(w_1, \cdots, w_n)$ を与えて, 標本 $(\theta_1^*, \cdots, \theta_n^*)$ を発生させる

とすればよい. Smith and Gelfand(1992) はこれを重み付ブートストラップと呼び, m が大きくなるに従って $h(\theta)$ から発生される確率標本に対する近似がよくなることを証明した.

インポータンス密度の選択は, 数値近似と計算効率の両方で重要である. 2 つの密度が似ていれば, それだけ数値近似がよくなる. インポータンス密度の選択で最も重要な点は, インポータンス密度の裾のカバレッジが少なくとも真の事後分布の裾のカバレッジに等しいか, そうでなければ少しだけ大きいことを保証することである. すなわち, インポータンス密度は θ の事後密度よりはいくらか裾が厚い. これによって, 事後密度の裾の方では標本をとりすぎることになるが, これらの観測値はインポータンス・ウェイト関数 $w(\theta)$ によって小さくウェイト付けされる. もし裾が非常に薄ければ裾のカバレッジが不適切であり, これらの観測値に過剰にウェイトを与え, 数値サンプリングをおこなう時に標本空間の全領域を完全に覆えない (Geweke, 1989a).

そのために, 通常適切なカバレッジを保証するためにスケールアップした分散・共分散行列のインポータンス密度を用いる (Tanner, 1996, p.56). たとえば, n 個の観測値と k 個の外生変数のある単一方程式回帰モデル $y = X\beta + \varepsilon$, $\varepsilon \sim$

$\mathcal{N}(0,\sigma^2 I)$ を仮定し, ここで $\theta=(\beta,\sigma^2)$ に対する事前分布は $h(\beta)/\sigma$ に比例する. すると, 回帰係数 β には報知事前情報を, そしてスケールパラメータには標準的な散漫な事前分布をもつ. データの尤度は標準多変量正規分布であり, $h(\beta)$ が一定なら事後分布も $\theta=(\beta,\sigma^2)$ に対して多変量正規-逆ガンマ型であるような標準的な形になる (Zellner, 1971). しかし, 多くの経済分析への応用の場合では, $h(\beta)$ は β の標本空間を越えて一定であることは滅多にない. 経済理論によってある点で切断されるかもしれないし, 正の値だけしかとらないかも知れない. また, 以前に行われた研究から得られた知識によって特別な情報を持つこともある. 多くの場合, インポータンス・サンプリングは β あるいは $g(\beta)$ の事後分布を調べるために用いる最適な方法である. データの尤度関数は, 事前分布の形によって尤度関数や事後分布が (事前分布のサポートがある標本空間の区域内で) 極端に異なるものではないと仮定すると, しばしばインポータンス密度は自然に選択される. (β,σ) についての確率抽出は簡単に行うことができ, したがって事前分布は評価することができ, 前述の全ステップで完全な分析ができる.

計算ソフトウェア問題

インポータンス・サンプリングを用いたモンテカルロ積分を行ういくつかのパッケージプログラムがある. たとえば, BRAP は正規あるいは t インポータンス関数を用いてモンテカルロ積分を行うプログラムである. 多分現在利用可能な最も手の込んだモンテカルロ積分パッケージは, Hop and van Dijk (1990) によって開発されたものであり, 2 つのアルゴリズムでプログラムした SISAM と MIXIN である. SISAM は, t インポータンス関数を用いたモンテカルロ積分を行う. 他方, MIXIN はもっと複雑な混合積分を用いて, モンテカルロと数値積分の両方を結合して歪んでいる可能性のある密度を含む積分の正確な評価を行う. 両方のサブルーティンを使うと, 広範囲のインポータンス関数の診断ができる. BRAP のモンテカルロ積分の機能を利用するのには Fortran が必要である. Fortran と NAG サブルーティン・ライブラリーが SISAM と MIXIN を利用するのに必要である.

しかし通常は GAUSS, S-Plus や Fortran, あるいはどれか共通のプログラム言語を用いて特定目的のプログラムを開発する場合が多い. このとき必要なの

は, サンプリング毎に事後分布を評価するサブルーティンである. すなわち, インポータンス関数から標本を抽出し, そのサンプリング[8] でのインポータンス関数を評価するサブルーティンと, すべての機能を含んだメインプログラムの部分は各サンプリングで評価し, 適切にウェイト付けする. インポータンス関数に対するパラメータ値を得るために, 事後分布を最初に最適化する必要がある場合もある. この場合では, GAUSS のようなパッケージは最適化サブルーティンをもっており, ユーザーはすでに最適化する関数に関してはプログラムしている.

受容/棄却法

インポータンス・サンプリングで $p(\theta)$ から標本を発生させにくい場合には, 次のような受容/棄却法 (A-R:*Acceptance-Rejection Sampling*) が用いられる. 目的とする密度 $\pi(x) = p(x)/K$ から標本を発生するとし, ここで $x \in \Re^d$, $p(x)$ は正規化されていない密度であり, K は (未知の) 正規化定数である. $g(x)$ をある既知の方法でシミュレートできる密度とし, すべての x に対して $\pi(x) \leq cg(x)$ であるような c が存在するとすると, $\pi(\cdot)$ から確率変数を得るには

(1) $g(\cdot)$ から確率標本の候補 z を発生させる.
(2) $(0,1)$ 区間上の一様分布 $\mathcal{U}(0,1)$ から乱数 u を発生させる.
(3) $u < \pi(z)/cg(z)$ ならば z を確率標本として受容し, そうでなければ棄却して (1) に戻る.
(4) (3) で得られた z を用いて $f(\theta|z)$ から確率標本 θ^* を発生させる.

とする. (1) から (3) のステップで受容された z の値は $\pi(\cdot)$ からの変量になることが証明されている (Ripley, 1987). この方法が有効であるためには, c を十分注意して選ぶ必要がある. この方法は $g(\cdot)$ の $p(\cdot)$ に対する近似がよくなければ受容率が低くなり, 効率の悪い乱数発生法になる.

---

[8] w_i の計算で, $g(\theta)$ あるいは $p(\theta|y)$ は, 任意のスケールファクターを掛けても結果に影響を与えない. しかし, 正規化定数は事後オッズ解析に重要であるので, 事後オッズを計算するとき注意しなければいけない.

SIR:Sampling/Importance Resampling 法

もう一つ公表されているデータを用いて統計的推測を簡単に行うために, 欠測値を埋める方法として Rubin(1987) は *Sampling/Importance Resampling* (SIR) 法を提案している. まず $g(Z|y)$ から目標とする m 個の乱数よりもずっと多い M 個の乱数を発生させる. ウェイトに従ってリサンプリングし, 目的の m 個の乱数を得る方法である. 未知の母数を θ とし, 観測されるデータを y, 観測されない観測値を Z として, 事後分布 $p(Z|y)$ から欠測値 Z を発生させる方法である. まず, 事後分布 $p(\theta, Z|y)$ に対してよく近似しており, 標本を発生させやすい密度関数

$$g(\theta, Z|y) = g(\theta|y)g(Z|\theta, y)$$

を選ぶ. $g(\theta|y)$ と $g(Z|\theta, y)$ はそれぞれ $p(\theta|y)$ と $p(Z|\theta, y)$ を近似した確率標本を発生しやすい密度関数である. そして

(1) $g(\theta, Z|y)$ から M 個の標本 $(\theta_i, z_i), (i = 1, 2, \ldots, M)$ を発生する.

(2) ウェイト w_i を次のように求める.

$$w_j = w(\theta_j, z_j|y) \propto \frac{p(y, z_j|\theta_j)p(\theta_j)}{g(\theta_j, z_j|y)}, j = 1, 2, \ldots, M$$

(3) $(z_1, z_2, \ldots, z_M)$ の中から $m(< M)$ 個の Z を確率 $p(Z = z_j) = w_j / \sum_{i=1}^{M} w_i$ で発生させる.

w_j の分布が非常にゆがんでいる場合は SIR による近似はきわめて悪いので, g の取り方に注意する必要がある.

例：ロジットモデル (続き)

再び Koop and Poirier (1993) のロジットモデルの例に戻ると, ロジットモデルに対する事後密度から直接にサンプリングするのは簡単ではない. 結果的に, ベイズ分析で $E(P_{ij}|y)$ のような数量を, 解析的あるいは単純なモンテカルロ積分を用いて計算することはできない. この場合は, インポータンスサンプリングを用いたモンテカルロ積分が必要であり, この方法を簡単に導入することができる.

弱い条件の下で, ロジット尤度関数はパラメータに関して凹である. 少なくとも Koop and Poirier (1993) で用いられたデータセットに対して, t インポータンス関数は事後密度に極めてうまく近似できる. 上記の論文では, 最初にロジット事後分布の対数を最適化し, 結果として得た事後モードとマイナスのヘッセ行列の逆行列をインポータンス関数の平均と分散として用いている. 自由度のパラメータの様々な値に対して実験し, 5 つの値を選択すると, それは正規分布の裾よりもかなり厚いことを示している. 次に, パラメータや P_{ij} の事後平均や事後分散のような様々な統計量を計算している.

これらのステップを遂行するのに必要なコンピュータプログラムは, 非常に単純である. 最初のステップでは, ロジット事後密度の対数が最適化され, 最適値とヘッセ行列が計算される. もし事前密度がフラットなら, このステップは MLE の計算とその分散の推定と同じである. この最初のステップは, インポータンス関数のパラメータを与える. 2 番目のステップは, インポータンス関数からサンプリングし, P_{ij} のような特性値を各サンプリング毎に計算する. 次に, そのサンプリングを以前に g^* に対して与えた単純な公式を用いて平均する.

2.2.3 Poly- t 密度

ベイジアン計量経済学における計算手法を説明する場合に, poly-t 密度についても多少ふれる必要があるだろう. Poly-t 密度は t カーネル (密度のカーネルは積分定数を無視した密度である) の積あるいは比であり, ベイジアン計量経済分析, 特に同時方程式モデルあるいは操作変数モデルでは頻繁に生じる. 線形回帰モデルでは, そのような密度は異分散が存在する場合に生じる. これらはしばしば解析的なモーメントをもたないので, 数値手法が通常必要とされる. この方法を poly-t 事後密度の性質を評価するのに用いることができるけれども, Bayesian Regression Package(BRP) として知られている特別なプログラムも利用できる (Richard and Tompa, 1980 参照). このプログラムを詳細に記述することはしない. Poly-t 密度で研究しようとする有効な数値積分法はないことに注意すれば十分である.

2.3 マルコフ連鎖モンテカルロ法

マルコフ連鎖モンテカルロ法 (*Markov chain Monte Carlo* : MCMC) は, 実際的な統計モデルの作成にとって大きな展望をもたらした. 最近まで, 多くの応用問題における分析対象の複雑性と構造を正確に認識する困難性とから, そのための特別の方法と特定分析目的のソフトウェアが必要であった. 現在では, MCMC 法を利用することによって, 多くの複雑な問題を汎用的なソフトウェアを用いて分析することができる統一的な環境が整えられつつある.

MCMC は基本的には, **マルコフ連鎖**を用いた**モンテカルロ積分**である. ベイズ分析では, 場合によっては古典的な分析でさえも, モデルのパラメータについて推測や予測を行うために, かなり高次元の確率分布を積分する必要がある. ベイズ的な方法では, データ所与のもとでのパラメータの事後分布を積分する必要があり, 古典的方法はパラメータの値が与えられたもとでの観測値の分布を積分する必要がある. **モンテカルロ積分**は必要な分布から標本を抽出し, 期待値に近似するために標本平均を計算する. **マルコフ連鎖モンテカルロ法**は, これらの標本を長期間に巧妙に作られたマルコフ連鎖を作動させることによって抽出する. これらの連鎖を作成する方法は多数あるが, Gibbs サンプラーを含むこれらの方法はすべて, Metropolis et al. (1953) と Hasting (1970) の一般的な枠組みの中の特別なクラスになる.

1990 年代はベイズ的な方法を応用する優れた実証経済分析が表れ始めた年代である. これらの応用のほとんどは事後分布をシミュレートするのに, マルコフ連鎖モンテカルロ法を用いている. シミュレーション・アルゴリズムは, その基本的な形において非常に簡単であるので, 多くのベイズ型応用研究ではだんだん標準的な方法になりつつある. さらに, この方法には長い歴史があり, 基本的な理論は確立されている. もともとは統計物理の分野で始められ, 空間統計学やイメージ解析の分野で長いこと使用されていたが, 最近数年の間に MCMC 法はベイズ統計学に大きな影響を与え, 古典的な統計分析の分野でもいくつかの応用が見られる. 最近の研究では多くの応用, 手法の豊富さ, そして理論的な精緻さにおいてかなりの成果が加えられた.

基本的には 2 種類の MCMC 法がある. ギブス・サンプラー (*Gibbs Sampler*)

は, **完全な条件付き分布** $p(\theta_i|\theta_{j\neq i}, y), i = 1, \ldots, k$ の集合から順々に標本をとり, かなり緩い条件の下でその安定分布として結合事後分布 $p(\theta|y)$ をもつマルコフ連鎖を生成する. このアルゴリズムは Geman and Geman (1984) によって名付けられ, Gelfand and Smith (1990) は広範囲のベイズ推定問題に適用できることを示した. 最近の優れた入門的な解説は Casella and George (1992) を参照するとよい.

もう一つの方法は, 完全な条件付き分布からシミュレートするのが難しい時に適用される. この場合, 別々の異なるマルコフ連鎖をシミュレートする代わりに, ある別の安定分布を持っているが, それを修正してその安定分布としての事後分布を持つ新しいマルコフ連鎖が生成できるようにする. これは**メトロポリス-ヘイスティング・アルゴリズム** (*Metropolis-Hastings Algorithm*) によって達成される. 以前に定めた各パラメータに対する**候補分布**から標本をとり, その後受容-棄却ステップを用いる. 重要な特徴はこの方法は標本をとる連鎖で正規化されていない事後密度 $p(\theta|y) = L(y|\theta)\pi(\theta)$ しか含まないということである. よく知られた参考文献は, Metropolis et al.(1953) と Hastings(1970) である. それに Chib and Greenberg (1995b) には, 優れた入門的な解説がある.

本章では最初に, 重要な導入部分であるベイズ推定についてまとめた後にマルコフ連鎖について Chib and Greenberg (1995b) に沿って述べる. 次に, Metropolis-Hastings アルゴリズムで与えられる MCMC の一般型を説明した後, ギブス・サンプラーがその特別の場合であることを述べる. 応用計量分析でよく用いられているギブス・サンプラーについて述べ, その計算方法と収束判断の方法について簡単にまとめる. 最後に, データ拡大アルゴリズムについて述べることにする.

マルコフ連鎖モンテカルロ法に関して, いろいろな面から取り上げた本が最近出版されている. たとえば, Carlin, B.P. and Louis, T.A. (1996) ではデータ解析の観点からベイズ法と経験ベイズ法がまとめられており, Gilks, Richardson, and Spiegelhalter (1996) は 25 編の興味ある論文が基礎から応用まで網羅されている, それに Gamerman (1997) にマルコフ連鎖モンテカルロに属する方法が詳細にまとめて述べられている. Kass, et al (1998) には, この問題に関する興味あるいろいろな示唆が述べられている. 日本語で書かれた文献は少ないが, 大森

(1996) にはマルコフ連鎖モンテカルロ法の紹介とその収束診断基準が要領よくまとめられている. 伊庭 (1997) には統計学への応用が参考になり, 繁桝 (1995) の本の付録にこの方法の説明がある.

2.3.1 ベイズ推定

現在での MCMC 法のほとんどの応用はベイズ推定を行うために始まった. ベイズ分析での観点では, 統計モデルの観測可能な変数とパラメータに間には何の基本的な差異はなく, すべて確率変数と考えられている. y を観測データ, θ はモデルパラメータと欠測データを表す. 推測の段階では, すべての確率変量に関して**事前分布** $p(\theta)$ と**尤度** $L(y|\theta)$ の 2 つの部分からなる同時分布 $p(y,\theta)$ をセットアップする. $p(\theta)$ と $L(y|\theta)$ が決まれば, 次の確率モデル (*full probability model*) が得られる.

$$p(y,\theta) = L(y|\theta)p(\theta)$$

データ y を観測すると, ベイズの定理を用いて y の条件のもとでの θ の分布を決定する.

$$p(\theta|y) = \frac{p(\theta)L(y|\theta)}{\int p(\theta)L(y|\theta)d\theta}$$

これは θ の **事後分布**と呼ばれ, すべてのベイズ推定の目的である.

事後分布のすべての特性値, モーメント, 四分位値, 最高事後密度区域等が, ベイズ推定にとって重要である. これらの統計量は θ の関数の分布全体ではなく, 事後期待値によって表すことができる. 関数 $f(\theta)$ の事後期待値は

$$E[f(\theta|y] = \frac{\int f(\theta)p(\theta)L(y|\theta)d\theta}{\int p(\theta)L(y|\theta)d\theta}$$

で得られる. この式の積分は最近まで, 特に高次元の場合にはベイズ推定における実際的な困難のほとんどの原因であった. ほとんどの応用分析では, $E[f(\theta|y)]$ の解析的な評価は不可能である.

期待値の計算

$\boldsymbol{x}$ は k 個の要素からなる確率変数ベクトルで $\pi(\cdot)$ の分布をもつ. ベイズ的な応用分析では, $\boldsymbol{x}$ はモデルパラメータと欠測データからなり, 古典的な応用分析

ではデータあるいはランダム効果である. ベイジアンでは $\pi(\cdot)$ は事後分布になり, 古典派では尤度になる. どちらも, 目的はある関数 $f(\cdot)$ の期待値を計算することである.

$$E[f(x)] = \frac{\int f(x)\pi(x)dx}{\int \pi(x)dx}$$

ここで x の分布は正規化定数まで既知である可能性を認めることにする. たとえば, ベイジアン推定では $p(\theta|y) \propto p(\theta)L(y|\theta)$ は知っているが, 正規化定数 $\int p(\theta)L(y|\theta)d\theta$ を簡単に評価することができない.

モンテカルロ積分

モンテカルロ積分は, 目標とする密度 $\pi(\cdot)$ から標本 $\{x_t, t = 1, \ldots, n\}$ を抽出することによって $E[f(x)]$ を評価し, 次に

$$E[f(x)] \approx \frac{1}{n}\sum_{t=1}^{n} f(x_t)$$

で近似する. $f(x)$ の母集団平均は標本平均によって推定される. 標本 $\{x_t\}$ が独立なら, 大数の法則によって近似は標本サイズ n を増やすだけ正確になる. ここで n は分析者がコントロールすることができ, 固定されたデータサンプルのサイズではない.

一般に, $\pi(\cdot)$ は標準的ではないので, 標本 $\{x_t\}$ は $\pi(\cdot)$ から独立に抽出されることはない. しかし, $\{x_t\}$ は必ずしも独立である必要ではない. $\{x_t\}$ は任意のプロセスによって生成することができる. 大胆に言えば, $\pi(\cdot)$ のサポートを通じて正しい割合で標本を取ればよい. これを行う一つの方法は, 安定分布として $\pi(\cdot)$ をもつマルコフ連鎖によって行う. これが**マルコフ連鎖モンテカルロ**である.

2.3.2 マルコフ連鎖

各時間 $t \geq 0$ で一連の確率変数 $\{x_0, x_1, x_2, \ldots\}$ を生成すると仮定すると, 次の状態 x_{t+1} は連鎖の現在の状態 x_t にのみ依存する分布 $P(x_{t+1}, x_t)$ からとられる. これは, x_t **所与のもとでの**次の状態 x_{t+1} は, 過去の連鎖 $\{x_0, x_1, \ldots, x_{t-1}\}$ にはこれ以上依存しない. この系列を**マルコフ連鎖**と呼び, $P(\cdot, \cdot)$ は連鎖の**推移核**と呼ぶ. 連鎖は時間同時性, すなわち $P(\cdot, \cdot)$ は t に依存しないと仮定する. 一般的には遷移核 $P(x, A)$ は条件付き分布関数で, x から集合 A 内の一点へ移

動する確率を表している. 分布関数であるから, $P(x, \Re^d) = 1$ で, ここで連鎖は点 x から x へ移動する場合もある, すなわち $P(x, \{x\})$ は必ずしもゼロにならない.

初期状態 x_0 がどのように x_t に影響を与えるのかを見るために, x_0 所与のもとでの x_t の分布を $P^{(t)}(x_t, x_0)$ と表す. ここで中間の値 $\{x_1, x_2, \ldots, x_{t-1}\}$ は与えられていないので, x_t は直接 x_0 に依存する. 正則条件のもとで, 連鎖は徐々にその初期状態を忘れ, $P^{(t)}(\cdot, x_0)$ は究極的には一意な**安定分布** (あるいは**不変分布**) に収束し, これは t あるいは x_0 に依存しない. 当面, 安定分布を $\phi(\cdot)$ と表す. すると, t を増やすにつれ, 標本点 $\{x_t\}$ は $\phi(\cdot)$ からの相関のある標本のように増加しているように見える.

連続状態空間に対するマルコフ連鎖理論の方法は, 遷移核 $P(x, A)$ から出発する. 遷移核は条件付き分布関数で, x から集合 A 内の一点へ移動する確率を表している. 分布関数であるから $P(x, \Re^d) = 1$ で, ここで連鎖は点 x から x へ移動する場合もある, すなわち $P(x, \{x\})$ は必ずしもゼロにならない.

マルコフ連鎖理論の主要な問題は, 安定分布 π^* が存在する条件と遷移核のイタレーションが安定分布へ収束する条件を決めることである. 安定分布は, 次の条件を満足する.

$$\pi^*(dy) = \int P(x, dy)\pi(x)dx \tag{2.1}$$

問題は適切な $P(x, dy)$ を見付けることである. 関数 $p(x, y)$ に対して遷移核は次のように表される.

$$P(x, dy) = p(x, y)dy + r(x)\delta_x(dy) \tag{2.2}$$

ここで, もし $x \in dy$ なら $p(x, x) = 0, \delta_x(dy) = 1$, その他はゼロである. そして $r(x) = 1 - \int_{\Re^d} p(x, y)dy$ は連鎖が x にとどまる確率である. $r(x) \neq 0$ の可能性から, y に関して $p(x, y)$ の積分が必ずしも 1 にならないことは明らかである.

さて, 関数 $p(x, y)$ が反転条件 (あるいは詳細均衡)

$$\pi(x)p(x, y) = \pi(y)p(y, x) \tag{2.3}$$

を満たせば, $\pi(\cdot)$ は $P(x, \cdot)$ の不変密度になる.

2.3.3 Metropolis-Hastings アルゴリズム

メトロポリス・ヘイスティング・アルゴリズムは, 標準的でない分布 $\pi^*(\cdot)$ からサンプルするのに用いることができる強力な MCMC 法である. アルゴリズムから一連の抽出を行う方法は, その安定分布 $\phi(\cdot)$ が関心ある分布 $\pi(\cdot)$ であるようなマルコフ連鎖を生成する方法を見つける必要があるが, このようなマルコフ連鎖を生成するのは非常にやさしい. 以下で, その形を Hatings(1970) にしたがって述べる. これは Metropolis et al. (1953) によって最初に提案された方法の一般化である.

前述の受容/棄却法と同様に, いくつかの候補を生成できる密度があると仮定する. マルコフ連鎖を用いていることから, 密度がプロセスの現在の状態に依存することを認める. したがって, **候補生成密度** (*candidate generating density* あるいは *proposal density*) は $q(x,y)$ と表され, ここで $\int q(x,y)dy = 1$ である. この密度はプロセスが x 点にあるとき, 密度は y を $q(x,y)$ から生成されることを意味している. もし $q(x,y)$ 自身が反転条件を満たせば探索は終了するが, ほとんどそういうことはない. たとえば, ある x, y について

$$\pi(x)q(x,y) > \pi(y)q(y,x) \tag{2.4}$$

となる場合がある. この場合, 大ざっぱには, プロセスは x から y へ頻繁に移動し, y から x へは滅多に動かない. この状態を修正する簡単な方法は, 移動が起こる確率 $\alpha(x,y) < 1$ を導入することによって x から y へ移動する数を減らす. ここで $\alpha(x,y)$ を**移動確率**と呼ぶ. 移動が起こらなければ, プロセスは再び x を目的分布からの値として戻す. そこで x から y $(y \neq x)$ への移行は

$$p_{\mathrm{MH}}(x,y) \equiv q(x,y)\alpha(x,y), \qquad x \neq y$$

に従って起こる. ここで $\alpha(x,y)$ も決められる.

再び不等式 (2.4) を考えると, これは y から x へはそれほど起こらないことを表している. したがって, $\alpha(x,y)$ をできるだけ大きくするようにし, これは確率であるから上限は 1 になる. しかし, 現在は移動確率 $\alpha(x,y)$ は $p_{\mathrm{MH}}(x,y)$ が反転条件を満足するように決められる. そのとき次のようになるためである.

$$\pi(x)q(x,y)\alpha(x,y) = \pi(y)q(y,x)\alpha(y,x)$$

$$= \pi(y)q(y,x)$$

ここから $\alpha(x,y) = \pi(y)q(y,x)/\pi(x)q(x,y)$ であることが分かる. もちろん不等号が逆の時は $\alpha(x,y) = 1$ である. 確率 $\alpha(x,y)$ と $\alpha(y,x)$ は前の不等式の両辺をバランスするように, 言い換えると $p_{\mathrm{MH}}(x,y)$ が反転条件を満足することを保証するために決められる. そこで, $p_{\mathrm{MH}}(x,y)$ が反転可能であるためには, 移動確率は次のように決められる.

$$\alpha(x,y) = \begin{cases} \min\left[\dfrac{\pi(y)q(y,x)}{\pi(x)q(x,y)}, 1\right] & \pi(x)q(x,y) > 0 \text{ のとき} \\ 1 & \text{その他} \end{cases} \tag{2.5}$$

M-H 連鎖に対する遷移核の定義を行うために, プロセスが x にとどまるゼロでない確率を考える. 先に定義したように, この確率は

$$r(x) = 1 - \int_{\Re^d} q(x,y)\alpha(x,y)dy$$

となる. したがって, $p_{\mathrm{MH}}(x,dy)$ と表される M-H 連鎖の遷移核は

$$P_{\mathrm{MH}}(x,dy) = q(x,y)\alpha(x,y)dy + \left[1 - \int_{\Re^d} q(x,y)\alpha(x,y)dy\right]\delta_x(dy)$$

で与えられる.

Metropolis-Hasting アルゴリズムでは, 各期 t で次の状態 x_{t+1} は**提案された分布** $P(\cdot, x_t)$ から**候補点** y を最初に抽出する. 提案された分布は現在の点 x_t に依存していても良い. たとえば, $P(\cdot, x_t)$ は多変量正規分布で, 平均が x, 一定の共分散行列である場合もある. 候補点 y は, 次の確率 $\alpha(x_t, y)$ で**受容**される. 最後のサンプルで x の値を得たとすると, 次の値は候補生成密度 $q(x,y)$ から y をサンプルすることによって生成される. このようにして得られた y は, 確率 $\alpha(x,y)$ で受容される. 候補点が受容されると, 次の状態は $x_{t+1} = y$ となる. 確率 $1 - \alpha(x,y)$ で候補が棄却されると, 連鎖は動かずサンプルした値は現在の値 $x_{t+1} = x_t$ となる.

このマルコフ連鎖の状態遷移を決める移動確率は, 反転条件を満たしていてマルコフ連鎖は時間に可逆 (*reversible*) である. このことから

$$\int \pi(x_t)q(x_{t+1}, x_t)dx_t = \pi(x_{t+1})$$

が言え, $\pi(x)$ はこのマルコフ連鎖の定常分布になる.

もし $x^{(0)} \sim \pi(x)$ なら, この式より各 $x^{(t)}$, $t \geq 1$ も $\pi(x)$ に従うことになるが, $x^{(0)}$ をそのように選ぶことはできない. 実際には, 十分に長い, たとえば m 回のイタレーションの後, 点 $x^{(m+1)}, \ldots, x^{(n)}$ まで発生させ, 最初の m 個のサンプルを捨ててマルコフ連鎖からの結果を用いて, 期待値 $E[f(x)]$ を推定する. ここで x は分布 $\phi(\cdot)$ をもつ. 初期のサンプルを捨てることは "burn-in" と呼ばれる. この burn-in 標本は通常計算から除外され, 次の推定量を得る.

$$\overline{f} = \frac{1}{n-m} \sum_{t=m+1}^{n} f(x_t) \tag{2.6}$$

これは**エルゴート的平均**と呼ばれる. 求める分布への収束は, エルゴード定理により保証される. 一般にマルコフ連鎖が proper な定常分布を持ち, 既約 (irreducible) で非周期的 (aperiodic) なら, 定常分布は一意で, $t \to \infty$ の時の $x^{(t)}$ の極限分布が $\pi(x)$ に収束する.

いくつかの重要な点がある. 第 1 に, $\alpha(x, y)$ に計算には $\pi(\cdot)$ の正規化定数の知識を必要としない. 第 2 に, もし候補生成密度が対称なら, すなわち $q(x, y) = q(y, x)$ なら, 受容確率は $\pi(y)/\pi(x)$ に縮小し, これは Metropolis et al.(1953) の本来の公式である. 最後に, Gibbs サンプラーは MH アルゴリズムの特別な場合であることが示すことができる (Chib and Greenberg, 1995b).

2.3.4 Gibbs サンプリング法

5 章で述べた手法は, 様々な計量経済モデルを分析することができるが, 多くの場合事後密度から直接サンプリングしたり, あるいは良いインポータンス関数を見つけるのが不可能な場合があり, その場合は上で述べたモンテカルロ法を直接適用することができない. 単純なモンテカルロ積分を事後密度からサンプリングし, これらのサンプルから事後密度の任意の特性値の期待値が計算できることを前に述べた. Gibbs サンプリング法は事後分布からのサンプリングに収束する乱数を生成するので, この性質のためにこの方法がよく使われるようになった. これらの乱数から, モンテカルロ積分が利用できない場合における事後密度の特性値を簡単に計算することができる[9]. 単一要素だけからなる

[9] Gibbs サンプリングはこのような場合に適しているが, 他の似たような方法も存在する. 関心がある人は, Tierney (1991), Tanner (1991), Gibbs サンプリングと関連手法に対する議論は,

Metropolis-Hastings の特別な場合が **Gibbs サンプリング**である. Gibbs サンプリングはその名称を Geman and Geman (1984) によって与えられ, 格子状の Gibbs 分布を解析するのに用いられたが, その応用可能性は Gibbs 分布に限られない. そこで Gibbs サンプリングは実際は誤解しやすい. 同じ方法はすでに統計物理の分野で使われており, そこでは **heat bath algorithm** として知られている. 言うまでもなく, Geman and Geman (1984) の仕事は Gelfand and Smith(1990) と Gelfand et al. (1990) の研究を通じて MCMC の統計学の主要な分析道具への道を拓いた. 今日まで, MCMC のほとんどの統計的な応用は Gibbs サンプリングが用いられている.

最近, 応用計量経済分析を行う人々の間に Gibbs サンプリング法に対する関心が生じてきた. ギブス法は, 次のような分野の分析に用いられている. 多項目反応データ (Albert and Chib, 1993a); トービットモデル (Chib, 1992); 多対プロビットモデル (McCulloch and Rossi, 1994); ARMA(p, q) 誤差を持つ回帰モデル (Chib, 1993; Chib and Greenberg, 1992); マルコフスウィッチング・モデル (Albert and Chib, 1993b; McCulloch and Tsay, 1994a); ストカスティック・フロンティア・モデル (Broeck, Koop, Osiewalski and Steel, 1994); ストカスティック・ボラティリティ・モデル (Jacquier, Polson and Rossi, 1994); 多重変化点モデル (Chib, 1998), そして非共役事前分布をもつ時系列モデル (Geweke, 1992a) である. この分野での最近の研究の発展から判断すると, ギブス法は応用ベイズ計量経済分析には重要な手法になっていることが分かる.

Gibbs サンプリング法の主要な仮定は, たとえ同時分布からサンプリングするのが難しくとも, いくつかの条件付分布からは簡単にサンプリングできることである. たとえば, A と B が確率変数であるとし, $p(A, B)$ からサンプリングしたい. たとえ $p(A, B)$ から直接にサンプルをとることができなくても, $p(A|B)$ と $p(B|A)$ からサンプリングできれば, Gibbs サンプリングを使って $p(A, B)$ から極めて簡単にサンプリングができる. 特に, 条件付分布から次々とサンプルをとることによって, 同時密度からの標本に収束することができる. 条件なしの密度からよりも, 条件付密度からサンプリングする方がずっと簡単であるから, ギブス・サンプラーは様々な計量経済学上の問題に対して役に立つことを示した.

Tanner and Wong (1987) あるいは Gelfand and Smith (1990) を参照すると良い.

もっと正確には, $\boldsymbol{\theta}$ を $(\theta_1', \ldots, \theta_k')$ のように分解されるパラメータ・ベクトルとする. ギブス・サンプラーが役に立つためには, $j = 1, \ldots, k$ に対して $p(\theta_j|\theta_1, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_k)$ から簡単にサンプリングできるような分割を選ばなければならない. Gibbs サンプリングは, 各サンプルが他の密度からの以前のサンプルについての条件付きである密度から連続してサンプリングを行う. すなわち, θ_j^i を $j = 1, \ldots, k, i = 1, \ldots, s$ に対する $p(\theta_j|\theta_1^i, \ldots, \theta_{j-1}^i, \theta_{j+1}^{i-1}, \ldots, \theta_k^{i-1})$ からのサンプルとし, ここで θ^0 は任意の初期値である. k はステップ数 (抽出される条件付密度の数) を示し, s はパスの数を示す. もし $\boldsymbol{\theta}^i$ を i 番目のパス, $\boldsymbol{\theta}^i = (\theta_1^{i\prime}, \ldots, \theta_k^{i\prime})'$, でとられた 1 組のサンプルとすると, 弱い条件の下で $\boldsymbol{\theta}^i$ は $p(\boldsymbol{\theta})$ からのサンプルに収束する. さらに, 次を示すこともできる.

$$g^* = \frac{1}{s}\sum_{i=1}^{s} g(\boldsymbol{\theta}^i) \to E(g(\boldsymbol{\theta}))$$

ここで例示のために, 2 変量の場合を考える. 初期値 $\boldsymbol{\theta}^0 = (\theta_1^0, \theta_2^0)$ を与えると, ギブス・サンプラーは 1 つのサンプル θ_1^1 を $p(\theta_1|\theta_2^0)$ からとり, 1 つのサンプル θ_2^1 を $p(\theta_2|\theta_1^1)$ からとり, 1 つのサンプルを $p(\theta_1|\theta_2^1)$ からとり, 等々と繰り返す. 何回かのパスの後, この方法によって $p(\theta_1, \theta_2)$ から連続してサンプルをとり, $E(g(\theta))$ 自身と $g(\cdot)$ を計算するのに使うことができる.

一般的には, $\pi(\theta), \theta \in \mathcal{S} \subseteq \Re^p$ を (正規化されていない) 目標密度とする. θ を $\theta_1, \theta_2, \ldots, \theta_d$ に何らかの分解し, k 番目のブロックの完全条件付き密度を $\pi(\theta_k|\theta_1, \ldots, \theta_{k-1}, \theta_{k+1}, \ldots, \theta_d)$ と表す. すると, Gibbs サンプリング・アルゴリズムは次の繰り返しによって定義される.

1. 初期値 $\theta^{(0)} = (\theta_1^{(0)}, \ldots, \theta_d^{(0)})$ を与え, $i = 0$ とおく

2. 次のようにシミュレートする

$$\begin{aligned}
&\theta_1^{(i+1)} \text{を} \pi(\theta_1|\theta_2^{(i)}, \theta_3^{(i)}, \ldots, \theta_d^{(i)}) \text{ から} \\
&\theta_2^{(i+1)} \text{を} \pi(\theta_2|\theta_1^{(i+1)}, \theta_3^{(i)}, \ldots, \theta_d^{(i)}) \text{ から} \\
&\theta_3^{(i+1)} \text{を} \pi(\theta_3|\theta_1^{(i+1)}, \theta_2^{(i+1)}, \theta_4^{(i)}, \ldots, \theta_d^{(i)}) \text{ から} \\
&\quad \vdots \quad \vdots \\
&\theta_d^{(i+1)} \text{を} \pi(\theta_d|\theta_1^{(i+1)}, \theta_2^{(i+1)}, \ldots, \theta_{d-1}^{(i+1)}) \text{ から}
\end{aligned}$$

3. $i=i+1$ とおいでステップ2に戻る

このアルゴリズムはマルコフ連鎖の次の項 $\theta^{(i+1)}$ を各完全条件付き密度をシミュレートすることによって与え, ここで条件付きの要素は一つのサイクル中に改訂される.

次に, Gibbs サンプリングを実行する際に生じるいくつかの問題に移ろう. 最初に, ブロックを計画するときに高い相関がある要素は一緒にまとめるべきである. そうしないと, マルコフ連鎖はゆっくりとしか減衰しない自己相関を示す傾向があり, 目標密度へゆっくりと収束することになる. 第2に, 扱いやすい完全条件付き構造は, しばしば x の定義に潜在データあるいは欠測データを導入することによって得られる. サンプラーに変数を追加するという考えはデータ拡大として知られており, Tanner and Wong(1987) で導入され, いくつかの応用例がある. 最後に, 完全条件付き密度のいくつかが伝統的な方法 (棄却サンプリング法や既知の発生器) によってサンプルすることが難しい場合, その密度は MH アルゴリズムによって (Müller, 1991) か, あるいは独立サンプルを生成することによってサンプルすることができる (Gilks and Wild, 1992).

各パスの最後で, ギブスサンプラーは $\theta_1,\ldots,\theta_k$ のそれぞれに対してサンプルを得る. パス毎の平均を, 任意の $g(\cdot)$ に対する $E(g(\theta)|y)$ を推定するのに使うことができる. しかし, 2つの問題がモンテカルロ積分を適用できない Gibbs サンプリング手法を使う際に生じる. それは初期条件とパス毎の系列相関である. 理論的には, ギブスサンプラーから得るサンプルは, パスの数が多いときだけは事後分布からのサンプルと等しくなる. もし θ^0 が不適切に選ばれると, 初期のサンプルは後のサンプルと大きく異なる場合がある. このために, 多くの分析ではギブスサンプラー中の m 個の初期のパスを棄てており, これによって θ^0 の選択への依存を減らす.

単純なモンテカルロ積分とは異なり, ギブスサンプラーは事後密度からの独立な標本を生成せず, したがって, サンプルがパス毎に系列相関があることを簡単に示すことができる. この特性から, 研究者によっては系列相関を断ち切る意味から r 回目のパス毎のサンプルを棄てるか, r 回目のパス毎にギブスサンプラーを再始動させている. 実証的な問題では様々な方策を試みるか, θ^0, m, r の異なる選択をし, 結果がこれらに敏感かどうかを見ることが重要である. もち

ろん, 最適な Gibbs サンプリング・アルゴリズムは問題毎に異なる. しかし, ギブスサンプラーから 1 つの長いランを得た方がよいか, あるいは再出発した方がよいか? という問題に関心が持たれている (Tanner, 1991; Zeger and Karim, 1991; Raftery and Lewis, 1991; and Gelman and Rubin, 1992; を参照).

モンテカルロ法が正規近似あるいはラプラス近似より優る 1 つの利点は, この方法には診断法があり, 近似の精度を反復数を選択することによりコントロールすることができることにある. Geweke (1992a) は, ギブス・サンプラーに対して同じ種類の診断を提案している. すなわち, もし $E(g(\theta)|y)$ の推定値を

$$g^* = \frac{1}{s^*} \sum_{i \in I} g(\theta^i)$$

とし, ここで I は推定に含まれる一組のパスを示す (I はすべてのパス, あるいは最初の m を除いたすべて, などでも良い), そして s^* は含まれるサンプル数 (s あるいは $s-m$) とすると, 1 組の $[g(\theta^i)](i \in I)$ は単変量確率過程になる. 漸近理論を用いて, Geweke は g^* は平均 $E(g(\theta)|y)$, 漸近分散 $s^{*-1}S_{(0)}$ の漸近正規密度にしたがうことを示した. ここで $S_{(0)}$ は周波数がゼロのところで評価した $\{g(\theta^i)\}$ のスペクトル密度である. この統計量は, $[s^{*-1}S_{(0)}]^{1/2}$ は g^* に対する NSE として用いることができ, 数値標準誤差の計算はギブスサンプラーを採用すると可能になることを示した. $S_{(0)}$ の推定値を得る多くの方法があり, これらは標準的な時系列のテキストで述べられている. さらに, RNE はよく知られた方法で計算することができる.

Geweke (1992b) はさらに別の収束診断法を作成した. ギブスサンプラーはパスの数が大きくなったときにだけ, 事後分布からのサンプルを生成する. したがって, 初期のパスを後のパスと比較すると収束の失敗が明らかになる. $E(g(\theta)|y)$ の 2 つの推定値が作成され, 次のように 1 つは初期のパスに基ずくものであり, もう 1 つは後のパスに基ずくものである.

$$g_A = \frac{1}{s_A} \sum_{i=1}^{s_A} g(\theta^i), \qquad g_B = \frac{1}{s_B} \sum_{i=s_C}^{s^*} g(\theta^i)$$

ここで $s_C = s^* - s_B + 1$. もし 2 つの推定値が大きく変われば, これは収束しなかったことを示す. より詳しくは, NSE_A と NSE_B を前のパラグラフでのよう

に計算した2つの推定値に対するNSEとするとして, Gewekeは次を示した.

$$\frac{(g_A - g_B)}{(\mathrm{NSE}_A^2 + \mathrm{NSE}_B^2)^{1/2}} \to \mathcal{N}(0,1)$$

もし $[g(\theta')]$ が定常なら, s_A/s^* と s_B/s^* は固定されており, $(s_A + s_B)/s < 1$ である. この応用では, Gewekeは $s_A = 0.1s^*$ と $s_B = 0.5s^*$ とおいた. 彼は上の方程式の左辺を収束診断とし, これを収束を確かめるのに用いることを薦めている.

本節では, Gibbsサンプリング法が基本的なモンテカルロ法の強力な拡張になっていることを述べた. ギブスサンプラーは同時事後密度がそれ自体扱い難いが, パラメータ空間を適切に分割した条件付分布から抽出しやすい場合には適している. ギブスサンプラーを利用した論文数の最近の急増ぶりから判断すると, 計量経済学の多くの問題は, そのような条件付の密度によってより簡単に扱えるようになるように見える.

例：マルコフ型スウィッチング回帰モデル

次のMcCulloch and Tsay (1994a) で用いられたマルコフ型スウィッチング回帰モデルの単純化したものを考える.

$$\begin{aligned} y_t &= X_t'\beta_1 + \varepsilon_{1t} \qquad s_t = 1 \text{ のとき} \\ y_t &= X_t'\beta_2 + \varepsilon_{2t} \qquad s_t = 2 \text{ のとき} \end{aligned}$$

ここで ε_{it} は, *iid.* $\mathcal{N}(0, \sigma_i^2)$ の確率変数 $(i = 1, 2, t = 1, \ldots, T)$ である. 任意の時間 t で, これらの回帰モデルのどちらか一方が成り立ち, この状態間のスウィッチはマルコフ過程にしたがう.

$$\begin{aligned} p(s_t = 2 | s_{t-1} = 1) &= p_1, \\ p(s_t = 2 | s_{t-1} = 2) &= p_2. \end{aligned}$$

いまパラメータ $\boldsymbol{\theta} = (\beta_1, \beta_2, \sigma_1^2, \sigma_2^2, p_1, p_2)'$ と状態ベクトル $s = (s_1, \ldots, s_T)'$ に関心がある. このモデルの尤度関数は複雑である. 実際, たとえ非報知事前密度を仮定しても, 事後密度から直接にサンプリングするのは難しい. しかし, 状態ベクトルを知っていれば, モデルは2つの単純な回帰を含んでいる. そして, パ

ラメータの値とデータが既知であれば, 状態について推測するのは簡単である. このことは, このモデルに対してギブス・サンプラーをセットアップできることを示唆している.

McCulloch and Tsay (1992a) は, ギブスサンプラーが次の条件付分布でセットアップできることを示した. $i = 1, 2$ に対して $p(\beta_i|y, s, \sigma_i^2), p(\sigma_i|y, \beta_i), p(p_i|s)$; そして $t = 1, \ldots, T$ に対して $p(s_t|y, s_{(-t)}, \beta_1, \beta_2, \sigma_1^2, \sigma_2^2)$, ここで $s_{(-t)} = (s_1, \ldots, s_{t-1}, s_{t+1}, \ldots, s_T)'$ とする. さらに, これらの条件付サンプルはそれぞれ簡単に抽出できる. 特に, $p(\beta_i|y, s, \sigma_i^2)$ が正規密度, $p(\sigma_i|y, \beta_i)$ は逆カイ2乗密度, $p(p_i|s)$ はベータ密度, そして $p(s_t|y, s_{(-t)}, \beta_1, \beta_2, \sigma_1^2, \sigma_2^2)$ は状態ベクトルのマルコフ性を利用することによって, 簡単に操作できる密度である. 証明と正確な密度関数の詳細は, McCulloch and Tsay を参照すると良い.

ギブス・サンプラーは, 各サンプルが以前のサンプルの条件付きであるときに, これらの条件付分布から連続してサンプリングする. McCulloch and Tsay (1994a) はこのタイプのモデルを用いて米国の GNP と失業率を調べ, McCulloch and Tsay (1994b) は同様のモデルを用いて米国の工業生産における単位根の存在を推定している. この種のモデルは最尤法では分析するのが非常に難しいので, ギブスサンプラーを用いたベイズ推定が極めて自然である.

例：打ち切りデータのある線形回帰モデル

次の線形回帰モデルを考える.

$$w_t = \boldsymbol{x}_t'\boldsymbol{\beta} + \varepsilon_t$$

ここで $\boldsymbol{x}_t$ は t 番目の個人に対する k 次元の外生変数ベクトル, そして ε_t は $iid.\mathcal{N}(0, \sigma^2)$, しかし w_t は観測されない変数である. そのかわり y が観測され, ここで $y_t \leq y^*$ なら $y_t = w_t$, その他の場合は $y_t = y^*$ である (従属変数はセンサーされている). もし従属変数がセンサーされていなければ, モデルは扱いやすい (事後密度は正規誤差を持つ線形回帰モデルに対する事後密度と同じになる). すなわち, w_t が既知なら事後密度は操作可能であるが, もしパラメータが既知なら, w_t についての推測は極めて簡単に行える. これは, 連続したサンプルを $p(\beta, \sigma^2|y, w)$ と $p(w|y, \beta, \sigma^2)$ からとるギブスサンプラーを示している. 非報

知事前密度, あるいは共役事前密度を仮定すると, $p(\beta, \sigma^2|y, w)$ は正規-逆ガンマ密度, $p(w|y, \beta, \sigma^2)$ は多変量正規密度であることを示すことができる. したがって2つの条件付分布は簡単にサンプルでき, ギブスサンプル法は拡張された事後密度の分析に採用することができる. より詳細は, Chib (1992) を参照のこと.

2.3.5 データ拡大 (Data Augumentation) アルゴリズム

Tanner and Wong (1987) によって開発されたデータ拡大 (*Data Augumentation*) アルゴリズムは, 観測できるデータ y に観測できない潜在変数 z を人工的に発生して拡大し, パラメータの事後分布の形状等を評価する方法である. y と z が与えられれば事後分布 $p(\theta|y, z)$ が簡単になることを仮定しており, この方法は次の imputation(埋め込み) と posterior(事後) ステップからなる.

imputation step 予測分布 $p(z|y)$ に対する現在の近似分布から, m 個の潜在変数 z の標本 $z_1, z_2, \ldots, z_m$ を発生させる.

posterior step 得られた拡大データ (y, z) を用いて, モンテカルロ法により $p(\theta|y)$ を近似する.

imputation ステップにおいて, $p(z|y)$ は

$$p(z|y) = \int p(z|\theta, y)p(\theta|y)d\theta$$

であり, $p(\theta|y)$ に関する知識を前提とする. すなわち, 埋め込みステップと事後ステップを繰り返すことを前提とする. 具体的には

(1) 第 i 回目の $p(\theta|y)$ の近似を $p^{(i)}(\theta|y)$ とする. $p^{(i)}(\theta|y)$ から θ^* を発生する.
(2) θ^* を所与として $p(z|y)$ から z を発生する.

(1) と (2) を繰り返すことによって $p(z|y)$ から m 個の z^* を発生する. $p(\theta|y)$ の近似は次の積分に基づく.

$$p(\theta|y) = \int p(\theta|y, z)p(z|y)dz$$

この積分を m 個の z^* を使ったモンテカルロ法によって近似する.

$$p^{(i+1)}(\theta|y) = \frac{1}{m}\sum_{j}^{m} p(\theta|y, z_j)$$

埋め込みステップの $p^{(i)}(\theta|y)$ もこのようにして得られる. このような混合分布に従う乱数を, 最初に m 個の分布の一つを無作為に選択し, 次に選ばれた $p(\theta|y, z_j)$ から θ を 1 つ発生する様にして作成する.

この考え方を応用したモデル化は最近増えており, たとえばランダム・ポイントにおける構造変化モデル (Carlin, Gelfand, and Smith, 1992), 打ち切りデータや離散データのモデル (Chib, 1992; Albert and Chib, 1993b), マルコフ・スウィッチングモデル (Albert and Chib, 1993b; Chib, 1993; McCulloch and Tsay, 1994), パラメータ制約のあるモデル (Gelfand, Smith, and Lee, 1992), それにセンサー化自己回帰モデル (Zangari and Tsurumi, 1996), 等々が含まれる.

例 : センサー化自己回帰モデル

Zangari and Tsurumi (1996) では, センサー化自己回帰誤差モデルのパラメータの事後モーメントと事後確率密度を 3 つの方法で導いている. ラプラス近似, データ拡大のある Gibbs サンプリング, そしてガウス求積法である. ベイズ点推定量を事後平均として求め, それをシミュレートした EM アルゴリズムと疑似最尤法で求めた最尤推定量と比較した.

これらの方法は, 日本の米国への乗用自動車輸出の 1974 年から 1992 年までの年データと四半期データを用いた回帰モデルに適用された. 日本の乗用自動車輸出は, 1981 年から 1986 年まで輸出自主規制として知られている実質数量割り当ての状態におかれていた. この自主規制は米国の自動車生産が自主規制の上限が無効になるまで 6 年間続いた. この間の時期をセンサーされていたと考え, 自動車輸出の需要関数を計測している. 彼らの研究の結果, 標準的な Tobit 回帰モデルとは異なり, センサーされている観測値がかなり多く, データに強い自己相関が見られる場合 Gibbs サンプラーと結びつけたデータ拡大アルゴリズムは計算上の問題に直面することが報告されている.

センサーされた観測値 $y^c = \{y_t^c\}$ とともに, 対応する潜在データ $y^{c*} = \{y_t^{c*}\}$ が利用できるとする. さらに, センサーされていない観測値 $y = y^*$ を $y' = \{y_t'\}$ と表す. 結果的に, $y = \{y', y^c\}$ と $y = \{y_t\}$ である. データ拡大 Gibbs サンプラーの基礎は $p(\theta|y, y^{c*}) = p(\theta|y', y^{c*})$, すなわち $p(\theta|y^c, y', y^{c*}) \stackrel{d}{=} p(\theta|y', y^{c*})$ である. y^{c*} は観測されなくとも, シミュレートできる. センサー化された自己相

関のある回帰の変換は, もとの式からその1期ラグを引いて得られる. すなわち L をラグオペレータとすると, $y_t^0 = (1-\rho L)y_t^*$ と $x_t^0 = (1-\rho L)x_t$ である. 変換したデータでのモデルは, 次のように表される.

$$y^0 = x^0\beta + u, \qquad u \sim \mathcal{N}(0, \sigma^2 I_{T-1})$$

事前分布 $p(\beta, \sigma^2)$ は ρ とは独立であるとすると, (β, σ^2, ρ) の事前分布は, $p(\beta, \sigma^2, \rho) = p(\beta|\sigma^2)p(\sigma^2)p(\rho)$ となる. ここで, 特に

$$p(\beta|\sigma^2) \sim \mathcal{N}(\beta_0, \sigma^2 A_0^{-1}), \quad p(\sigma^2) \sim \mathcal{IG}(\nu_0/2, \delta_0/2), \quad p(\rho) \sim \mathcal{N}(\psi_0, \Psi_0^{-1})$$

とする. Gibbs サンプラーでは, 次の条件付き分布が必要である:$p(\beta|\sigma^2, \rho, y, y^{c*})$, $p(\sigma^2|\beta, \rho, y, y^{c*})$, $p(\rho|\beta, \sigma^2, y, y^{c*})$, そして $p(y^{c*}|\beta, \sigma^2, \rho, y,)$ である. y^{c*} の条件付きの構造は

$$p(y^{c*}|y, \beta, \rho, \sigma^2) \sim \text{truncated } \mathcal{N}_{v^c}(\eta^c, \Sigma^c)$$

である. β の正規事前分布を尤度関数を結びつけて, 次式を得る.

$$p(\beta|y, \rho, \sigma^2, y^c) \sim \mathcal{N}(\beta^*, \sigma^2, A^{*-1})$$

ここで, $\beta^* = (A_0 + x^{0\prime}x^0)^{-1}(A_0\beta_0 + x^{0\prime}y^0)$, $A^* = (A_0 + x^{0\prime}x^0)$ である. β と ρ が与えられたときの σ^2 の尤度はすぐに得られ, 仮定された事前分布は

$$p(\sigma^2|\beta) \sim \mathcal{IG}((\nu_0 + k)/2, (\delta_0 + Q_\beta)/2), \quad p(Q_\beta) = (\beta - \beta_0)' A_0 (\beta - \beta_0)$$

を意味する. σ^2 の条件付き分布は

$$p(\sigma^2|y, \rho, \beta, y^{c*}) \sim \mathcal{IG}((T - 1 + \nu_0 + k)/2, (\delta_0 + Q_\beta + d_\beta)/2)$$

ここで, $d_\beta = (y^0 - x^0\beta)'(y^0 - x^0\beta)$ である. $p(\rho|y, \sigma^2, \beta, y^{c*})$ の条件付き事後分布は, 次のように導出される. AR(1) モデルでは, 誤差 $\epsilon_t = y_t^* - x_t\beta$ は次のように書け

$$\epsilon_t = \rho\epsilon_{t-1} + u_t, \quad t = 2, \ldots, T$$

ρ は単変量正規分布に従う. 尤度関数と ρ の正規事前密度を結合すると, 次の切断された正規事後密度を得る.

$$p(\rho|y, \sigma^2, y^{c*}, \beta) \propto \text{truncated}\mathcal{N}(\psi, \Psi^{-1})$$

ここで, $\epsilon = (\epsilon_1, \ldots, \epsilon_T)'$, $\psi = \Psi^{-1}(\Psi_0\psi^0 + \sigma^{-2}M'\epsilon)$, $\Psi = (\Psi_0 + \sigma^2 M'M)$ である. そして M は $(T \times 1)$ ベクトルで, その第 i 行は ϵ_{t-1} である.

2.4 計量経済分析における MCMC シミュレーション法

MCMC 法は, このようにして**目標分布** (*target distribution*) から標本 (多重観測値) を生成するシミュレーション技法である. マルコフ過程の推移核は, その極限の不変分布が目標分布になるという性質を持つように定式化され, マルコフ連鎖は, コンピュータで生成されるモンテカルロ・シミュレーションを多数回反復する. そして一時的な不安定状態の後, 様々な条件の下での出力が目標分布からのサンプルになる. そのような一般的な方法は, Metropolist-Hastings(MH) アルゴリズムとして知られている. このアルゴリズムでは, マルコフ連鎖の次の値は提案密度から生成され, 現在の点での密度と比べたときの候補点での密度に従って, 受容したり棄却されたりする. MH 法の特別な場合は Gibbs サンプリング・アルゴリズムで, 次の標本は一連の完全条件付き分布からのランダム・ベクトルの部分要素をサンプリングすることによって得られる. その他の MCMC 法には, ハイブリッド MH と棄却サンプリング (Tierney, 1994) と EM アルゴリズムの確率バージョン (Celeux and Diebolt, 1985) などの方法が含まれる.

MCMC を用いた方法は統計学上での応用分析できわめて役に立ち, 多くは複雑な, 高次元問題に適用するのが難しい伝統的な独立サンプリング法よりもずっと役に立つことが証明されている. MCMC 法は目標密度の正規化定数が分からない場合にも適用することができ, このことは目標 (事後) 密度の正規化定数がほとんどわからないベイズ分析では非常に重要である. 加えて, 扱いにくい尤度関数を持つモデルをシミュレートできるような MCMC スキームを組み立てることができようになる. これは通常, 特に Gibbs サンプリングではパラメータ空間を欠測データ, あるいは潜在変数を含めるように拡張する方法である**データ拡大**によって行う.

以下で, 代表的なモデルについて Chib and Greenberg (1996) に沿って説明する.

2.4.1 応用例

MCMC シミュレーション法をいろいろな種類の計量経済モデルに適用する方法を, 以下で示すことにする. まず最初に, Gibbs サンプラーをデータ拡大のない場合に適用することができ, シミュレーションが標準的な分布だけからなる単純な例から始める. 目的はこの方法の考え方を示し, それを他の場合に適用する方法の役に立つことである. 例を示す前に, この後使う事前密度に対して次の仮定を導入する. ベクトル β は $\mathcal{N}_k(\beta_0, B_0^{-1})$ に, 分散 σ^2 は逆ガンマ分布 $\mathcal{IG}(\nu_0/2, \delta_0/2)$ に, そして精度行列 Ω^{-1} はウィッシャート分布 $\mathcal{W}(\rho_0, R_0)$ に従う. 事前密度のハイパーパラメータはゼロの添え字を付け, 既知であると仮定する.

SUR:Seemingly Unrelated Regression Model

最初の例はSURモデルで, 計量経済学では広く用いられているモデルである. 誤差項が正規分布に従うとの仮定の下で, 観測されたデータ y_{it} は

$$y_{it} = x'_{it}\beta_i + \epsilon_{it} \qquad \epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{pt})' \sim iid.\mathcal{N}_p(0, \Omega),$$

$$1 \leq i \leq p,\ 1 \leq t \leq n$$

ここで, β_i は $k_i \times 1$ のベクトルで Ω は正定符号行列である. 各期で観測値を縦にスタックすると, モデルを次ぎのようなベクトル型で書き直すことができる.

$$\boldsymbol{y}_t = \boldsymbol{X}_t\boldsymbol{\beta} + \boldsymbol{\epsilon}_t$$

ここで $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{pt})', \boldsymbol{X}_t = \mathrm{diag}\,(x'_{1t}, \ldots, x'_{pt}), \boldsymbol{\beta} = (\beta'_1, \ldots, \beta'_p) : k \times 1$ そして $k = \sum_i k_i$ である. $p = 1$ の時, 単一方程式正規回帰モデルである. データ $\boldsymbol{Y}_n = (y_1, \ldots, y_n)$ の標本に対する最尤推定値は反復過程を通じてのみ得られ, これらの推定量の有限標本分布は取り扱いにくい. 対照的に, Gibbs サンプリング・アルゴリズムは, このモデルに対する厳密な小標本ベイズ推定を与える (Percy, 1992; Chib and Greenberg, 1995a).

(β, Ω^{-1}) に関する事前情報が密度 $\pi(\beta)\pi(\Omega^{-1})$ によって表されると, β と Ω^{-1} (精度行列) は独立と仮定していることになる. すると, (事前密度と尤度関数の

積に比例する) パラメータの事後密度は, 次式で与えられる.

$$\pi(\beta)\pi(\Omega^{-1}) \times |\Omega^{-1}|^{n/2} \exp\left[-\frac{1}{2}\sum_{t=1}^{n}(y_t - X_t\beta)'\Omega^{-1}(y_t - X_t\beta)\right]$$

これは, (正規化定数が未知の) 目標密度であり, ここからシミュレートする. さて, もし β と Ω^{-1} をパラメータの 2 つのブロックとして扱うと, 完全条件付き密度 $p(\beta|Y_n, \Omega^{-1})$ と $p(\Omega^{-1}|Y_n, \beta)$ は簡単にシミュレートできる. 特に, すでに述べた事前分布のもとでは

$$p(\beta|Y_n, \Omega^{-1}) \sim \mathcal{N}(\hat{\beta}, B_n^{-1}) \quad \text{と} \quad p(\Omega^{-1}|Y_n, \beta) \sim \mathcal{W}_p(\nu_0 + n, R_n)$$

ここで $\hat{\beta} = B_n^{-1}(B_0\beta_0 + \sum_{t=1}^{n} X_t'\Omega^{-1}y_t)$, $B_n = (B_0 + \sum_{t=1}^{n} X_t'\Omega^{-1}X_t)$, そして $R_n = [R_0^{-1} + \sum_{t=1}^{n}(y_t - X_t\beta)'\Omega^{-1}(y_t - X_t\beta)']^{-1}$ である. これらの 2 つの分布を Gibbs アルゴリズムでシミュレートすると, $\beta^{(i)}$ が周辺密度 $\pi(\beta|Y_n), \Omega^{-1(i)} \sim \pi(\Omega^{-1}|Y_n)$ に従って分布するような標本 $\{\beta^{(i)}, \Omega^{-1(i)}\}$ を生成する. そして $p(\beta^{(i)}, \Omega^{-1(i)})$ は目標 (同時) 密度に従って分布する. 抽出されたサンプルは, インポータンス・サンプリング関数や尤度関数を評価することなしに得られる.

Tobit と Probit 回帰モデル

SUR モデルでは, Gibbs サンプラーをモデルのパラメータに直接適用した. しかし直接適用できない場合でも, 扱いやすい完全条件付き分布のセットは, tobit モデルや probit モデルのようにパラメータ空間を直接観測できない潜在データで拡張することによって得ることができる. 興味あることに, サンプラーで定義されるパラメータ空間が非常に大きい (probit モデルの場合はサンプルサイズよりも大きい) と, シミュレーションにおけるブロックの数はきわめて小さくなる (tobit モデルでは 3, 二値 probit モデルでは 2 である).

Tobin の打ち切り回帰モデル (censored regression model) では, 観測値 y_i は次のように生成される.

$$z_i \sim \mathcal{N}(x_i\beta, \sigma^2) \quad \text{そして} \quad y_i = \max(0, z_i), \qquad 1 \le i \le n$$

n 個の独立な観測値のセットが与えられると, β と σ^2 に対する尤度関数は次のように表すことができる.

$$\prod_{i \in C}[1 - \Phi(x_i'\beta/\sigma)] \prod_{i \in C'}(\sigma^{-2}) \exp\left[-\frac{1}{2\sigma^2}(y_i - x_i'\beta)^2\right]$$

ここで C は一組の打ち切り観測値, そして Φ は標準正規確率変数の累積分布関数である. 明らかに, この関数は (事前分布を掛けた後)Gibbs サンプリング・アルゴリズムのなかで簡単に扱えない. 計量経済学で Gibbs サンプリングを適用した最初の研究の一つである Chib (1993) は, パラメータ空間が打ち切り観測値に対応する潜在データによって拡張される場合には, 非常に簡単になることを示した.

その理由は, ベクトル $z = \{z_i\}, i \in C$ が利用できるためである. y_z は $n \times 1$ ベクトルで, もし i 番目の観測値がセンサーされていなければ i 番目の要素は y_i で, センサーされていれば z_i である. ブロックを β, σ^2, z とし, それぞれの完全条件付き密度を $p(\beta|Y_n, z, \sigma^2)$, $p(\sigma^2|Y_n, z, \beta)$, $p(z|Y_n, \beta, \sigma^2)$ とした場合の Gibbs サンプリング・アルゴリズムを考えると, これらの分布はすべて扱いやすいことが分かり, Gibbs シミュレーションはすぐに適用可能である. 最初の 2 つの分布は次のようになる.

$$p(\beta|y_z, \sigma^2) \sim \mathcal{N}_k(\hat{\beta}, (B_0 + \sigma^{-2}X'X)^{-1})$$

そして

$$p(\sigma^2|y_z, \beta) \sim \mathcal{IG}\left(\frac{\nu_0 + n}{2}, \frac{\delta_0 + \delta_n}{2}\right) \tag{2.7}$$

ここで $X = (x_1, \ldots, x_n)', \hat{\beta} = (B_0 + \sigma^{-2}X'X)^{-1}(B_0\beta_0 + \sigma^{-2}X'y_z)$, そして $\delta_n = (y_n - X\beta)'(y_n - X\beta)$. 他方, 潜在データの完全条件付き分布は n 個の独立分布の積に簡略化される. $p(z|Y_n, \beta, \sigma^2) = \prod_{i \in C} p(z_i|y_i = 0, \beta, \sigma^2)$, ここで

$$p(z_i|y_i) = 0, \; p(\beta, \sigma^2) \sim \mathcal{TN}_{(-\infty,0)}(x_i'\beta, \sigma^2), \qquad i \in C$$

は, 区間 $(-\infty, 0]$ でサポートされる切断正規分布である. $\mathcal{TN}_{(a,b)}(\mu, \sigma^2)$ からシミュレートするには, 最初に一様確率分布 U をシミュレートし, 次に $\mu + \sigma\Phi^{-1}\{p_1 + U(p_2 - p_1)\}$ として必要なサンプルを抽出する. ここで, Φ^{-1} は正規分布の逆分布関数で, $p_1 = \Phi[(a - \mu)/\sigma]$ と $p_2 = \Phi[(b - \mu)/\sigma]$ である. あるいは Geweke (1991) の方法をこの分布をサンプルするのに用いることができる. たとえば, β と σ^2 の分布が, 潜在データが与えられたときのセンサーされたデータと独立であるような場合に観測される, 条件付独立になるように単純化することは通常データ拡大を行ったときに生じ, これはなぜデータ拡大が有効な道具であるかを示している.

データ拡大が有効であることは probit モデルでも明白である. n 個の独立観測値 $Y_n = \{y_i\}$ が与えられており, 各 y_i は $\text{Prob}(y_i = 1) = \Phi(x_i'\beta)$ であるベルヌーイプロセスに従う. このモデルやこのクラスに入る他の多くのモデルに対して, Albert and Chib (1993b) は, 潜在的なガウス型データを Gibbs サンプリング・アルゴリズムにおけるもう一つの未知パラメータとして導入する単純で, 有力な方法を開発した. 次式

$$z_i = x_i'\beta + u_i, \qquad u_i \sim iid.\mathcal{N}(0,1), \quad \text{そして} \quad y_i = I[z_i > 0] \tag{2.8}$$

は probit モデルを表すことを示した. (データ拡大を含む)Gibbs サンプリング・アルゴリズムは, 次の条件付き分布を通じて定義される.

$$p(\beta|Y_n, Z_n) \overset{d}{=} p(\beta|Z_n), \quad \text{そして} \quad p(Z_n|Y_n, \beta) \overset{d}{=} \prod_{i=1}^{n} p(z_i|y_i, \beta)$$

ここで, $Z_n = (z_1, \cdots, z_n)'$ である.

β の完全条件付き分布は, y_z を Z_n で置き換え $\sigma^2 = 1$ とした (2.7) 式と同じ形をしている. 完全条件付き密度 $p(Z_n|Y_n, \beta)$ は, 独立な項の積として分解され, $y_i = 1$ あるいは $y_i = 0$ に依存している. (2.8) から, $z \leq 0$ なら $y_i = 0$, $z_i > 0$ なら $y_i = 1$ となる. そこで

$$p(z_i|y_i) = \begin{cases} 0, & \beta \sim \mathcal{TN}_{(-\infty,0)}(x_i'\beta, 1) \\ 1, & \beta \sim \mathcal{TN}_{(0,\infty)}(x_i'\beta, 1), \end{cases} \qquad 1 \leq i \leq n$$

この MCMC アルゴリズムは, 自由度が ν の独立な t リンク関数を持つモデルを推定するように簡単に修正することができる (Albert and Chib, 1993a). t 分布は, 正規分布と混合するガンマ分布 $\Gamma(\nu/2,\ \nu/2)$ のスケール混合であるという結果から, パラメータ空間をこのガンマ変数で各観測値ごとにさらに拡張することができる. 完全条件付き密度は再び扱える (この考え方を線形回帰に用いたものは, Carlin and Polson, 1991; Geweke, 1993b である). Albert and Chib (1993b) は ν を未知とし, probit モデルの一般的な頑健性を導いている.

$AR(p)$ 誤差の回帰モデル

この節では MH アルゴリズムを用いたシミュレーションの例を示す. ARMA(p, q) 誤差のある回帰モデルの詳細な分析は, Chib and Greenberg (1994) と Marriott, Ravishanker, and Gelfand (1995) に見られる.

次のモデルを考える.

$$y_t = x_t'\beta + \epsilon_t, \qquad 1 \le t \le n \tag{2.9}$$

ここで, y_t はスカラーの観測値である. 誤差項は定常な AR(p) プロセス

$$\epsilon_t - \phi_1\epsilon_{t-1} - \cdots - \phi_p\epsilon_{t-p} = u_t \quad \text{or} \quad \phi(L)\epsilon_t = u_t \tag{2.10}$$

によって生成されていると仮定する. ここで, $u_t \sim iid.\mathcal{N}(0,\sigma^2)$ そして $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ はラグオペレータ L の多項式である. 定常性の仮定は, $\phi(L)$ の根が単位円の外にあることを意味しており, これは $\phi = (\phi_1, \ldots, \phi_p)$ を $\Re^p$ の部分空間内 (S_ϕ) にあるように制約を課すことになる. この制約に従うようにするために, ϕ の事前分布を $\mathcal{N}_p(\phi|\phi_0, \Phi_0^{-1})I_{S_\phi}$ とし, 正規分布は定常区域で切断されている (ただし, β と σ^2 に対しては標準的な事前分布を仮定する). このモデルに対する尤度関数は次のように表される.

$$f(Y_n|\beta,\phi,\sigma^2) = \Psi(\phi) \times (\sigma^2)^{-(n-p)/2} \exp\left[-\frac{1}{2\sigma^2}\sum_{t=p+1}^{n}(y_t^* - x_t^{*'}\beta)^2\right]$$

ここで, $t \ge p+1$ に対して $y_t^* = \phi(L)y_t, x_t^* = \phi(L)x_t$, そして

$$\Psi(\phi) = (\sigma^2)^{-p/2}|\Sigma_p|^{-1/2}\exp\left[-\frac{1}{2\sigma^2}(Y_p - X_p\beta)'\Sigma_p^{-1}(Y_p - X_p\beta)\right] \tag{2.11}$$

は最初の p 個の観測値の (定常な) 密度である. 以前に, $Y_p = (y_1, \ldots, y_p)', X_p = (x_1, \ldots, x_p)'$, と $\Sigma_p = \Phi\Sigma_p\Phi' + e_1(p)e_1(p)'$, ここで

$$\Phi = \begin{bmatrix} \phi_{-p} & \phi_p \\ I_{p-1} & 0 \end{bmatrix}$$

$e_1(p) = (1, 0, \ldots, 0)'$, そして $\phi_{-p} = (\phi_1, \ldots, \phi_{p-1})'$ である.

事後密度を次のようにシミュレートする. まず, 3 つの事実がある. 最初に, Gibbs の方法は β, ϕ, σ^2 をブロックとしてとるとよい. 第 2 に, β と σ^2 の完全条件付き分布は, 標本密度の 2 つの指数項を結びつけた後で簡単に得られる. 第 3 に, ϕ の完全条件付き密度は, MH アルゴリズムでシミュレートすることができる. 次に, いくつか詳細に検討する.

$Y_p^* = Q^{-1}Y_p$ と $X_p^* = Q_{-1}X_p$ と定義し, ここで Q は $QQ' = \Sigma_p$ を満足する. $y^* = (y_1^*, \ldots, y_n^*)'$ とし, X^* に対しても同様である. 最後に, $e = (e_{p+1}, \ldots, e_n)'$

とし, E は $n-p \times p$ 行列で, t 番目の行は $(e_{t-1}, \ldots, e_{t-p})$ で与えられ, ここで $e_t = y_t - x_t'\beta, t \geq p+1$ である. 完全条件付き分布は, 次のようになることを簡単に示すことができる.

$$
\begin{aligned}
p(\beta|Y_n, \phi, \sigma^2) &\sim \mathcal{N}_k(\hat{\beta}, B_n^{-1}) \\
p(\phi|Y_n, \beta, \sigma^2) &\propto \Psi(\phi) \times \mathcal{N}_p(\hat{\phi}, \Phi_n^{-1}) I_{S_\phi} \\
p(\sigma^2|Y_n, \beta, \phi) &\sim \mathcal{IG}\left(\frac{\nu_0 + n}{2} \frac{\delta_0 + \delta_\beta}{2}\right)
\end{aligned}
\tag{2.12}
$$

ここで $\hat{\beta} = B_n^{-1}(B_0\beta_0 + \sigma^{-2}X^{*\prime}y^*), B_n = (B_0 + \sigma^{-2}X^{*\prime}X^*), d_\beta = ||y^* - X^*\beta||^2, \hat{\phi} = \hat{\Phi}_n^{-1}(\Phi_0\phi_0 + \sigma^{-2}E'e)$, そして $\hat{\Phi}_n = (\Phi_0 + \sigma^2 E'E)$ である.

β と σ^2 の完全条件付き分布は簡単にシミュレートできる. ϕ をシミュレートするためには, 候補生成密度として $\mathcal{N}_p(\hat{\phi}, \hat{\Phi}_n^{-1}) I_{S_\phi}$ をもつ MH 独立連鎖を採用する. すると, MH ステップは次のように実行する. i 番目のイタレーションで, 平均 $\hat{\phi}$ で共分散 $\sigma^{2(i)}\Phi_n^{-1}$ の正規密度から, 候補 $\phi^{(i+1)}$ を抽出する. もしそれが定常性を満足するなら, この点に確率

$$
\min\left\{\frac{\Psi(\phi^{(i+1)})}{\Psi(\phi^{(i)})}, 1\right\}
$$

で移動する. そして, そうでない場合は $\phi^{(i+1)} = \phi^{(i)}$ とおく. Chib and Greenberg (1994) はこのアルゴリズムの収束の十分条件を確かめ, いくつかの実証例を示している.

多重時点における構造変化

時系列データをあつかったモデルでの変化点モデルは重要である. 計量経済分析では, 構造変化問題といわれ従来は構造変化点があるか否か, あったとすればいつあったのかに関心が持たれていた. 最初は構造変化点がただ一つの場合について分析され, 多重変化点についての分析は最近行われるようになった. 古典的な方法としては変化点が既知の場合についての分析は, Chow(1960), Toyoda(1974), Schmidt and Sickles(1977) が, 変化点が未知の一般的な場合は BDE(1975) による CUSUM, CUSUMSQ テストがよく知られている. ベイジアンの立場では, 最近 Chib(1997) がパラメータ・ベクトルの動きを記述するのに hidden Markov chain に基づく新しいモデル化を提案している. 多重変化点を扱

うために制約付きの遷移確率行列を使用し, 状態変数が現在の値にとどまるか次のより高い値にジャンプするかを離散時間離散状態のマルコフ過程によってモデル化した. モデルはマルコフ連鎖モンテカルロシミュレーション法によって, 混合モデルに基づいて推定される.

また, 単一変化点についての分析は, Carlin, Gelfand, and Smith (1992) が, 多重変化点についての分析は, Inclan (1993), Stephens (1994), Yao (1984), Barry and Hartigan (1993), West, Müller and Esobar (1994), 等によって研究されている.

$Y_n = \{y_1, y_2, \ldots, y_n\}$ は時系列で, そのパラメータ ξ_t は未知の変化点 $\Upsilon_m = \{\tau_1, \ldots, \tau_m\}$ で変化し, 他は一定であると仮定する.

$$\xi_t = \begin{cases} \theta_1 & \text{if} \quad t \leq \tau_1 \\ \theta_2 & \text{if} \quad \tau_1 < t \leq \tau_2 \\ \vdots & \text{if} \quad \vdots \\ \theta_m & \text{if} \quad \tau_{m-1} < t \leq \tau_m \\ \theta_{m+1} & \text{if} \quad \tau_m < t \leq n \end{cases}$$

ここで $\theta_k \in \Re^d$ であり, パラメータ・ベクトル $\Theta = (\theta_1, \ldots, \theta_{m+1})$ の推定と, 未知の変化点 $\Upsilon_m = (\tau_1, \ldots, \tau_m)$ を検出し, 変化点の数が異なるモデルと比較することを目的としている.

この多重変化点モデルに関して多くの研究があり, 1つの問題は ξ_t のジャンププロセスを決めることである. ベイジアンでは, 変化点で生じる別のレジームのパラメータである τ_k と θ_k の結合事前分布を求めることと同じである.

通常は階層モデルの形で定式化される：最初に, 以前の変化点での変化の確率分布をモデル化し, 次に現在の変化点と以前のパラメータで, 現在のレジームでのパラメータのプロセスをモデル化し, 最後にパラメータと変換点によってデータが生成されるとする. パラメータが与えられたときの, データの分布を考える. $Y_i = (y_1, \ldots, y_i)$ を時点 i までの経過とし, $Y^{i,j} = (y_i, y_{i+1}, \ldots, y_j)$ を時間 i から j までの経過とする. すると, (Θ, Υ) の下でのデータの同時密度は

$$f(Y_n|\Theta, \Upsilon_m) = \prod_{k=1}^{m+1} f(Y^{\tau_{k-1}+1,\tau_k}|Y_{\tau_{k-1}}, \theta_k, \tau_k)$$

で表される. ここで $\tau_0 = 0, \tau_{m+1} = n$ である.

各期間での離散的な確率変数 s_t を導入し, 時点 t でのシステムの状態を表し, これは整数値 $\{1, 2, \ldots, m+1\}$ で, そこから特定の観測値 y_t がとられるレジームを示す. すなわち $s_t = k$ は, 観測値 y_t が $f(y_t|Y_{t-1}, \theta_k)$ からとられることを示す. 変数 s_t は, 制約付の遷移確率行列を持つ離散時間, 離散状態マルコフ・プロセスとしてモデル化されるので変化点モデルになる. 1 期先遷移確率行列は, 次のように表される.

$$P = \begin{pmatrix} p_{11} & p_{12} & 0 & \cdots & 0 \\ 0 & p_{22} & p_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \vdots & 0 & p_{mm} & p_{m,m+1} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

このモデルは, ジャンプ確率 $p_{ii}(i \leq m)$ がレジームに依存し, 状態の遷移は変化点 $\Upsilon_m = (\tau_1, \ldots, \tau_m)$ を識別する一般的な変化点モデルとして見ることができる. k 回目の変化は, $s_{\tau_k} = k$ と $s_{\tau_k+1} = k+1$ なら τ_k で起こる. もう一つの見方は hidden マルコフモデルとして見るもので, hidden 状態変数 s_t の遷移確率に前に述べたような制約を課す. このモデルの見方は MCMC の計算図式の基礎になっている. パラメータの事前密度 $\pi(\Theta, P)$ を決め, データ Y_n が得られると, 事後密度 $\pi(\Theta, P|Y_n) \propto \pi(\Theta, P) f(Y_n|\Theta, P)$ に関心を集中する. ここで, π は (Θ, P) の事前密度と事後密度を表す. この事後密度は, パラメータ空間を観測されない状態 $S_n = (s_1, \ldots, s_n)$ で拡大した後, マルコフ連鎖モンテカルロ法によって最も適切に要約される. 言い換えると, モンテカルロ・サンプリングを事後密度 $\pi(S_n, \Theta, P|Y_n)$ に適用することになる.

他方, Kozumi and Hasegawa (1998) はベイズの観点から構造変化を調べる新しい方法を提案した. この方法は Dirichlet 事前分布を用いた階層モデルを基にしている. Dirichlet プロセスの特徴は, 構造変化を検出するのに役立つ離散性にあり, それを利用して事前分布を形成する. この論文で彼らは多重変化点を認め, その数が未知であるモデルに対するベイズ・アプローチを提案した. Dirichlet プロセスは魅力があるが, 計算が難しいため応用は単純な場合に限定されていたが, 最近のマルコフ連鎖モンテカルロ (MCMC) 法の発展により, この点は解消された (Gelfand and Smith, 1990; Besag and Green, 1993 参照). Dirichlet プロセスは従来は密度の推定に利用されていたが (Escobar, 1994; Escobar and West,

1995), Dirichlet プロセスの離散性がベイズ的なモデル化にも有効な道具であることが最近の多くの研究で示されている (Kuo and Mallich, 1997; Mukhopadhyay and Gelfand, 1997; West, Müller and Esobar, 1994) この離散性の性質を用いて構造変化を検出する.

実証分析では, 1957 年から 1995 年の年データを用いて日本の政府支出に構造変化があるかどうかを検証し, それを CUSUM と CUSUMSQ テストの結果と比較している.

次の単純な線形回帰モデルを考える.

$$y_t = \beta_0 + \beta_1 z_t + u_t, \qquad u_t \sim \mathcal{N}(0, \sigma^2), \quad t = 1, \ldots, T$$

ここで, y_t と x_t はそれぞれ従属変数と独立変数であり, 構造変化を考えるのでパラメータ・ベクトルがある変化点で変化するすることを認め, 次のように書き直す.

$$y_t = \beta_{0t} + \beta_{1t} z_t + u_t, \qquad u_t \sim \mathcal{N}(0, \sigma_t^2)$$

ベイズ分析ではパラメータ・ベクトル $\theta_t = (\beta_{0t}, \beta_{1t}, \sigma_t^2)$ の分布が必要であり, 前述の Chib(1997) は θ_t の分布を hidden マルコフ連鎖でモデル化した.

$\{\theta_t\}_{t=1}^T$ を Dirichlet プロセスからの標本とする. 次に, θ_t を $k(< T)$ 個のグループに分割し, 同じグループ内の θ_t は等しく, 異なるグループの θ_t は異なり, k 個の θ_i の値 $\theta^* = (\theta_1^*, \ldots, \theta_k^*)$ は G_0 からサンプルされる. 何回構造変化が起こったか, いつ構造変化が起こったのかは, 異なる θ^* を調べると分かるので, Dirichlet プロセスの離散性は構造変化点を検出するのに適しているといえる.

パラメータベクトル $\theta_t = (\beta_{0t}, \beta_{1t}, \sigma_t^2)$ の分布 G は未知で, G の事前分布は Dirichlet プロセスにしたがうと仮定する.

$$y_t \sim \mathcal{N}(x_t' \beta_t, \sigma_t^2), \quad \theta_t \sim G, \quad G \sim \mathcal{DP}(\alpha G_0)$$

ベイズモデルとするため, 残りのパラメータに対し事前分布を決める. 基本的な分布 G_0 に対しては, 正規分布とガンマ分布とする.

$$dG_0 = \mathcal{N}(\mu, \Sigma)\mathcal{IG}(n_0/2, \tau/2)$$

さらに, 階層モデルでの μ, Σ, τ について, 事前分布を次のようにおく.

$$\mu \sim \mathcal{N}(\mu_0, V_0)$$

$$\Sigma^{-1} \sim \mathcal{W}(\nu_0, \Sigma_0)$$
$$\tau \sim Ga(m_0/2, \tau_0/2)$$
$$\alpha \sim Ga(a_0, b_0)$$

多重変化点を認めるがその数は未知であるモデルで, 誤差分散は変化点にしたがって変化する. ここで $x_t = (1, z_t)'$ と $\beta_t = (\beta_{0t}, \beta_{1t})'$ である. このモデルは階層構造をしており, 複雑そうに見えるが Gibbs サンプリングを用いて事後密度の推定を行うことができる.

その他のモデル

今までに述べたモデルに加えて, その他のモデルとしては MCMC 法とデータ拡大を用いる Gibbs サンプリングは特に役に立つ分野である. 回帰分析の枠組みでは, 欠測値をサンプラーに加えてパラメータの分布から標本を生成する. 多対プロビットモデルの重要なクラスは, データ拡大を通じた MCMC シミュレーションによって分析することができ, これらは Albert and Chib (1993a), McCulloch and Rossi (1994), Geweke, Keane, and Runkle (1994) で議論されている. McCulloch and Rossi (1994) にはこの分野での Gibbs サンプリングが役立つ拡張した議論が例とともになされている. 変量効果のある一般線形モデルへの応用は, Zeger and Karim(1991) が参考になる. もう一つの重要な分野は混合 (mixture) モデルであり, このモデルではサンプルの各観測値は K 種類の母集団の一つから生じたものである. 2 種類のモデルが研究されている. 最初のものは, 母集団が一つの観測値から次のものへ独立にサンプルされる (Diebolt and Robert, 1994). 2 番目のものは, 母集団はマルコフプロセスに従ってサンプルされ, これはマルコフ・スウィッチング・モデル (Albert and Chib, 1993b; Chib, 1993) と呼ばれる. MCMC 法の多様さを示している新しい計量分析の応用は, 引き続き多数現れている.

本稿で扱わなかった重要な分野で, 最近多くの研究が集中的になされている金融時系列 (株価, 収益率, 外国為替, 先物等の日次, 週次などの高周波データ) の変動に関する計量分析がある. これらには, Stochastic Volatility Model, Garch Model, Semi-Non Parametric Model, 線形状態空間モデル, 非線形非ガウス・フィ

ルター・モデル, 等が含まれる.

2.5 おわりに

このサーベイでは, ベイズ統計学の計量経済分析への応用に関連するいくつかの重要な問題について議論した. 本サーベイは2つの部分に分けられる : ベイジアン計量経済学の考え方と実証分析への応用に関する問題と, 大きな部分を占める計算問題である. 第2章では, ベイジアン計量経済学の簡単な解説を行った. 第3章では, 応用研究を行う場合におけるベイジアン・パラダイムのいくつかの関連する利点と欠点を述べた. ここで, 単位根とEBAに特に焦点を当てたベイジアンの実証分析の簡単な要約を行った. 第4章以降では, ベイズ型の実証分析を行う場合に重要な計算問題について, 最初に基本的な計算問題について第4章で述べ, ベイジアン研究者が実証研究で利用するいくつかの役に立つ道具について述べた. 特に, 4.1節では通常出合う問題に接近するためのいくつかの基本的な技術を示し, これらの多くの技法は古典的な研究者が用いているTSPやRATSのようなパッケージと全く同じような方法で, パッケージプログラムを用いて簡単に導入できる. パッケージプログラムが利用できない場合でも, これらの手法のプログラムに必要な時間は, GAUSSやFortranのような言語を使った基本的なプログラムの能力を持っている研究者にとっては最小で済む. 第5章では, ある種のクラスの問題を分析するのに適している, もう少し手の込んだベイジアンの手法について述べた. この方法をプログラムするのに必要な要件は基本的な手法の場合よりもずっと大きいにも拘わらず, 僅かなプログラムの知識しか持っていない計量経済学者の能力をまだ越えていない. 第6章では, 現在の計量経済分析を用いた実証研究において主要な手法になったマルコフ連鎖モンテカルロ法について, Metropolis-Hastingアルゴリズム, Gibbsサンプリング, そしてデータ拡大・アルゴリズムについて述べた. MCMC法の理論と実際における最近の発展のサーベイを, 計量経済分析への応用に重点を置いて, このアルゴリズムをデータ拡大と結合する方法を用いてベイズ推定への体系的方法を示した. この考え方をセンサー化モデル, 離散応答モデル, パネルデータ・モデル, 自己回帰モデル, そして時変パラメータモデルの枠組みの中で示したが, この考え方は多くの他の計量モデルに適用できる.

MCMC 法が好まれる多くの理由は, 他の方法で分析するのが今まで難しかったモデルを分析することを可能にするためである. この方法を利用できるのは, 厳密に定式化したモデルと事前分布からなるベイズ分析だけではない. すでに示したように, 扱いにくい尤度関数を含む多くのモデルは MCMC 法でシミュレートできる. 特に予測, モデルと事前分布の摂動, そしてモデル適合性に関係する様々な推定方程式は, シミュレーションの結果を用いて効果的に記述することができる.

最近では多くの分野での実証研究問題にベイズ法を利用した例が増えてきている. それらの応用例, 研究動向などは国際ベイズ分析学会 (ISBA), 4 年ごとに開催されるベイズ統計学に関する国際バレンシア研究集会, それに米国統計学会ベイズ統計部会の年次大会などで発表される数多くの論文や Proceedings に見られる.

最後に, このサーベイではできるだけ単純な形で, ベイジアンの手法を用いて実証分析を行おうとする応用経済分析が利用できる手法を示すことを試みた. したがって, マルコフ連鎖モンテカルロ法で問題になる多くの事柄, 特に収束診断法を含む多くの問題点について説明する余裕がなかったので省略した. また, ここで説明した方法で重要な多くの証明や他で示されている式の導出の詳細も省略したので, 参考文献を参考にしてほしい. この分野の進展は早いので, 最新の時系列分析や金融時系列モデル等の分野についての多量の研究がなされているが, この分野についてのサーベイは次回に譲ることにしたい. また関連する参考文献も完全ではない.

ベイズコンピュータ・パッケージについて

ベイズ部分のある標準的な計量分析プログラム

いくつかの標準的なコンピュータパッケージは、よく用いられる経済問題に対して役立つ部分がある。たとえば、RATS マニュアルはベイズ型ベクトル自己回帰モデル (BVAR) に対して RATS を用いたり、SHAZAM はパラメータに線形不等号制約に従う回帰モデルの分析に対するベイズモンテカルロ積分の節に、大きな部分をさいている。

A.1 初期に開発されたプログラム

BAYES FOUR 汎用のベイズプログラムで、Prof. Adrian F.M. Smith と彼の同僚である Drs. John Naylor, Allan Skene, Ewart Shaw らによって University of Nottingham で開発された。プログラムではパラメータの期待値としての事後モーメントを計算する。1 次元と 2 次元の周辺事後密度と予測密度も数値的に得られ、グラフィック・ディスプレーも *S-Plus* のプロットライブラリーを利用することによって可能である。10 次元ぐらいまでの問題に対しては数値積分 (求積法) を用いて、それ以上の次元 (20 次元ぐらいまで) は非反復モンテカルロ法を用いる。`Bayes Four` は PC と Unix に対しては、Fortran 77 で書かれており、その他は `C` あるいは `C++` が使える。

BRAP (PC バージョンは PCBRAP) シカゴ大学の Arnold Zellner と同僚が書いた Fortran ベースのプログラムである。正規誤差を持つ線

形回帰モデルとして書くことができるいかなるモデルの分析が可能である。加えて、ユーザが与えた関数の (正規あるいは t インポータンス関数を用いた) モンテカルロ積分ができる。レビューは Koop(1992) で与えられている。このプログラムのコマンドは TSP のコマンドと非常に似ているので、単純な実証的問題の教育や実行に対しては、最も汎用性があるプログラムである。

BRP CORE の何人かの研究者によって書かれた Fortran ベースのプログラムである。線形モデルの処理と様々な積分手法の両方の広範囲の機能を含んでいる。このサーベイで特に関心があるのは、Poly-t 密度を取り扱う数値積分手法である。このパッケージは BRAP よりは汎用性があるが、ユーザ・フレンドリーではない。プログラムの諸機能を使うには、Fortran の知識が必要である。さらに、BRP は適度の込み入った問題に対処するのに適している。

MICRO-EBA GAUSS ベースのプログラムで EBA を分析するためのユタ大学の Richard Fowles が作成したプログラムである。

SEARCH Ed Leamer の推奨する多くの方法を行うプログラムである。

SISAM, MIXIN 2 本の Fortran ベースのプログラムで、モンテカルロ積分と混合積分を行う。汎用性はないが、このプログラムが採用している手法は、非常に複雑である。Fortran でプログラムされた poor な動きのする事後分布を持つ研究者に、このプログラムは適している。事後分布が非対称であり、従って t インポータンス関数を用いて扱うのは難しい場合には特に有益である。

A.2 特定目的のプログラム

BATS *Bayesian Analysis of Time Series* プログラムは、カルマンフィルター・モデルを含んだ動的線形モデル (DLM) と非線形モデルの推定を中心とし、APL で書かれている。もともとは University of Warwick で開発され、West, Harrison and Pole (1987) による論文に記述されている。現在は、強化されたバージョンが Pole, West and Harrison(1994) のベイズ予測と時系列の本で得られる。

bicreg, bic.logit and glib これらの3つのプログラムは、ワシントン大学の Prof. Adrian Raftery が作成した比較的短い S-plus 関数で、ベイジアンモデル選択と BIC を用いたモデル不確実性の評価を行う。最初のものは線形回帰モデルでこれを行い、2番目はロジスティック回帰で、3番目は一般線形モデルでこれを行う。

BART *Bayesian Approach to Randomized Trial* は、clinical trial monitoring と indifference zone による分析を行うためのベイズ法をプログラム化した S 関数で、ケンブリッジ大学の MRC Biostatistics Unit の Prof. David Spiegelhalter によって開発された。プログラムは正規尤度を標準事前分布を用いて更新できる。

B/D belief 解析を行うために、Wooff(1992) が作成した [B/D] と呼ばれるプログラムは、*Belief adjusted by Data* の略称で、University of Hull and Durham で開発した Pascal で書かれたプログラムである。未知の数量のネットワーク上で確信を部分的な式を与えると、プログラムは内部的な整合性 (コヒーレンス) をチェックし、他の要素とデータが存在するときにそれぞれを調整する。

Expert System 最近のベイズ的な方法の重要な応用分野はエキスパートシステムである。現在のデータセット集が与えられたときに意思決定者が利用できる情報を自動的に更新する、最も単純で自然な方法はベイズの法則を利用することである。

Hugin

BELIEF 1.2, 2.0

Neural networks software ニューラルネットワークは、エキスパートシステムと密接に関連しており、これらのモデルでの MCMC 法を利用したベイズ学習アルゴリズムが Neal(1994) に書かれており、ソフトウェアが無料で入手できる。

A.3 教育用プログラム

First Bayes Nottingham University の Prof. Tony O'Hagan によって作成され、学生が初等的なベイズ統計学を学習し、理解するのに役

立つ使いやすいパッケージである。含まれている機能の一部は次の通り

- 13 種類の標準的な分布族とこれらの混合分布のプロットと要約
- 一元配置 ANOVA と単純線形回帰モデルの分析と、位置パラメータの任意の線形結合に対する周辺分布
- ユーザが行ったデータあるいは事前分布に対する変更にしたがって、自動的に事後分布と予測分布の要約統計量を更新する。
- 現在の事後分布を事前分布に等しくおくことによるシーケンシャル分析

プログラムは APL で書かれている。

Minitab Macros Albert(1996) に書かれている Minitab 用のマクロ集。

A.4 マルコフ連鎖モンテカルロプログラム

BUGS は、英国の MRC の Medical Research Unit で開発されたシステムで、*Bayesian inference Using Gibbs Sampling* の略称で、Gibbs サンプリングを手軽に行うために便利なソフトウェアである。D. Spiegelhater を中心としたグループで開発されている。所与のモデルに対する様々な事前分布：ベルヌーイ分布、ベータ分布、正規分布、2 項分布、ディリクレ分布、指数分布、ガンマ分布、対数正規分布、多項分布、正規分布、パレート分布、ポワソン分布、一様分布、ワイブル分布にしたがう乱数を発生させることができ、完全条件付き分布からのサンプリングを与えることによって多くのベイズ問題を解くことができる。モデルの確率構造 (観測値とパラメータ) に対するモデルのスペシフィケーションと確率分布を記述する 1 組の関数からなる。モデルの記述はモデルが複雑でも簡単である。BUGS を使って既に解析されたモデルと、マニュアルに記述されているのはランダム効果を持つ一般線形モデル、生存デー

タの線形回帰分析、空間的な従属性を持つデータとノン・パラメトリック平滑化モデルである。

CODA *Convergence Diagnosis and Output Analysis* : 一般的な MCMC アウトプットやBUGS アウトプットの収束診断を調べるメニュー方式の S-plus 関数で、Gelman and Rubin (1992a)、Geweke (1992)、Ruftery and Lewis (1992) そして Heidelberger and Welch (1983) の要約統計量と収束診断を含んでいる。

Bpois Doss and Narasimhan(1994) によって書かれたベイズポワソン回帰に対する Gibbs サンプリング・プログラム。

MCSim マルコフ連鎖モンテカルロによってベイズ推定を行えるシミュレーションパッケージ。

glmm glimm S 関数

BINGO *Bayesian Inference - Numerical, Graphical and Other stuff*

参考文献

1. Albert, J and Chib, S. (1993a), "Bayesian analysis of binary and polychomotonous response data," *Journal of the American Statistical Association,* 88, 669–679.

2. Albert, J and Chib, S. (1993b), " Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts," *Journal of Business and Economics Statistics,* 11, 1–15.

3. Bauwens, Luc and M. Lubrano, eds (1995), "Bayesian and classical econometric modelling of time series," *Journal of Econometrics*

4. Bauwens, Luc , W. Polasek and H. van Dijk, eds (1996) "Bayes, Bernoullis, and Basel," *Journal of Econometrics*

5. Berger, J. (1985), *Statistical decision theory and Bayesian analysis,* (second edition), Springer–Verlag: New York.

6. Besag, J. and Green, P.J. (1993), "Spatial statistics and Bayesian computation," *Journal of the Royal Statistical Society,* B.55, 25–37.

7. Brown, R.L., Durbin, J, and Evans, J.M. (1975), "Techniques for testing the constancy of regression relationships over times," *Journal of the Royal Statistical Society,* B.37, 149–192.

8. van den Broeck, J., Koop, G., Osiewalsky, J. and Steel, M. (1994), "Stochastic frontier models : a Bayesian perspective," *Journal of Econometrics,* 61, 273–303,

9. Carlin, B.P., A.E. Gelfand, and A.F.M. Smith (1992), "Hierarchical Bayesian Analysis of change point problems," *Journal of the Royal Statistical Society C,* 41, 389–405.

10. Carlin, B.P. and Louis, T.A., (1996) *Bayes and empirical Bayes methods for data analysis*, Chapman & Hall

11. Carlin, B.P. and N.G. Polson (1991), "Inference for nonconjugate Bayesian models using the Gibbs sampler," *Canadian Journal of Statistics,* 19, 399–405.

12. Carlin, B.P., N.G. Polson and D.S. Stoffer (1992), "A Monte carlo Approach to nonnormal and nonlinear state-space modeling," *Journal of American Statistical Association,* 87, 493–500.

13. Carter, C. and R. Kohn, (1994), "On Gibbs sampling for state space models," *Biometrika,* 81, 541–554.

14. Casella, G., and George, E. (1992), " Explaining the Gibbs sampler," *American Statistician,* 46, 167–174.

15. Celeux, G. and Diebolt, J. (1985), "The SEM algorithm: A probabilistic teacher algorithm derived from the EM algorithm for the Mixture problem," *Computational Statistics Quarterly,* 2, 73–82.

16. Chib, S. (1992), "Bayes Inference for the Tobit censored regression models, " *Journal of Econometrics,* 51, 79–99.

17. Chib, S. (1993), "Bayes regression with autoregressive errors : a Gibbs sampling approach," *Journal of Econometrics,* 58, 275–294.

18. Chib, S. (1995), "Marginal Likelihood from the Gibbs Output," *Journal of the American Statistical Association,* 90, 1313–1321.

19. Chib, S. (1996), "Calculating posterior distributions and modal estimates in Markov mixture models," *Journal of Econometrics,* 75, 79–98.

20. Chib, S. (1998), "Estimation and comaprison of multiple change-point models," *Journal of Econometrics,* 86, 221–241.

21. Chib, S. and Greenberg, E. (1994), "Bayes inference via Gibbs sampling in regression models with ARMA(p, q) errors," *Journal of Econometrics,* 64, 183–206.

22. Chib, S. and Greenberg, E. (1995a),"Hierarchical analysis of SUR models with extensions to correlated serial errors and time-varying parameter models," *Journal of Econometrics,* 68, 339–360.

23. Chib, S. and Greenberg, E. (1995b),"Understanding the Metropolis-Hastings algorithm," *American Statistician,* 49, 327–335.

24. Chib, S. and Greenberg, E. (1996), "Markov chain Monte Carlo simulation methods in econometrics," *Econometric Theory,* 12, 409–431.

25. Chow, G.C. (1960), "Tests of equality between sets of coefficients in two linear regressions," *Econometrica,* 28, 591–605.

26. Cooley, T. and LeRoy, S. (1986)," What will take the con out of econometrics : a reply to McAleer, Pagan and Volker," *American Economic Review,* 76, 504–507.

27. Cooley, T. and LeRoy, S. (1991), "Identification and estimation of money demands," *American Economic Review,* 71, 825–844.

28. DeJong, D. (1992), "Co-integration and trend-stationarity in macroeconomic time series: Evidence from the likelihood function," *Journal of Econometrics*, 52, 347–370.

29. DeJong, D. and N. Shephard (1992), "The Simulation smoother for time series Models," *Biometrika,* 82, 339–350.

30. DeJong, D.N. and C.H. Whiteman (1989a) "Trends and random walks in macroeconomic time series: A reconsiderration based on the likelihood principle," *Journal of Monetary Economics*

31. DeJong, D. and Whiteman, C. (1989b), "Trend and cycles as unobserved components in GNP: A Bayesian perspective," *Proceedings of the American Statistical Association, Business and Economics Section,* 63–70.

32. DeJong, D. and Whiteman, C. (1991a), "The temporal stability of dividends and stock prices : Evidence from the likelihood function," *American Economic Review,* 81, 600–617.

33. DeJong, D. and Whiteman, C. (1991b), "Reconsidering trends and random walks in macro-economic time series," *Journal of Monetary Economics,* 28, 221–254.

34. DeJong, D. and Whiteman, C. (1991c), "The Case for trend-stationarity is stronger than we thought," *Journal of Applied Econometrics,* 6, 413–421.

35. DeJong, D. and Whiteman, C. (1996), "Trend and cycles as unobserved components in GNP: A Bayesian Perspective," in *Advances in Econometrics*, eds by T.B.Fomby and R.C. Hill Vol.11 B, 29-49.

36. DeJong, D., Nankervis, J. Savin, N. and Whiteman, C. (1992), " The power problems of unit root tests in time series with autoregressive Errors, " *Journal of Econometrics,* 53, 323–343.

37. DeJong, D. and Whiteman, C. (1993), "Estimating moving average parameters: Classical pileups and Bayesian posterios," *Journal of Business and Economic Statistics,* 11, 311–317.

38. Dempster, Laird and Rubin (1977), "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society B,* 39, 1–38.

39. Dickey, D.A. and Fuller, W.A. (1979), "Distribution of the estimators for autoregressive time–series with a unit root," *Journal of the American Statistical Association*, 74, 427–431.

40. Doan, T. (1988), *RATS: Users Manual, version 4.*, Estima, Evanston, Illinois.

41. Dorfman, J.H. (1997), *Bayesian Economic through Numerical Method,* Springer Verlag.

42. Dreze, J. (1977), "Bayesian regression Aaalysis using poly-t densities," *Journal of Econometrics,* 6, 329–354.

43. Escobar, M.D. (1994) "Estimating normal means with a Dirichlet process prior," *Journal of the American Statistical Association*, 89, 268–277.

44. Escobar, M.D. and West, M. (1995) "Bayesian density estimation and inference using mixture," *Journal of the American Statistical Association*, 89, 577–587.

45. Frühwirth-Schnatter, S. (1994) "Data augumentation and dynamic linear models," *Journal of Time Series Analysis,* 15, 183–202.

46. Gamerman, D. (1997), *Markov chain Monte Caro : Stochastic simulation for Bayesian inference*, Chapman & Hall.

47. Gelfand, A., Hills, S.E., Racine-Poon, A., and Smith, A.F.M., (1990), "Illustration of Bayesian inference in normal data models using Gibbs sampling," *Journal of the American Statistical Association,* 85, 972–985.

48. Gelfand, A. and Smith, A.F.M. (1990), "Sampling-based approaches to calculating Marginal densities," *Journal of the American Statistical Association,* 85, 398–409.

49. Gelfand, A., Smith, A.F.M., and T.M. Lee (1992), "Bayesian analysis of constrined parameter and truncated data problems," *Jounal of the American Statistical Association,* 85, 398-409.

50. Gelman, A. and Rubin, D. (1992), "A Single series from the Gibbs sampler provide a false sense of security," *Bayesian Statistics 4,* J.M. Bernardo, J.M. Berger, A.P.Dawid, and A.F.M. Smith (eds), Oxford University Press, Oxford.

51. Geman, S. and Geman, D. (1984), "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 6, 721-741.

52. Geweke, J., (1986), "Exact inference in the inequality constrained normal linear regression models," *Journal of Applied Econometrics,* 1, 127–141.

53. Geweke, J.,(1988), "Antithetic accelerations of Monte carlo integration in Bayesian econometrics," *Journal of Econometrics,* 38, 73–90.

54. Geweke, J., (1989a), "Bayesian inference in econometric models using Monte carlo integration," *Econometrica,* 57, 1317–1340.

55. Geweke, J.,(1989b), "Modeling with normal polynomial expansions," in *Economic Complexity : Chaos, Sunspots, Bubbles and Nonlinearity,* edited by W. Barnett, J. Geweke, and K. Shell, Cambridge University Press.

56. Geweke, J. (1992a), "Priors for macroeconomic time series and their applications," *Econometric Theory,* 10, 609–632.

57. Geweke, J., (1992b), "Evaluating the accurating of sampling-based approaches to the calculation of posterior moments," *Bayesian Statistics 4,* J.M. Bernardo, J.M. Berger, A.P.Dawid, and A.F.M. Smith (eds), Oxford University Press, Oxford, 169–193.

58. Geweke, J. (1997) "Posterior Simulators in Econometrics," D.M. Kreps and K.F. Wallis eds., *Advances in economics and Econometrics : Theory and Applications,* Seventh World Congress Vol.3, Cambridg University Press.

59. Gilks, W.R., Rchardson, S., and Spiegelhalter, D.J.,(1996), "Introducing Markov chain Monte Carlo," in Gilks,Richardson, and Spiegelhalter eds. *Markov Chain Monte Carlo in Practice,* Chapman & Hall.

60. Gilks, W.R. and Wild, P. (1992), "Adaptive rejection sampling for Gibbs sampling," *journal of the Royal Statistical Society, C* 41, 337–348.

61. Hastings, W.K. (1970), "Monte carlo sampling methods using Markov chains and their applications," *Biometrika,* 57, 97–109.

62. Hatanaka, M. (1996), *Time-series-based econometrics: Unit roots and cointegration*, Advanced Texts in Econometrics, Oxford University Press.

63. Heidelberger, P. and Welch, P.D. (1983), "Simulation run length control in the presence of an initial transient," *Operations Research,* 31, 1109–44.

64. Hop, J. and van Dijk, H. (1990), "SISAM and MIXIN : Two algorithms for the computation of posterior moments and densities using using Monte Carlo integration," Econometric Institute, Erasmus University Report 9031/A.

65. Inclan, C. (1993), " Detection of multiple changes of variance using posterior odds," *Journal of Business and Economic Statistics,* 11, 289–300.

66. Jacquier, E., Polson, N. and Rossi, P. (1994), "Bayesian analysis of stochastic volatility models (with discussion)," *Journal of Business and Economic Statistics,* 12, 371–417.

67. Jeffreys, H. (1961) *Theory of Probability*, 3rd ed., Oxford University Press, London.

68. Kadane, J. (1984), *Robustness of Bayesian Analysis,* North-Holland, Amsterdam.

69. Kadane, J., Dickey, J. Winkler, R., Smith, W. and Peters, S. (1980), "Interactive elicitation of option for a Normal linear models," *Journal of the American Statistical Association,* 75, 845–854.

70. Kass, R.E., L. Tierney, and J.B. Kadane (1988), "Asymptotics in Bayesian computation (with discussion)," In *Bayesian Statistics 3,* J.M. Bernardo, M.H. DeGroot, A.P. Dawid, and A.F.M. Smith, eds. Oxford University Press, 261–278.

71. Kass, R.E., L. Tierney, and J.B. Kadane (1990), "The Validity of posterior expansions based on Laplace's method," in S.Geisser, J.S.Hodges,

S.J.Press, and A. Zellner eds. *Bayesian and Likelihood Methods in Statistics and Econometrics,* 473–487.

72. Kass, R.E., Carlin, B.P., gelman, A., and Neal, R.M. (1998), "Markov Chain Monte Carlo in Proctice: A Roundtable Discussion," *American Statistician,* 52, 93-100.

73. Kloeck, T. and van Dijk, H. (1978), "Bayesian estimates of equation system Parameters : an Application of integration by Monte carlo," *Econometrica,* 46, 1–20.

74. Koop, G. (1992a), "Objective' Bayesian unit root tests," *Journal of Applied Econometrics,* 7, 65–82.

75. Koop, G., (1992b), "Aggregate shocks and macroeconomic fluctuations : a predictive approach," *Journal of Applied Econometrics,* 7, 395–412.

76. Koop, G., (1992c), "Intertemporal properties of real output: a Bayesian approach," *Journal of Business and Economic Statistics,* 9, 253–265.

77. Koop, G. and Poirier, D. (1993), "A Bayesian analysis of Logit models using natual conjugate priors," *Journal of Econometrics,* 56, 323–340.

78. Koop, G. and Steel, M. (1994), "A Decision theoretic analysis of the unit root hypothesis using mixtures of elliptical models," *Journal of Business and Economic Statistics,* 12, 95-107.

79. Kozumi, Hideo and Hasegawa, H. (1998), "A Bayesian analysis of structural changes with an application to displacement effect," paper presented at Japan Statistical Association annual meeting.

80. Kuo, L. and Mallick, B.K. (1997), "Bayesian semiparametric inference for the accelerated failure time model," *Canadian Journal of Statistics,*

81. Leamer, E. (1985), "Sensitivity analysis would help," *American Economic Review,* 75, 308–313.

82. Leamer, E., (1983), "Let's take the con out of Econometrics," *American Economic Review,* 73, 31–43.

83. Leamer, E., (1978), *Specification Searches : Ad-hoc inference with non-experimental data,* Wiley, New York.

84. Lee, Peter (1997), *Bayesian statistics : An introduction,* 2nd. Ed. Arnold.

85. Marriott, J., N. Ravishanker, and A.E. Gelfand (1995), " Bayesian analysis of ARMA processes: Complete sampling-based inferences under full likelihood," in D. Berry, K. Chaloner, and J. Geweke eds. *Bayesian Statistics and Econometrics: Essays in Honor of Arnold Zellner*

86. Marsh, T.A. and R.C. Merton (1986) "Dividend variability and variance bounds tests for the rationality of stock market prices," *American Economic Review,* 76, 483–494.

87. McAleer, M., Pagan, A., and Volker, P. (1985), "What will take the con out of Econometrics?", *American Economic Review,* 75, 293–307.

88. McCulloch, R. and Rossi, P., (1992), "Bayes factors for nonlinear hypothesis and likelihood distributions," *Biometrika,* 79, 663–676.

89. McCulloch, R. and Rossi, P. (1994), "An Exact likelihood analysis of the multinomial probit models," *Journal of Econometrics,* 64, 207–240.

90. McCulloch, R. and Tsay, R., (1994a), "Statistical analysis of economic time series via Markov switching models," *Journal of Time Series Analysis,* 15, 235–250.

91. McCulloch, R. and Tsay, R., (1994b), "Bayesian inference of trend- and difference- stationarity," *Econometric Theory,* 10, 596–608.

92. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E (1953), "Equations of state calculations by fast computing machines," *Journal of Chemical Physics,* 21, 1087–1092.

93. Mukhopadhyay, S. and Gelfand, A.E. (1997), "Dirichlet process mixed generalized linear model," *Journal of the American Statistical Association,* 92, 633–639.

94. Müller, P. (1991), "A Generic approach to posterior integration and Gibbs sampling," Technical Report 91-09, Department of Statistics, Purdue University.

95. Naylor, J. and Shaw, J. (1991), *Bayes Four User Guide,* Nottingham Polytechnic, manuscript.

96. Naylor, J. and Smith, A., (1982), "Applications of method for the efficient consumption of posterior distributions," *Applied Statistics,* 31, 214–225.

97. Naylor, J. and Smith, A. (1988), "Econometric illustrations of novel numerical integration strategies for Bayesian inference," *Journal of Econometrics,* 38, 103–126.

98. Nelson, C. and Plosser, C. (1982), "Trends and random walks in macroeconomics time series," *Journal of Monetary Economics,* 10, 139–162.

99. Phillips, P. (1991), "To criticize the critics: an objective Bayesian analysis of stochastic trend (with discussion)," *Journal of Applied Econometrics,* 6, 333–364.

100. Phillips, P. and Ploberger, W. (1991), "Time series modelling, with a Bayesian frame of reference: I. Concepts and Illustrations, Cowles Foundation Discussion Paper No.980.

101. Phillips, P. and Perron, P. (1988), "Testing for unit roots in time series regression," *Biometrika,* 71, 599–607.

102. Poirier, D. (1988), "Frequentist and subjectivist perspectives on the problems of model building in Econometrics, (with Discussion)", *Jounal of Economic Perspectives,* 2, 121–170.

103. Poirier, D. ed. (1991), "Bayesian empirical studies in economics and finance," *Journal of Econometrics,*

104. Poirier, D. (1995), *Intermediate Statistics and Econometrics : A Comparative Approach,* MIT Press.

105. Raftery, A. and Lewis, S. (1992), "How many iterations in the Gibbs sampler?" *Bayesian Statistics 4,* J.M. Bernardo, J.M. Berger, A.P.Dawid, and A.F.M. Smith (eds), Oxford University Press, Oxford, 115–130.

106. Richard, J. and Tompa, H. (1980), "On the evaluation of poly-t density functions," *Journal of Econometrics,* 12, 335–352.

107. Ripley, B. (1987), *Stochastic Simulation,* New York, Wiley.

108. Rubin, D,B., (1987), 'Comment on "The calculation of posterior distributions by data augumentation" by M.A. Tanner and W.H. Wong, *Journal of the American Statistical Association,* 82, 543–546.

109. Schmidt, P. and P.C.B. Phillips (1989) "Testing for a unit root in the presence of deterministic trends," Cowles Foundation Discussion Paper No.933, New Haven, Ct.

110. Schmidt, P. and P.C.B. Phillips (1992) "LM tests for a unit root in the presence of deterministic trends," *Oxford Bulletin of Economics and Statistics,* 54, 257–87.

111. Schmidt, P. and Sickles, R. (1977), "Some further evidence on the use of the Chow test under heteroscedasticity," *Econometrica,* 45, 1293–1298.

112. Schotman, P. and van Dijk, H. (1991a), "A Bayesian analysis of the unit root in real exchange rates," *Jounal of Econometrics,* 49, 195–238.

113. Schotman, P. and van Dijk, H., (1991b), "On Bayesian routes to unit roots," *Journal of Applied Econometrics,* 6, 387–402.

114. Schwert, G. (1989), "Tests for unit roots: a Monte carlo Investigation," *Journal of Business and Economic Statistics,* 7, 147–160.

115. Shiller (1981) "Do stock prices ove too much to be justified by subsequent changes in dividends?" *American Economic Review,* 71, 421–436.

116. Sims, C. (1988), "Bayesian skepticism on unit root econometrics," *Journal of Economic Dynamics and Control,* 12, 463–474.

117. Sims, C. and Uhlig, H. (1991), "Understanding unit rooters: a helicopter tour," *Econometrica,* 59, 1591–1599.

118. Smith, A.F.M. and Gelfand, A.E. (1992), "Bayesian statistics without tears," *American Statistician,* 46, 84–88.

119. Smith, A., Skene, A., Shaw, J., Naylor, J. and Dransfield, M. (1985), "The implementation of the Bayesian paradigm," *Communications in Statistics,* 14, 1079–1102.

120. Stephens, D.A. (1994), "Bayesian retrospective multiple-change point identification," *Applied Statistics,* 43, 159–178.

121. Tanner, M. (1991), *Tools for Statistical Inference,* Springer-Verlag, Lecture Notes in Statistics Series, edited by J. Berger, S. Fienberg, J. Gani, K. Krickberg, I. Olkin and B. Singer.

122. Tanner, M. and Wong, W. (1987), "The Calculation of posterior distributions by data augmentations, (with Discussion)," *Journal of the American Statistical Association,* 82, 528–550.

123. Tierney, L. (1991), "Exploring posterior distributions using Markov chains, " in *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface,* E.M. Keramides and S.M. Kaufman (eds.), Interface Foundation of North America.

124. Tierney, L. (1994), "Markov chains for ecploring posterior distributions," *Annals of Statistics,* 22, 1701–1728, with discussion.

125. Tierney, L. and Kadane, J. (1986), "Accurate approximations for posterior moments and marginal distributions," *Journal of the American Statistical Association,* 81, 82–86.

126. Tierney, L., Kass, R.E., and Kadane, J. (1987), "Interactive Bayesian analysis using accurate asymptotic approximations," In *Computing Science and Statistics: 19th Symposium on Interface,* R.M. Heiderberger, 15–21.

127. Tierney, L., Kass, R.E., and Kadane, J. (1989), "Fully exponential Laplace approximations to expectations and variances of nonpositive functions," *Journal of the American Statistical Association,* 84, 710–716.

128. Thisted, R.A. (1988), *Elements of Statistical Computing,* New York, Chapman & Hall.

129. Toyoda, T. (1974), "The use of the Chow test under heteroscedasticity," *Econometrica,* 42, 601–608.

130. Tsurumi, H. and Wago, H. (1996), "A Bayesian analysis of unit root and cointegration with an application to Yen-Dollar exchange rate model," in *Advances in Econometrics*, eds by T.B.Fomby and R.C. Hill, Vol.11 B, 51–86, JAI Press.

131. Yao, G.C.G. (1984), "Estimation of a noisy didcrete-time step function: Bayes and empirical Bayes approaches," *Annals of Stastistics,* 12, 1434–1447.

132. West, M., Müller, P., and Escobar, M.D. (1994), "Hierarchical priors and mixture models, with application in regression and density estimation," In *Aspects of Uncertainty : A Tribute to D. V. Lindley,* ed. by Smith, A.F.M. and Freeman, P., 363–386, New York: Wiley.

133. White, K., Haun, S. and Horsman, N. (1987), "SHAZAM: *Users Reference Manual,* Version 7."

134. Zangari, P.J. and H. Tsurumi (1996), "A Bayesian analysis of censored autocorrelated data on exports of Japanese passenger cars to United States," in *Advances in Econometrics*, eds. by T.B.Fomby and R.C. Hill, Vol.11 A, 111–143, JAI Press.

135. Zeger, S. and Karim, M. (1991), "Generalized liner models with random effects: a Gibbs sampling approach," *Journal of the American Statistical Association,* 86, 79–86.

136. Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics,* John Wiley and Sons, Inc., New York.

137. Zellner, A., (1983), "Applications of Bayesian analysis in econometrics," *The Statistician,* 32, 23–34.

138. Zellner, A., (1985), "Bayesian Econometrics," *Econometrica,* 53, 253–269.

139. Zellner, A., (1986), "On assessing prior distributions and Bayesian regression analysis with g-prior distributions," in *Bayesian Inference and Decision techniques,* edited by P. Goel and A. Zellner, Elsevier Science, London.

140. Zellner, A. and Rossi, P. (1984), "Bayesian analysis of dichotomous quantal response models," *Journal of Econometrics,* 25, 365–393.

141. Zivot, E., (1992) "Bayesian analysis of the unit root hypothesis within an unobserved component model," *mimeograph*

142. Zivot, E. and P.C.B. Phillips, (1991) "A Bayesian analysis of trend determination in economic time series," Cowles Foundation Discussion Paper, No.1002, New Haven, CT.

143. 伊庭幸人 (1996),「マルコフ連鎖モンテカルロ法とその統計学への応用」統計数理、44, 49–84.

144. 大森裕浩 (1996),「マルコフ連鎖モンテカルロ法」千葉大学経済研究, 10, 237–287.

145. 繁桝算男 (1995),「意志決定の認知統計学」朝倉書店、付録 137–148.

146. 下平英寿 (1997),「ベイズ的方法における MCMC の利用」応用統計学会チュートリアル 講演予稿集、1-10.

147. 伴 金美 (1991)、「マクロ計量モデル分析」有斐閣

Part II

ベイズ計量経済分析の応用

第3章　単位根のベイズ分析

3.1　はじめに

回帰モデルで定数項とトレンド項がある場合における単位根のベイズ分析は, 階差定常過程 (DSP) 対トレンド定常過程 (TSP) を検定するのに適用することができる. しかし, 単位根の下で事後pdfの退化を避けるために事前確率密度 (pdf) を十分に注意して選択する必要がある. Phillips[1991a] やSchotman and van Dijk[1991a,b] で示されたように, 単位根の下での事後pdfの退化はJeffreys型の事前分布を使用することによって避けることができる. この退化の問題に加えて, 単位根の下で定数項が識別されない問題があり, これがフラットな事前分布を用いた場合に事後オッズ比を発散させることになる. Schotman and van Dijk[1991a,b] は, 事後オッズ比の発散を避けるために, 対立仮説の下で定数項にproper事前分布を使用することを示唆した. 他方, Zivot [1992] はこの退化と非識別性の問題を回避するのに観測されない要素モデルを用いている.

本論文では, 誘導型方程式に基づいた単位根に対する事後オッズ比を導き, 次に単位根仮説をモデル選択問題として考えて, 予測の平均平方誤差の予測密度を導く. 単位根問題のモデル選択アプローチはPhillips and Ploberger [1992] で与えられており, 予測密度に関連したベイズ的な方法を開発している. 本論では, DSP対TSPを検定する事後オッズ比を導く. さらに, 単位根とトレンドパラメータに対する事後pdfを導く. 退化を避けるために, これらのパラメータにproperな事前分布を用い, 事前pdfのパラメータを計算するのにブートストラップ法を用いる.

本論の構成は次の通りである. 第2節で, 今までのベイズ単位根検定について簡単に振り返り, 第3節では, 線形トレンド項のあるAR(1)モデルにおける退化と非識別性の問題を回顧した後で事後オッズを示し, 異なる事後オッズ比と比

較するための標本実験を行う. そして, 平方予測誤差の予測密度を導く. 第4節では, 誤差項が未知の次数のARMA(p,q)プロセスにしたがう場合に, 単位根をテストするための平方予測誤差の予測密度と事後pdfを導く. そして, このベイズ検定の方法を円の外国為替レートの月次データに応用し, 結語を第5節で述べる.

3.2 ベイズ単位根モデルのレビュー

Sims [1988] と Sims and Uhlig [1991] による議論の発端となった論文は, 単位根仮説を調べるのに古典的な計量経済学的な方法を用いることを批判し, フラットな事前分布を用いることを支持しており, 最近の多くの実証分析では, フラットな事前分布を用いたベイズ法を採用していることを示した. これらの論文は, ベイズ的な方法による単位根計量経済分析に大きな関心を呼び起こした. 最近の実証分析を見ると, フラットな事前分布を用いるベイズ法では, 経済時系列の単位根仮説に関する推測が古典的な検定による結果とはしばしば異なることを示している. 特に, DeJong and Whiteman [1989a,b,c] によるフラットな事前分布を用いたベイズ分析では, Nelson and Plosser [1982] によって研究されたマクロ時系列や, Shiller [1981] と Marsh and Merton [1986] で分析されたストック価格や配当系列に, 単位根はほとんど見い出されていない. 同様に, Schotman and van Dijk [1991] によるフラットな事前分布を用いたベイズ分析では, 実質為替レートや実質GNPデータに単位根があるという事実は, 古典的な単位根検定の結果によって支持されよりもずっと弱いことを見い出した.

しかし, Peter Phillips は時系列モデルにフラット事前分布を用いるのに強く反対した. それは, フラット事前分布はある種の不変性を欠いており, 自己回帰係数がコレログラムの形状に影響を及ぼし, データやその標本積率に含まれる期待情報量に影響を与えるという経路で, 一般的なモデルに基づいた重要な情報を無視しているからである. 彼は, フラット事前分布を用いるとデータの確率構造の推定の精度に誤解を与えることを示し, 単位根モデルが実際真であるときに, そのモデルの事後分布の推定をかなり偏らせることになることを示した. フラット事前分布に対する対立仮説として, Phillips は情報行列に基づいた (Jeffreys 型) 事前分布の利用を提唱した. この事前分布は, 確率的な非定常性に適応し, 係数

が標本積率とデータに含まれる情報に影響を与える方法を記述した時系列モデルで利用できる事前情報を利用している．これらのモデルに基づいた事前分布を用いて，ベイズ法による事後解析は，単位根の有無について古典的な有意性検定と同じくらい不確実であることを示した．加えて，Phillipsは Jeffreys 型事前分布を利用した方法を Nelson-Plosser データに応用し，もっと一般的な Jeffreys 型の事前分布を採用し，同様な結果を得ている DeJong and Whiteman [1989a] や Zivot and Phillips [1991] で報告されたよりは，単位根モデルに対して決定的な証拠がないことを見い出している．

DeJong and Whiteman [1989a,b], Phillips [1991a] それに Zivot and Phillips [1991] によって行われた単位根仮説のベイズ分析では，モデルのパラメータに関して線形である推定された定数項とトレンドをもつ共通の AR(k) モデルを用いている．Schotman and van Dijk [1991a] は Phillips への返答で，定数項とトレンドのある AR(k) モデルの非線形の観測されない要素モデルを用いて，モデルのパラメータに関してフラット事前分布と Jeffreys 型事前分布の動きと，主要な自己回帰係数 ρ の結果として得られた事後分布を調べた．彼らは，Schmidt and Phillips [1989,1992] におけるように，この要素モデルによる表現は単位根仮説の分析に対してより適切であるとした．それは要素モデルにおける確定項パラメータは，単位根帰無仮説とトレンド定常対立仮説の下では同じ解釈ができるが，線形モデルでのパラメータでは同じ解釈ができないからである．しかし，彼らの見方は DeJong and Whiteman や Phillips によるものとはかなり異なる．特に，要素モデルの中ではフラット事前分布と比較した場合，Jeffreys 型の事前分布は単位根モデルのウェイトを小さくすることを示した．そして，DeJong and Whiteman [1989a,b] の結果と対照的に，フラット事前分布から導かれた事後分布の推定は確率 1 で単位根仮説になることも示した．さらに，単位根モデルを受け入れるという事後オッズ比は，レベルパラメータに improper prior を用いると発散することを示した．レベルパラメータに対する Jeffreys 型の事前分布は improper であるから，Schotman and van Dijk は Jeffreys 型の事前分布は「単位根仮説のベイズ分析には不適切である」と批判した．

Zivot [1992] は，Schotman and van Dijk の発見は単位根仮説の下での観測されない要素 (UC) モデルの誘導型モデルにおける退化の結果であるということを示

した. 彼らの言っていることは, Schotman and van Dijk に対する Phillips[1992] の返答と密接に関連している. もっと重要なことは, Schotman and van Dijk のパラドックス的な結果は, もし要素モデルの構造をそのまま保持し, 初期値の役割に十分な注意すると消えてしまうと言うことを Zivot は示した. 初期値を別な取り扱い方にすると, 構造型の要素モデルは単位根モデルの下で退化する必要はないことを示した. この構造モデルを用いると, Jeffreys 型の事前分布は単位根モデルを必ずしも過小評価せず, フラット事前分布から計算された事後分布は十分に形が良く, 単位根仮説の事後オッズ比は, 適切に正規化されたフラット事前分布と Jeffreys 型事前分布を用いて形成することができる. 加えて, 彼らはこの要素モデルの枠組みの中で, フラット事前分布と Jeffreys 型事前分布から計算された主要な自己回帰パラメータの事後分布がしばしば似たものになり, 線形モデルから得られた対応する事後分布よりも, 単位根が真であったときにはそれを否定するような繰り返し標本でずっとバイアスが小さいことを示した. 興味あることに, いくつかの繰り返し標本という状況において, 要素モデルでの ρ のフラットな事後分布は, 線形モデルでの ρ の Jeffreys 型の事後分布よりも, 単位根モデルを否定するだけの偏りが少ないことが見い出される. 要素モデルの事後分布はある場合には初期値に対してセンシティブであり, 真のモデルが定常であるときに単位根モデルを受容するように結果を偏らせることになる. これらの結果はモデルの定式化は, 事前分布と同様に, 単位根仮説に関する推測へ重大な影響を与えることを示唆している.

3.3 線形トレンドをもつ AR(1) モデル

まず退化と非識別性の問題を振り返ろう. 定数項とトレンドがある AR(1) モデルは, 次式で与えられる.

$$\begin{aligned} y_t &= \mu + \alpha t + \epsilon_t \\ \epsilon_t &= \rho \epsilon_{t-1} + u_t, \quad t = 1, \cdots, n \end{aligned} \tag{3.1}$$

ここで $u_t \sim \mathrm{NID}(0, \sigma^2)$ である. 階差定常過程 (DSP) 対トレンド定常過程 (TSP) を検定するのに (3.1) 式を用いる. DSP の帰無仮説

$$\mathrm{H} : \rho = 1 \quad \text{vs.} \quad \alpha = 0$$

は, TSP の対立仮説に対して検定される.

$$\mathrm{K} : |\rho| < 1$$

方程式 (3.1) から, 次の誘導型を導く.

$$\begin{aligned} y_t - \rho\, y_{t-1} &= \mu(1-\rho) + \alpha[t - \rho(t-1)] + u_t \\ &= [\mu(1-\rho) + \alpha\,\rho] + \alpha(1-\rho)\, t + u_t \end{aligned} \tag{3.2}$$

(3.2) から明らかなように, 定数項 μ は $\rho = 1$ のとき識別されない. フラットな事前分布

$$\mathrm{p}_F(\rho, \mu, \alpha, \sigma) \propto \sigma^{-1} \tag{3.3}$$

を用いると, ρ の事後pdf

$$\mathrm{p}(\rho|\mathrm{data}) \propto |\tilde{X}'\tilde{X}|^{-\frac{1}{2}} (\nu\, s^2)^{-\frac{n-2}{2}} \tag{3.4}$$

を得る. ここで $\tilde{X}$ は $n \times 2$ 行列で, その第 t 行は $\tilde{x}_t = \{1, t - \rho(t-1)\}$ で与えられ, $\nu\, s^2 = y'(\rho) M_{\tilde{x}} y(\rho)$, $y(\rho) = (y_1(\rho), \cdots, y_n(\rho))'$, $y_t(\rho) = y_t - \rho\, y_{t-1}$, と $M_{\tilde{x}} = I - \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'$ である. $\rho = 1$ のとき, $\tilde{X}$ は要素が1の $n \times 2$ 行列になるが, $\nu\, s^2$ はこの線形従属に対して不変であり, したがって, (3.4) における ρ の事後pdfは $\rho = 1$ で無限になる.

Phillips [1991] は, Jeffreys 型事前分布を用いると, (3.4) 式の $|\tilde{X}'\tilde{X}|^{-1/2}$ 項が相殺し, ρ の周辺事後pdfは次のようになることを示した.

$$\mathrm{p}_J(\rho|\mathrm{data}) \propto \alpha_0(\rho)^{\frac{1}{2}} (\nu\, s^2)^{-\frac{n}{2}} \tag{3.5}$$

ここで $\alpha_0(\rho) = (1-\rho^2)^{-1}\{n - (1-\rho^{2n})/(1-\rho^2)\}$ である. (3.5) は $y_0 = \mu$ とおくか, ρ の周辺pdfを得るためにラプラス近似を用いることによって得ることができる.

Schotman and van Dijk [1991a,b] は, レベルパラメータ μ は $\rho = 1$ のとき識別されないので, μ に対するフラット事前分布は無限に発散する事後オッズ比になり, したがって, 事後確率が単位根仮説に対しては1になることを意味することを示した. Zivot [1992] は Schotman and van Dijk の事後オッズが発散するという結果は, 誘導型 (2) を使ったためであると述べた. 彼は観測されない要素

の形は, もし初期値の役割に十分注意すればこの問題は生じないことを示した. 観測されない要素の形は, 次のように表される.

$$y_t = \begin{cases} \mu + \alpha + \rho\epsilon_0 + u_1, & t = 1 \\ [\mu(1-\rho) + \alpha\rho] + \alpha(1-\rho)t + u_t, & t = 2, \cdots, n \end{cases} \tag{3.6}$$

(3.6) の観測されない要素の形では, $\rho = 1$ のときのレベルパラメータ μ の非識別性を避ける一つの方法であるにもかかわらず, 潜在的な初期値 ϵ_0 を評価する必要がある. 誘導型 (3.2) では ϵ_0 の代わりに y_0 を用いており, この場合 y_0 は観測可能であるので, ρ に対する事後分布を導くのに誘導型を用いることにする.

誘導型 (3.2) は次のように表すことができる.

$$y_t(\rho) = \gamma + \delta\, t + u_t \tag{3.7}$$

ここで $y_t(\rho) = y_t - \rho\, y_{t-1}$, $\gamma = \mu(1-\rho) + \alpha\,\rho$ そして $\delta = \alpha(1-\rho)$ である. $\rho = 1$ のとき $\delta = 0$ であり, 方程式 (3.7) は redundant な変数の場合になる. redundant な変数は ρ, γ, δ に対してはより大きな事後分散になるが, これらのパラメータのベイズ推定量の一致性は保持する. 方程式 (3.2) ではなく (3.7) を用いると, (3.4) におけるような $\rho = 1$ での ρ の事後pdfの退化と, ρ に関してもっと散漫な推定を行った場合の非退化事後pdfの間のトレードオフに直面する. 以前に述べたように, このトレードオフはproperな事前分布かJeffreys型の事前分布を用いることによって避けることができる. しかし, どのproper事前分布を使用すべきかについて確信が持てない場合は, 方程式(3.7)で ρ にフラットな事前分布を用いて行った方がよい.

(3.7) の誘導型の形を用いて, 次のフラット事前分布を用いる.

$$\mathrm{p}(\rho, \gamma, \delta, \sigma) \propto \sigma^{-1} \tag{3.8}$$

次に γ, δ, σ に関して積分した後で, ρ の事後密度を得る:

$$\mathrm{p}(\rho|\mathrm{data}) \propto (\nu\, s^2)^{-\frac{\nu}{2}} \tag{3.9}$$

ここで $\nu\, s^2 = y(\rho)' M y(\rho)$, $M = I - X(X'X)^{-1}X'$, $X = (x_1', \cdots, x_n')$, $x_t = (1, t)$, そして $\nu = n - 2$ である.

ρ のこの事後pdfを用いて, 単位根仮説を検定する.

$$\mathrm{H}_0 : \rho = 1 \qquad \text{vs.} \qquad \mathrm{H}_1 : 0 \leq \rho < 1$$

ρ の符号に確信が持てなければ, 対立仮説を $|\rho| < 1$ に変更することができる. ベイジアンは仮説を比較する方法として事後オッズ比か, あるいは最高事後密度区間 (HPDI) をしばしば用いる. HPDIは古典的な信頼区間のベイズ版であるといわれ (Box and Tiao [1973] と Berger [1985] を参照), 推測が単に事後分布全体を報告するだけで, HPDIが厳密ではないが便利な事後分布の要約であると述べているベイジアンによって用いられている. 他方, 事後オッズ比はベイジアンに仮説の事後確率を計算することを認め, 決定理論に基づいてベイズ推定を行うことができる. H_0 を受容する事後オッズ比は

$$\mathrm{K}_{01} = \frac{\pi_0}{1-\pi_0} \frac{\mathrm{p}(\rho = 1|\mathrm{data})}{\int_0^1 \mathrm{p}(\rho|\mathrm{data})\mathrm{d}\rho} \tag{3.10}$$

であり, ここで π_0 は H_0 が真である事前確率である. 事後オッズ比が与えられると, 単位根仮説が真である事後確率は

$$\mathrm{P}(\mathrm{H}_0|\cdot) = \frac{K_{01}}{1+K_{01}} \tag{3.11}$$

である.

誘導型 (3.7) を用いると $\rho = 1$ での情報を無視し, t の係数はゼロになる. $\rho = 1$ のとき, (3.2) は次のようになる.

$$y_t - y_{t-1} = \alpha + u_t \tag{3.12}$$

他方, $\rho \neq 1$ のとき

$$y_t - \rho\, y_{t-1} = [\mu(1-\rho) + \alpha\,\rho] + \alpha(1-\rho)t + u_t \tag{3.13}$$

である. (3.12) と (3.13) では, 単位根仮説 H_0 と H_1 はモデル選択問題へ翻訳される. (3.12) と (3.13) はネストしている (入れ子型) ので, パラメータに関する仮説検定は通常の方法でなされる. このように仮説検定の方法はそれ自身の問題があるので, われわれはモデル選択の方法を採用することにする.

Phillips and Ploberger [1992] はこのような方法を採っている. 彼らはベイズ予測密度に関連したベイズ流モデル選択法を開発した. われわれはこのモデル選択問題を別の方法で行う. (3.12) と (3.13) 式のそれぞれに対して, 予測の平均平方誤差の密度を導き, 期待平均平方誤差と比較する.

y_t について m 期の事後標本観測値 $(y_{n+1}, \cdots, y_{n+m})$ があるとする. y_n 所与の下でのその pdf は, (3.12) の下で次のように与えられる.

$$\begin{aligned} &\mathrm{f}(y^{(m)}|\alpha, \sigma, y_n) = (2\pi)^{-\frac{m}{2}}|G|^{-\frac{1}{2}}\sigma^{-m} \\ &\exp\left[-\frac{1}{2\sigma^2}(y^{(m)} - \alpha\tau^{(m)})'G^{-1}(y^m - \alpha\tau^{(m)})\right] \end{aligned} \tag{3.14}$$

ここで

$$y^{(m)} = \begin{bmatrix} y_{n+1} - y_n \\ \vdots \\ y_{n+m} - y_n \end{bmatrix}, \quad \tau^{(m)} = \begin{bmatrix} 1 \\ \vdots \\ m \end{bmatrix}, \quad G = BB', \quad B = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & & 0 \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

(3.12) の下での予測密度は

$$\mathrm{p}(y^{(m)}|y_n) = \int_0^\infty \int_{-\infty}^\infty \mathrm{f}(y^{(m)}|\sigma^2, \alpha, y_n)\mathrm{p}(\alpha, \sigma|\mathrm{data})\mathrm{d}\alpha\,\mathrm{d}\sigma \tag{3.15}$$

であり, ここで $\mathrm{p}(\alpha, \sigma|\mathrm{data})$ は次式で与えられ

$$\mathrm{p}(\alpha, \sigma|\mathrm{data}) \propto \sigma^{-(n+1)}\exp\left[-\frac{n}{2\sigma^2}(\alpha - \widehat{\alpha})^2\right], \tag{3.16}$$

そして $\widehat{\alpha} = \frac{1}{n}\sum(y_t - y_{t-1})$ である. α と σ に関して積分すると, 次を得る.

$$\mathrm{p}(y^{(m)}|y_n) = \text{const.}\ \left[\nu s^2 + (y^{(m)} - \tau^{(m)}\widehat{\alpha})'H(y^{(m)} - \tau^{(m)}\widehat{\alpha})\right]^{-\frac{n+m-1}{2}} \tag{3.17}$$

ここで $\nu s^2 = \sum(\Delta y_t - \widehat{\alpha})^2$, $\Delta y_t = y_t - y_{t-1}$, $\nu = n - 1$ そして $H = [G + \frac{1}{n}\tau^{(m)}\tau^{(m)\prime}]^{-1}$ である. $w = y^{(m)} - \tau^{(m)}\widehat{\alpha}$ を予測誤差のベクトルとしよう. すると $z = \frac{1}{m}w'w$ は, 予測の平均平方誤差になる. z の予測密度は Tsurumi and Wago [1991] にある. $m = 1$ に対して, これは

$$\mathrm{p}(z|y_n, \widehat{\gamma}, s^2) \propto z^{-\frac{1}{2}}(\nu s^2 + z/\mu_1)^{-(\nu+1)/2} \tag{3.18}$$

となり, ここで $\mu_1 = H^{-1}$ である.

(3.13) 式の下で, 予測密度は次のようになる.

$$\mathrm{p}(y^{(m)}|y_n) = \int_0^1 \int_0^\infty \int \mathrm{f}(y^{(m)}|\sigma,\beta,\rho,y_n)\mathrm{p}(\beta,\sigma,\rho|\mathrm{data})\mathrm{d}\beta\,\mathrm{d}\sigma\,\mathrm{d}\rho \tag{3.19}$$

ここで $\beta=(\mu,\alpha)'$, そして

$$\begin{aligned}\mathrm{p}(\beta,\sigma,\rho|\mathrm{data}) &= (2\pi)^{-\frac{n}{2}}\sigma^{-(n+1)} \\ &\quad \exp\left\{\frac{1}{2\sigma^2}\left[y(\rho)-X(\rho)\beta\right]'\left[y(\rho)-X(\rho)\beta\right]\right\} \qquad (3.20)\\ \mathrm{f}(y^{(m)}|\sigma,\beta,\rho,y_n) &= (2\pi)^{-\frac{m}{2}}|G|^{-\frac{1}{2}}\sigma^{-m} \\ &\quad \exp\left\{-\frac{1}{2\sigma^2}\left[y^{(m)}-X^{(m)}\beta\right]'G^{-1}\left[y^{(m)}-X^{(m)}\beta\right]\right\} \qquad (3.21)\end{aligned}$$

$$y(\rho)=\begin{bmatrix} y_1-\rho\,y_0 \\ \vdots \\ y_n-\rho\,y_{n-1}\end{bmatrix},\quad X(\rho)=\begin{bmatrix} 1-\rho & 1 \\ \vdots & \vdots \\ 1-\rho & n-\rho(n-1)\end{bmatrix},\quad y^{(m)}=\begin{bmatrix} y_{n+1}-\rho\,y_n \\ \vdots \\ y_{n+m}-\rho^m y_n\end{bmatrix}$$

$$X^{(m)}=\begin{bmatrix} 1-\rho & (n+1)-\rho\,n \\ \vdots & \vdots \\ 1-\rho^m & (n+m)-\rho^m n\end{bmatrix},\quad G=BB',\quad B=\begin{bmatrix} 1 & 0 & \cdots & 0 \\ \rho & 1 & & \vdots \\ \vdots & & \ddots & \\ \rho^{m-1} & \cdots & \rho & 1\end{bmatrix}.$$

である. β と σ に関して積分すると, 予測の平均平方誤差の密度 z は, $m=1$ に対して次のように与えられる.

$$\mathrm{p}(z|y_n) \propto \int_0^1 z^{-\frac{1}{2}}(\nu\,s^2+z/\mu_1)^{-\frac{\nu+1}{2}}\mathrm{p}(\rho|\mathrm{data})\,\mathrm{d}\rho \tag{3.22}$$

ここで

$$\begin{aligned}\mathrm{p}(\rho|\mathrm{data}) &\propto |X(\rho)'X(\rho)|^{-\frac{1}{2}}(\nu\,s^2)^{-\frac{n-2}{2}} \\ \nu\,s^2 &= y(\rho)'[I-X(\rho)(X(\rho)'X(\rho))^{-1}X(\rho)']y(\rho), \\ \mu_1 &= H^{-1}=[G+X^{(m)}(X(\rho)'X(\rho))^{-1}X^{(m)'}].\end{aligned}$$

である. z のより小さな事後平均になるようなモデルを選択することによって, 仮説 H_0 と H_1 の検定に予測の平方誤差のベイズ予測密度を用いる.

事後オッズ比 K_{01} と Schotman and van Dijk (SVD と表す), Zivot (UC と表す), そして (3.11) 式 (DTREND と表す) によって得られた単位根の事後確率のシミュレーション結果を示そう. さらに, 平方した予測誤差の平均の差 $\mathrm{E}(z|y_n,\mathrm{H}_0)-$

表 3.1: 期待事後オッズ比, 平方予測誤差の平均の事後確率と差

	$n=50$		$n=10$		$n=250$	
	$\rho=1$	$\rho=.85$	$\rho=1$	$\rho=.85$	$\rho=1$	$\rho=.85$
DTREND						
Mean of K_{01}	.95	.24	2.76	.11	4.91	.001
Mean of $P(H_0\|\cdot)$	.34	.15	.44	.08	.62	.001
SVD						
Mean of K_{01}	1.25	.54	2.98	.34	6.04	.007
Mean of $P(H_0\|\cdot)$	.41	.27	.53	.17	.69	.014
UC($\epsilon_0=0$)						
Mean of K_{01}	4.44	1.78	9.34	1.40	18.39	0.24
Mean of $P(H_0\|\cdot)$	.58	.40	.71	.38	.83	.12
$E(z\|H_0)-E(z\|H_1)$	−.001	.0013	−.003	.0021	−.002	.0028

注: (1) データ発生プロセス:

$$y_t = 6.0 - .11t + \epsilon_t,$$
$$\epsilon_t = \rho\epsilon_{t-1} + u_t \quad u_t \sim N(0,1), \quad \epsilon_0 \sim N(0,1)$$

(2) DTREND: K_{01} は (3.11) の事後pdfを用いて計算
(3) SVD: K_{01} は Schotman and van Dijk [1991] の (3.9) 式から計算
(4) UC: $K_{01} = K_{FJ,01}$ は Zivot [1992] で $\epsilon_0 \sim N(0,1)$.
(5) $E(z|H_i)$ は H_i, $i=0,1$ の下での平方予測誤差の事後平均
(6) 反復回数 = 1000.

$E(z|y_n, H_1)$ に対するシミュレーション結果も示す. 表3.1には K_{01} の実証的な期待値, $P(H_0|\cdot)$ と 50, 100, と 250 の標本サイズに対してと, $\rho=1$ と $\rho=0.85$ に対しての平方予測誤差の平均の差が報告されている. データ発生プロセスは表3.1の注にある. $\rho=1$ のとき, 事後オッズ比の期待値のサイズの順序は, 小さいものから順にDTREND, SVD, そしてUCである. UCモデルに対しては, $\epsilon_0 \sim N(0,1)$ とした. $\rho=.85$ に対しては, すべてのオッズ比は $\rho=1$ に対する対応するそれよりも小さい. サンプルサイズが増加するにつれて, すべての K_{01} と $P(H_0|\cdot)$ は, $\rho=1$ のときに大きく $\rho=0.85$ のときに小さいという望ましい性質を示している. UCモデルに対する表3.1の結果はZivot [1992]のそれと, データ発生プロセスと反復回数が異なるけれども似ている. 平方予測誤差の平均の差の実証的期待値は, $\rho=1$ のとき負で, $\rho=0.85$ のとき正である.

検定の検出力は標本理論での概念であるが, 事後オッズ比を単位根のいくつかの標本検定量と比較してみるのは興味がある. 表3.2は3つの事後オッズ比(DTREND, SVD, UC ($\epsilon_0 = 0$))の経験的検出力と3つの標本検定量(Dickey-Fullerの $\hat{\tau}_\tau$ と $\hat{\rho}_\tau$ 検定とDurbin-Watson検定)のそれを示している. 単位根に対するDurbin-Watson検定は, Sargan and Bhargava [1983] と Bargava [1986] で議論されている. Imhofの方法[1961]を用いると, 単位根仮説に対するDurbin-Watson検定の厳密なサイズとパワーを計算できる. 表3.2の結果は, 事後オッズ比のパワーと標本検定のパワーは, $n = 50$ に対しては平方予測誤差の平均の差のパワーは他の検定のパワーよりも大きくなる傾向になるにもかかわらず, 多かれ少なかれ比較可能である. しかしながら, $n = 100$ と250に対しては, Durbin-Watson(DW)検定は他の検定を圧倒する傾向にある. Durbin and Watson [1951] は, DW検定が $\rho = 0$ の帰無仮説の検定で局所最良不変(LBI)であることを示した. Sargan and Bhargava [1983], Bhargava [1986], そして Tsurumi [1993] で述べられているように, DW検定は単位根 ($\rho = 1$) の帰無仮説に対してはLBIではない. 単位根に対するLBIは検定は, 線形タイムトレンドがある回帰モデルには使うことができない. 表3.2の標本実験の結果は, DW検定は大標本では良好なパワーをもっていることを示している.

3.4 ARMA(p,q)誤差の次数が未知の場合の線形トレンドモデルにおける単位根

事後オッズ比と平方予測誤差の平均の差を, タイムトレンドモデルで誤差項にさらに系列相関がある場合に拡張しよう. モデルは次のようになる.

$$y_t = \mu + \alpha t + \epsilon_t \tag{3.23}$$

$$\epsilon_t = \rho \epsilon_{t-1} + \frac{\Theta(B)}{\Phi(B)} u_t \tag{3.24}$$

ここで $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ と $\Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ である. $\Theta(B)$ は反転可能で $\Phi(B)$ は定常であると仮定されている.

自己回帰誤差のある回帰モデルで単位根を検定するのに, 一般に2つのモデル化の方法がある. DeJong and Whiteman [1991a,b] と Said [1991], その他は, 最大根モデル化にしたがう. $\Pi(B) = \Phi(B)/\Theta(B) = 1 + \pi_1 B + \pi_2 B^2 + \cdots$ とおく

表 3.2: 事後オッズ比と単位根の標本検定の経験パワーの比較
$n = 50$, $n = 100$, $n = 250$ の場合

ρ	1.0	.95	.90	.85	.80
	$n = 50$				
DTREND	.05	.068	.10	.16	.27
SVD	.05	.055	.06	.09	.17
UC($\epsilon_0 = 0$)	.05	.065	.09	.14	.21
$\widehat{\tau}_\tau$	.05	.065	.08	.12	.20
$\widehat{\rho}_\tau$	.05	.055	.09	.17	.27
DW	.05	.066	.10	.17	.26
E(z|H$_0$)−E(z|H$_1$)	.05	.070	.14	.21	.29
	$n = 100$				
DTREND	.05	.13	.25	.45	.75
SVD	.05	.10	.17	.38	.61
UC($\epsilon_0 = 0$)	.05	.11	.19	.40	.69
$\widehat{\tau}_\tau$	.05	.13	.24	.42	.73
$\widehat{\rho}_\tau$	.05	.13	.25	.46	.76
DW	.05	.10	.26	.46	.77
E(z|H$_0$)−E(z|H$_1$)	.05	.060	.17	.35	.68
	$n = 250$				
DTREND	.05	.27	.75	.99	1.00
SVD	.05	.26	.74	.99	1.00
UC($\epsilon_0 = 0$)	.05	.27	.75	.99	1.00
$\widehat{\tau}_\tau$	.05	.17	.60	1.00	
$\widehat{\rho}_\tau$	.05	.25	.74	1.00	
DW	.05	.37	.92	1.00	
E(z|H$_0$)−E(z|H$_1$)	.05	.21	.73	.99	1.00

注: (1) データ発生プロセスは表3.1と表3.2は同じ.
(2) DTREND, SVD, UC($\epsilon_0 = 0$), と経験パワーの計算に用いられる検定の5%サイズに対応する E(z|H$_0$)−E(z|H$_1$) の臨界値
(3) $\widehat{\tau}_\tau$ = Dickey-Fuller t-test; $\widehat{\rho}_\tau$ = Dickey-Fuller ρ-test.
(4) DW = the Durbin-Watson 検定統計量.
(5) DW 検定を除く反復回数 = 1000.
DW 検定のサイズとパワーは Imhof の方法によって計算.

と, 最大根表現は次のように表される.

$$(1-\rho B)(1+\pi_1 B+\pi_2 B^2+\cdots) \tag{3.25}$$

そして (3.23) は

$$e_t=\pi_1 e_{t-1}+\pi_2 e_{t-2}+\cdots+u_t \tag{3.26}$$

となる. ここで

$$\begin{aligned} e_t &= y_t(\rho)-x_t(\rho)\beta \\ y_t(\rho) &= y_t-\rho\, y_{t-1} \\ x_t(\rho) &= x_t-\rho\, x_{t-1}, \qquad x_t=(1,\, t) \end{aligned}$$

である.

もう一つの表現法は, 単位根や共和分の文献でより多く使用されており, (3.25) を

$$a(B)=\eta B+(1-B)(1-\psi_1 B-\psi_2 B^2-\cdots) \tag{3.27}$$

とし, ここで $\eta=a(1)$ である. この表現法では (3.23) 式は

$$\Delta y_t=\mu^*+\eta\, y_{t-1}+\alpha^* t+\sum_{i=1}^{\infty}\psi_i \Delta y_{t-i}+u_t \tag{3.28}$$

となり, ここで μ^* は新たな定数項, そして $\alpha^*=\alpha\eta$ である. DST と TSP 仮説は, この場合次のようになる.

$$\mathrm{H}_0'\colon\ \eta=0 \quad と \quad \alpha^*=0 \qquad \text{vs.} \qquad \mathrm{H}_1'\colon\ \eta\neq 0 \tag{3.29}$$

DeJong and Whiteman [1991b] によるコメントに対する返答で, Phillips [1991b] は (3.28) のモデル化を支持する分かりやすい説明をしている. これに加え, われわれは (25) のモデル化の下で ρ に対する事後 pdf は π_1 の推定値に対して敏感であることが分かった. (3.25) のモデル化の下では, $\mathrm{H}_0\colon\ \rho=1$ を $\mathrm{H}_1\colon\ 0<\rho<1$ のような片側対立仮説に対して検定することができる. π_1 に対して ρ の事後 pdf が敏感になるのを避けるために, $\hat{\pi}_1, \hat{\pi}_2, \cdots,$ が与えられたときの ρ の条件付事後 pdf を用いることができる. ここで $\hat{\pi}_i$ は帰無仮説の下での π_i の推定値である. $\hat{\pi}$ が与えられた下での ρ の事後 pdf はパワーが低くなる傾向にあり, 対

立仮説が真であるときに, 単位根の帰無仮説を受容することを見い出した. したがって, 以後(3.28)のモデル化にしたがうことにする.

(3.28) の誘導型で Δy_{t-i} が k のラグで切断されていると

$$\Delta y_t = \mu^* + \eta\, y_{t-1} + \alpha^* t + \sum_{i=1}^{k} \psi_i \Delta y_{t-i} + u_t \tag{3.30}$$

となる. 散漫な事前分布

$$\mathrm{p}(\mu^*, \eta, \alpha^*, \psi_1, \cdots, \psi_k, \sigma) \propto \sigma^{-1}$$

を用いると, η の事後pdfは, 次式のようになる.

$$\mathrm{p}(\eta|\mathrm{data}) \propto \left[\nu\, s^2 + (\eta - \widehat{\eta})^2 / h^{11}\right]^{-\frac{\nu+1}{2}} \tag{3.31}$$

ここで h^{11} は $(Z'Z)^{-1}$ の逆行列の第 1-1 要素であり, Z の第 t 番目の行は $z_t = (1, y_{t-1}, t, \Delta y_{t-1}, \cdots, \Delta y_{t-k})$ である. $\widehat{\eta}$ は η の事後平均で $\nu = n - k - 3$ である. (3.30) を用いると, 事後オッズ比, K_{01}, 帰無仮説の確率, そして $\mathrm{P}(\mathrm{H}_0|\cdot)$ を計算できる. 同様にして, 次の2つのモデルそれぞれに対して, 平方予測誤差に対する密度を得ることができる. $\mathrm{H}_0' : \eta = 0$ の下で

$$\Delta y_t = \mu^* + \sum_{i=1}^{k} \psi_i \Delta y_{t-i} + u_t \tag{3.32}$$

であり, $\mathrm{H}_1' : \eta \neq 0$ の下で

$$\Delta y_t = \mu^* + \eta\, y_{t-1} + \alpha^* t + \sum_{i=1}^{k} \psi_i \Delta y_{t-i} + u_t \tag{3.33}$$

である.

K_{01}, $\mathrm{P}(\mathrm{H}_0|\cdot)$, そして1期先予測の平方誤差の平均の差はラグの長さ k の選択に敏感である. しかし, そのような敏感さは k を十分大きくすれば消えてしまうことが期待される. 標本実験によって, このことをARMA(3,1)誤差プロセスの下で調べることにする. Schwert [1988] や Said [1991] による研究によって, MAパラメータ θ_1 の値を変化させる.

表3.3には, 帰無仮説の事後確率の経験的な期待値と k と θ_1 の異なる値と, それに $n = 50, 100$ と 250のサンプルサイズに対する平方予測誤差の平均の差が

示されている. ラグの長さが12と18に対して, 事後確率 $P(H_0'|\cdot)$ は θ_1 の所与の値と近い. $\theta_1 = -.8$ に対して, 事後確率は $\theta_1 \in (-.5, .5, .8)$ に対するそれよりもずっと小さい. 表3.1のホワイトノイズの場合と比べてみると, 事後確率はかなり小さい. ラグの長さが12と18に対して, 平方予測誤差の平均の差の経験的な期待値は θ_1 のさまざまな値に対して敏感ではない.

表3.4は, $P(H_0'|\cdot)$, 平方予測誤差の平均の差, とサンプルサイズが $n = 250$ に対する拡張Dickey-Fuller検定(ADF)である. 表3.2のホワイトノイズの場合と比べると, パワーは誤差項がARMAプロセスにしたがうときすべて低い. さらに, 検定のパワーは θ_1 の値に対して敏感である. 一般的に, $P(H_0'|\cdot)$ のパワーはADFのそれよりも大きく, 平方予測誤差の平均の差のパワーは3つのテスト中で最も小さい.

表 3.3: 事後確率と平方予測誤差の平均の差の計算した期待値

θ_1	ϕ_1	ϕ_2	$k=0$		$k=6$		$k=12$		$k=18$	
			$P(H_0'\|\cdot)$	DIF	$P(H_0'\|\cdot)$	DIF	$P(H_0'\|\cdot)$	DIF	$P(H_0'\|\cdot)$	DIF
			$n=50$							
−.8	.5	−.3	.014	.015	.108	.014	.148	−.009	.141	−.010
−.5	.5	−.3	.111	−.018	.258	−.015	.252	−.016	.229	−.014
.5	.5	−.3	.580	−.026	.428	−.014	.424	−.015	.386	−.016
.8	.5	−.3	.591	−.026	.481	−.016	.434	−.015	.413	−.016
			$n=100$							
−.8	.5	−.3	.010	.016	.147	.013	.224	−.011	.233	−.009
−.5	.5	−.3	.132	.016	.147	.013	.358	−.012	.356	−.011
.5	.5	−.3	.667	−.017	.531	−.013	.522	−.013	.518	−.011
.8	.5	−.3	.680	−.017	.584	−.013	.554	−.012	.547	−.010
			$n=250$							
−.8	.5	−.3	.015	.014	.169	−.008	.354	−.010	.398	−.011
−.5	.5	−.3	.172	.005	.540	−.011	.570	−.011	.564	−.012
.5	.5	−.3	.801	−.016	.710	−.010	720	−0.12	.733	−.012
.8	.5	−.3	.812	−.016	.761	−.012	.753	−.012	.758	−.012

注: (1) データ生成プロセスについては表3.1の注を参照, ただし誤差項は次式で生成されている.

$$\epsilon_t = \epsilon_{t-1} + \frac{1+\theta_1 B}{1-\phi_1 B - \phi_2 B^2}$$

(2) $P(H_0'|\cdot) = H_0' : \eta = 0$ が真である事後確率.

(3) $DIF = E(z|H_0') - E(z|H_1')$.

(4) $k = 0$ は誤って $\epsilon_t \sim N(0, \sigma^2)$. と仮定した場合

表 3.4: H_0' と平方予測誤差の平均の差の事後確率のパワーの計算値

$k = 15$, $n = 250$ の場合

θ_1	ϕ_1	ϕ_2		$\rho = 1$	$\rho = .95$	$\rho = .90$	$\rho = .85$	$\rho = .80$
			$P(H_0'\|\cdot)$	.05	.30	.75	.92	.98
−.8	.5	−.3	DIF	.05	.13	.45	.70	.81
			ADF	.05	.28	.71	.86	.92
			$P(H_0'\|\cdot)$	.05	.33	.65	.86	.93
−.5	.5	−.3	DIF	.05	.19	.40	.55	.71
			ADF	.05	.23	.45	.62	.73
			$P(H_0'\|\cdot)$	.05	.33	.68	.85	.93
.5	.5	−.3	DIF	.05	.19	.39	.58	.66
			ADF	.05	.22	.41	.63	.74
			$P(H_0'\|\cdot)$	.05	.34	.68	.85	.89
.8	.5	−.3	DIF	.05	.21	.38	.60	.65
			ADF	.05	.25	.43	.64	.74

注: (1) データ生成プロセスについては表3.3の注を参照.
(2) $P(H_0'|\cdot) = H_0'$: $\eta = 0$ が真である事後確率.
(3) $DIF = E(z|H_0') - E(z|H_1')$.
(4) 検定の 5% サイズに対応する臨界値が経験的なパワーの計算に用いられる.

単位根あるいはDSPを検定するのに, 事後pdfを $\eta = 0$ の方向に引っ張る事前pdfを選択する必要がある. 以前に Jeffreys 型事前分布は適切でないと述べた. DSP対TSPを検定するのに, DeJong and Whiteman (1991a) は AR(3) モデルを用いて, 最大根に対する事後pdfを導き, それが1であるか否か検定した. 彼らはARパラメータにフラット事前分布を用い, AR(3) 過程の根に変換した後のこのフラット事前分布は, 1の近傍にもっと多くのウェイトをおくことによって事後pdfを僅かに単位根の方に引っ張る最大根に対する事前分布を生じる. 彼らの事前pdfは1の近傍で切断されている.

本論文では, 次のような自然共役事前分布を用いることを勧める.

$$p(\mu^*, \eta, \alpha^*, \psi_1, \cdots, \psi_k, \sigma) \propto \sigma^{-1}\, p(\eta, \alpha^*|\sigma) \tag{3.34}$$

ここで

$$p(\eta, \alpha^*|\sigma) \propto \exp\left\{-\frac{1}{2\sigma^2}(\delta - \delta_0)' V_0 (\delta - \delta_0)\right\} \tag{3.35}$$

ここで $\delta' = (\eta, \alpha^*)$, そしてこの事前分布を $\eta = 0$ と $\alpha^* = 0$ の帰無仮説の値の

中心におく. δ の事後pdfを導くために, (3.29)式を次のように書きなおす.

$$w = Z\delta + X\beta + u$$

ここで

$$w = \begin{bmatrix} \Delta y_1 \\ \vdots \\ \Delta y_n \end{bmatrix} \quad Z = \begin{bmatrix} y_{-1} & 1 \\ \vdots & \vdots \\ y_{n-1} & n \end{bmatrix} \quad X = \begin{bmatrix} 1 & \Delta y_0 & \cdots & \Delta y_{1-k} \\ \vdots & \vdots & & \vdots \\ 1 & \Delta y_{n-1} & \cdots & \Delta y_{n-k} \end{bmatrix}$$

そして $\beta' = (\mu^*, \psi_1, \cdots, \psi_k)$ である. δ の事後pdfは

$$\mathrm{p}(\delta|\mathrm{data}) \propto \left[\nu s^2 + d + (\delta - \tilde{\delta})'(V_0 + V)(\delta - \tilde{\delta})\right]^{-(n-k-1)/2} \tag{3.36}$$

であり, ここで $V = Z'M_xZ$, $M_x = I - X(X'X)^{-1}X'$, $\nu s^2 = w'M_xw$,

$$d = (\delta_0 - \widehat{\delta})'V_0(V_0 + V)^{-1}V(\delta_0 - \widehat{\delta}) \qquad \widehat{\delta} = V^{-1}Z'w$$

そして

$$\tilde{\delta} = (V_0 + V)^{-1}(V_0\delta_0 + V\widehat{\delta})$$

である.

方程式(3.36)を用いるためには δ の事前分布におけるパラメータ V_0 と δ_0 を見つける必要がある. ここで, 次のようなパラメトリック・ブートストラップ法を提案する.

散漫事前分布に基づく(3.29)の回帰パラメータの事後平均を $(\widehat{\mu}^*, \widehat{\eta}, \widehat{\alpha}^*, \widehat{\psi}_1, \cdots, \widehat{\psi}_k, \widehat{\sigma})$ と表す. $\widehat{u}_t$ を $\mathrm{N}(0, \widehat{\sigma}^2)$, $t = 1, \cdots, n^*$ から抽出し

$$\widehat{\Delta y}_t = \widehat{\mu}^* + \widehat{\eta}\,\widehat{y}_{t-1} + \widehat{\alpha}^* t + \sum_{i=1}^{k} \widehat{\psi}_i \widehat{\Delta y}_{t-k} + \widehat{u}_t \tag{3.37}$$

によって, 初期値 $(y_0, \Delta y_0, \cdots, \Delta y_{-k})$ として標本平均を用いて $\{y_t\}$ の観測値系列を発生する. 各試行 i に対して, n^* 個の観測値を得て, $\widehat{V}_{0,i} = Z_i'M_{x,i}Z_i$ と $\widehat{\delta}_{0,i} = V_{o,i}^{-1}Z_i'M_{x,i}w_i$ を得る.

$$V_0 = \frac{1}{r}\sum_{i=1}^{r} \widehat{V}_{0,i} \qquad \delta_0 = \frac{1}{r}\sum_{i=1}^{r} \widehat{\delta}_{0,i} \tag{3.38}$$

とおく. ここで添字 i は第 i 番目の試行を表す. 単位根分析におけるブートストラップがMark [1993] と Campbell [1992] で用いられている. 彼らは t 比のブートストラップ分布を発生した.

表 3.5: ADF-t と F 検定の事後オッズ比, 事後平均と標準偏差:
円の月次での外国為替レートの対数, 1973.01–1992.12

			η		α^*			
	O_{HK}	P(H|·)	PM	PSD	PM	PSD	ADF-t	F-test
(3.36) 式の δ_0	26.32	.96	.003	.013	0	.00006	−2.19	2.43
$\delta_0 = 0$	8.10	.89	−.021	.013	−.0001	.00006		

注: (1) ブートストラップ法で $n^* = 100$ と $r = 1000$ とおいた.
(2) (3.36) 式の δ_0 = (3.36) のブートストラップ分布の平均として計算した δ_0 の値
(3) P(H|·) = 帰無仮説の事後確率:
$\mathrm{P(H|\cdot)} = \frac{O_{HK}}{1+O_{HK}}$
(4) PM = 事後平均, と PSD = 事後標準偏差
(5) ADF-t = augmented Dickey-Fuller t 検定 と
F test = Dickey-Fuller [1981] の Φ_3 検定

仮説 (3.29) を検定するために, 事後オッズ比を用いる.

$$O_{HK} = \frac{\pi_0}{1-\pi_0} \cdot \frac{\mathrm{p}(\delta = 0|\mathrm{data})}{\int \mathrm{p}(\delta|\mathrm{data})\mathrm{d}\delta} \tag{3.39}$$

そして η と α^*の事後pdfを調べることによって, 事後オッズ比を拡張する. 事後オッズ比は K を棄却することを示唆するが, η と α^* の事後pdfを調べると, $\eta = 0$ であるが $\alpha^* \neq 0$ であることを示している.

この方法を円の外国為替レートに関する月次データに適用しよう: $\{y_t = \log(\mathrm{exch}_t)\}$, ここで exch_t は t 月の月末における U.S. ドルの円価格である. 1973年1月 (1973.01) から 1992年12月までの月次データ (1992.12) を用いる. データは日本経済新聞社のNEEDSデータから採用した. 事後オッズ比 (POR) を計算し, 帰無仮説の確率 Pr(H|·) は δ_0 の2つの選択：1つは (3.36) 式と, もう1つは δ の帰無仮説の値 $\delta = 0$ を用いて計算する. 結果は表3.5にまとめてある.

表3.5 で報告された結果は, ラグの長さが $k = 6$ である. ラグの長さを $k = 3$ から $k = 12$ へ変えたが, 結果は表3.5と同じであった. 事後オッズ比と帰無仮説の事後確率は, データがDSPの帰無仮説を支持することを示す. η と α^* の事後pdfは, 分布がゼロを中心に集中して分布していることを示し, $\alpha^* \neq 0$ であるチャンスはほとんどないことを示す. ADF-t と F-検定もDSPの仮説を支持し

ている. (3.29) の帰無仮説は η と α^* の同時検定であることを示し, したがって F 検定が ADF-t 検定より優れた検定であることを示している.

3.5 おわりに

本論文では誘導型に基づいた単位根の検定にベイズ事後オッズ比を提案し, さらに予測密度基準も提案した：1期先予測の平方誤差の平均の差である. 事後オッズ比は, Schotman and van Dijk [1991] や Zivot [1992] のそれと似た結果を得た. 次に事後オッズと1期先予測の平方誤差の平均の差を, 次数が未知の場合の ARMA(p,q) 誤差プロセスのモデルに拡張した. 標本実験は十分に長いラグに対しては, これらの基準は ARMA パラメータの値に敏感でない結果を生じる.

ラグの長さを先見的に決めるよりも, ある基準に基づいて選択した方がよいだろう. Campbell and Perron [1991] は k を選択する一つの方法を述べている：ラグの最大長, $k_{\max}$ から出発し, もしそのラグ係数が有意なら $k = k_{\max}$ を選択する. もう一つの方法は AIC, BIC, あるいは PIC のようなモデル選択基準を用いることができる. PIC は Phillips and Ploberger [1992] で与えられた.

DSP 対 TSP の検定に, 単位根とトレンドパラメータに対する事後オッズ比と事後 pdf を導いた. 単位根とトレンドパラメータに対して proper な事前分布を用い, 事前 pdf のパラメータの値をブートストラップ法によって生成した. このベイズ法を円の対ドル為替レートの月次データに適用した. ベイズテストの方法は, 標本理論における検定と似た結果を得た.

参考文献

1. Berger, J.O., (1985) *Statistical decision theory and Bayesian analysis*, 2nd ed. New York: Springer-Verlag.

2. Bhargava, A., (1986) "On the theory of testing for unit roots in observed time series," *Review of Economic Studies*, 52, 369–384.

3. Box, G.E.P. and G.C. Tiao, (1973) *Bayesian inference in statistical analysis*, Reading: Addison-Wesley.

4. Campbell, J.Y., (1992) "Why long horizon? Power against persistent alternative," *mimeigraph.*

5. Campbell, J.Y. and P. Perron, (1991) "Pitfalls and opportunities: what macroeconomists should know about unit roots," *NBER Macroeconomics Annal* 141–201.

6. DeJong, D.N. and C.H. Whiteman (1989a) "Trends and random walks in macroeconomic time series: A reconsiderration based on the likelihood principle," *Journal of Monetary Economics*

7. DeJong, D.N. and C.H. Whiteman (1989b) "Trends and cycles as unobserved components in real GNP: A Bayesian perspective," *Proceedings of the American Statistical Association,* 63–70.

8. DeJong, D.N. and C.H. Whiteman, (1991a) "The temporal stability of dividends and stock prices: evidence from the likelihood function," *American Economic Review* 81, 600–617.

9. DeJong, D.N. and C.H. Whiteman, (1991b) "The case for trend-stationarity is stronger than we thought," *Journal of Applied Econometrics*, 6, 413–421.

10. Durbin, J. and G.S. Watson, (1951) "Testing for serial correlation in least squares regression," *Biometrika*, 58, 1–19.

11. Imhof, P.J., (1961) "Computing the distribution of quadratic forms in normal variables," *Biometrika* 48, 419–426.

12. Mark, Nelson C., (1993) "Exchange rates and fundamentals: evidence on long-horizon predictability," *mimeograph.*

13. Marsh, T.A. and R.C. Merton (1986) "Dividend variability and variance bounds tests for the rationality of stock market prices," *American Economic Review*, 76, 483–494.

14. Nelson, C.R. and C.I. Plosser (1982) "Trends and random walks in macroeconomic time series: some evidence and implications," *Journal of Monetary Economics*, 10, 139–162.

15. Phillips, P.C.B., (1991a) " To criticize the critics: an objective Bayesian analysis of stochastic trends," *Journal of Applied Econometrics*, 6, 333–364.

16. Phillips, P.C.B., (1991b) "Bayesian routes and unit roots: De rebus prioribus semper est disputandum," *Journal of Applied Econometrics*, 6, 413–421.

17. Phillips, P.C.B., (1992) "Bayes methods for trending multiple time series with an empirical application to U.S. economy," Cowles Foundation discussion paper no.1025 (Yale University, New Haven, CT.)

18. Phillips, P.C.B. and W. Ploberger, (1992) "Posterior odds testing for a unit root with data-based model selection," *mimeograph.*

19. Said, E.S., (1991) "Unit root test for time-series data with a linear time trend," *Journal of Econometrics* 47, 285–304.

20. Sargan, J.D. and A. Bhargava, (1983) "Testing residuals from least squares regression for being generated by the Gaussian random walk," *Econometrica*, 51, 153–174.

21. Schmidt, P. and P.C.B. Phillips (1989) "Testing for a unit root in the presence of deterministic trends," Cowles Foundation Discussion Paper No.933, New Haven, Ct.

22. Schmidt, P. and P.C.B. Phillips (1992) "LM tests for a Unit root in the presence of deterministic trends," *Oxford Bulletin of Economics and Statistics,* 54, 257–87.

23. Schotman, P. and H. van Dijk, (1991a) "A Bayesian analysis of the unit root in real exchange rate," *Journal of Econometrics* 45, 195–238.

24. Schotman, P. and H. van Dijk, (1991b) "On Bayesian routes to unit roots," *Journal of Applied Econometrics*, 6, 387–402.

25. Schotman, P. and H. van Dijk, (1991c) "Posterior analysis of possibly integrated time series with an application to real GNP," in E.Parzen et. al eds., *New Directions in Time Series Analysis*, IMA Vol. in Mathematics and it Applications, Springer Verlag.

26. Schwert, G.W., (1988) "Tests for unit roots: A Monte Carlo investigation, " *Journal of Business and Economic Statistics* 7, 147–160.

27. Shiller (1981) "Do stock prices move too much to be justified by subsequent changes in dividends?" *American Economic Review*, 71, 421–436.

28. Sims, C., (1988) "Bayesian skepticism on unit root econometrics," *Journal of Economic Dynamics and Control*, 12, 463–474.

29. Sims, C.A. and H. Uhlig (1991) "Understanding unit rooters: a helicopter tour," *Econometrica*, 59, 1591–1599

30. Tsurumi, H., (1993) "Sampling unit root tests," *mimeograph.*

31. Tsurumi, H. and H. Wago, (1991) "Mean squared errors of forecast for selecting nonnested linear models and comparison with other criteria," *Journal of Econometrics*, 48, 215–240

32. Zivot, E., (1992) "Bayesian analysis of the unit root hypothesis within an unobserved component model," *mimeograph*

33. Zivot, E. and P.C.B. Phillips, (1991) "A Bayesian analysis of trend determination in economic time series," Cowles Foundation Discussion Paper, No.1002, New Haven, CT.

第4章　為替レートモデルにおける時系列分析

4.1　はじめに

名目あるいは実質の為替レートは, ほとんどランダム・ウォークに近い動きをすると言われている. これは, ほとんどが最近の為替レートモデルに関する実証研究から生じたものである. もし, 実質為替レートがランダム・ウォークにしたがうとすれば, 為替レートの動きについての経済理論に重大な影響を与える. この場合には実質為替レートに対するショックは累積し, 実質為替レートの時系列は平均に回帰する傾向を示さない. これは, 一定の均衡為替レートが存在すると主張する**購買力平価** (PPP) の問題とは相いれない考え方である. 従来行われていた実証分析での問題点の多くは, 単位根の検定力の低さと, フロート期の為替レートの観測期間が比較的短いことから生じる結果の不安定性である. さらに, 帰無仮説としてランダム・ウォークを考えることの問題がある. 本論では, ベイズ分析の立場から新しい方法を提案し, 為替レートの分析に適用しようとするものである.

単位根のベイズ分析は, Sims and Uhlig (1991) が単位根の検定に対してベイズ分析と標本理論に基づく分析の対比に焦点を合わせて以来, かなり多くの関心を集めている. 本論では, Tsurumi and Wago(1996), 和合 (1996, 1997) で述べられた方法に基づいて単位根に対するベイズ分析を導き, さらに最高事後密度区間を用いた共和分の方法を導く. 帰無仮説の値の方向に事後分布を引っ張る自然共役事前分布を用いる方法を用いて, 事前パラメータの値をブートストラップ法によって経験ベイズの考え方で求める. これらの方法を円ドル為替レートモデルのデータに適用する.

本書の構成は, 次の通りである. 次の第2節で, 為替レートモデルと購買力平価の実証分析を簡単に振り返り, 第3節では伝統的な方法で行った購買力平価

モデルの分析を示す. 第4節で, 誤差項が未知の次数の ARMA プロセスにしたがう場合のベイズ的な単位根の検定の方法を紹介し, 変数が階差定常プロセスとトレンド定常プロセスのどちらに属するかを中心に考える. さらに第5節でベイズ的な共和分分析の方法を示す. まず散漫な事前分布に基づく最高事後密度区間を導き, これが Johansen の尤度比検定と同等であることを示し, 共和分関係を含む誘導型係数行列のランクを示す特異値の事後分布を導く. 特異値を用いたのは, VAR モデルで方程式の順序に関して不変であるためである. 第6節では, これらの2つの節で説明した方法を用いて, 為替レートモデルのベイズ分析を行い, 最後に第7節で結語を述べる.

4.2 為替レートモデルにおける実証分析

名目あるいは実質の為替レートは, ほとんどランダム・ウォークに近い動きをする. この結論は, ほとんどが最近の為替レートモデルに関する実証研究から生じたものである. Meese and Rogoff (1983a,b, 1988) の一連の論文では, 構造モデル, 制約なしの VAR モデル, そして単変量時系列モデルを含む多くの経済モデルの予測パフォーマンスを比較し, どれもランダム・ウォークモデルを越えるものはなかったと結論している. 同様な結論は, 他の研究にも見られる. Frankel and Meese (1987) や Dornbush and Frankel (1987) は, この問題やその意味を適切にレビューしている. 特に, **実質為替レート**がランダム・ウォークにしたがうことは, 為替レートの動きについての経済理論に重大な影響を与えている. もし実質為替レートがランダム・ウォークにしたがえば, 実質為替レートに対するショックは累積し, 実質為替レートの時系列は平均に回帰する傾向を示さない. これは, 一定の均衡為替レートがあると主張する**購買力平価** (PPP) の考え方とは著しく異なっている. ほとんどのマネタリータイプの為替レートモデル [Dornbush(1976) and Frankel(1979) 参照] では, 一定の PPP 水準は価格と名目為替レートを連続的に調整する繋ぎ役として機能する.

多くの予測実験は, 実質為替レートに対してランダム・ウォークモデルを示唆する実証的な事実だけではない. 公式的な統計的検定を用いれば, 単位根という帰無仮説を定常な自己回帰モデルという対立仮説に対して棄却しない. この結論は別の検定を用いた研究によっても得られている. Meese and Rogoff (1988) は,

フロート期間における米ドルに対するドイツマルク, 日本の円, 英国のポンドの月次データにおける単位根を検定し, 標準的な Dickey-Fuller 検定ともっと頑健な Phillips (1987) 検定を用いて単位根仮説を棄却できなかった. Edison and Fisher (1988) もオランダのギルダーに対するドイツのマルクの場合を除いて, EMS 内のほとんどの通貨に対して同じ結論を得ている. Huizinga (1987) と Kaminsky (1988) は, 実質為替レートに対するランダム・ウォークをもっと一般的な時系列モデルに対して検定するために Cochrane(1987) によって開発された分散比を用いている. これらの著者は平均回帰に対するいくつかの傾向を見い出しているが, 公式にランダム・ウォークという帰無仮説を棄却できない.

これらの著者がすべて認めているように, これらの結果は単位根検定の検出力の低さと, 1973 年から始まるフロート期の為替レートという比較的短い期間のためである. このことは, われわれは高々280 個の月次観測値しかないことを意味している. 真のモデルが AR(1) であるという理想的な状態であっても, 1 次の自己相関のような古典的な統計量は, 単位根の仮説を棄却するためには 0.92 よりも小さくなくてはならない. 実質為替レート系列の観測された自己相関は, ランダム・ウォークという帰無仮説と整合的である. このことは, 必ずしもランダム・ウォークが実質為替レートを記述するのに尤も妥当であることを意味してはいない。1 次の自己相関のパラメータが 0.97 であれば定常であるという対立仮説でも, データを同様にうまく説明することができる.

実質為替レートに対する単位根検定を応用する場合の 2 番目の問題は, これらの検定における帰無仮説が「理論が誤っている」というものである. 実質為替レートに対するランダム・ウォーク・モデルは, 長期の PPP が成り立たないことを意味している. これは 5%水準で経済的に興味ある仮説を検定するということは, 経済理論をそれが正しいのに棄却するのに 5%のチャンスがあることを意味している検定方法とは矛盾する考え方である. ここで正しいとは時系列が定常であることを意味している. 古典的な仮説検定は, 定常な AR(1) でなくランダム・ウォークを帰無仮説として検定する. すなわち, ランダム・ウォーク・モデルは AR(1) モデルのクラスの特別な部分であると考えている. ベイジアンの方法では, 帰無仮説と対立仮説は対称的に扱われる. データを定常あるいは非定常という特定の観点から見る必要はない. データが与えられると, 2 つの中

どちらの方がもっともらしいか決めることができる. 異なるモデル間で選択する場合に用いる公式的なベイズ流の道具は, 事後オッズの計算である [たとえばZellner(1971) を参照]. ランダム・ウォークであることに離散的な確率を与えると, 事後オッズを精密な帰無仮説 (*sharp null hypothesis*) を検定するのに用いることができる.

ベイズ統計学の観点からの最初の結果は, Sims(1988) と DeJong and Whiteman(1991a, b) に見られる. Sims(1988) は定数項のある AR(1) モデルを考えた. Sims の検定方法は単位根の帰無仮説を定常な対立仮説に対する事後オッズ比を計算する. ベイズ流の単位根検定に関するサーベイは, 和合 (1996) を参照のこと.

外国為替レートに関する実証研究は, 3 つのグループに分類される. 最初のグループは外国為替レートの時系列データに分布関数を当てはめるか, ランダム・ウォーク仮説を検定することによって分析する研究からなる. 主要な関心は, いくつかの変数グループと外国為替レート間の関係を調べることなしに時系列自身を分析する. たとえば, Boothe and Glassman (1987) は為替レートの対数変化の日次データに 3 つの非正規確率密度関数を当てはめており, 分布パラメータが時間で変化する証拠があると結論している. Westerfield (1977) と So (1987) も為替レートの時系列データにいくつかの密度関数を当てはめている. Hakkio (1986) は, 外国為替レートのランダム・ウォークの検定がいろいろな結果を与える一つの理由として, ランダム・ウォークの検定は検出力が低いことを議論している. Kariya and Matsue (1989) は, 日次あるいは週次の為替レートデータがランダム・ウォークにしたがうかどうか調べるために検定を行っている. 彼らは Taylor (1986) の非線形条件付不均一分散モデルを用いて, 日次の為替レートは確率トレンドにしたがい, 月曜日のレートはランダム・ウォークにしたがうが, 他の曜日の為替レートデータは自己回帰プロセスにしたがうことを見い出した.

2 番目のグループの研究は, 外国為替レートの回帰モデルである. 経済理論に基づいて一組の説明変数と外国為替レートとの間の関係を定式化し, 推定した. 総合的なサーベイは Levich (1985) や Isard (1988) にある. 外国為替レートに関する多くの回帰モデルは, いくつかの金融理論に基ずく定式化を用いており, その用い方によって価格弾力的なマネタリーモデル (*flexible-price monetary*

models) [Frankel (1979), Bilson (1978,1979)], 価格硬直的なマネタリーモデル (*sticky-price monetary models*) [Dornbusch (1983)], そして価格硬直的な資産モデル (*sticky-price asset models*) [Frankel (1983), Hooper and Morton (1982)] と分類される. Kariya and Fukao (1988) は, 同時方程式体系に基づく合理的期待モデル (*rational expectation model*) を開発した.

3番目のグループの研究は, 最初の2つのグループを組み合わせたものである. すなわち, 回帰の枠組みの中でランダム・ウォークあるいは共和分検定を行うか, 回帰モデルを回帰係数がランダム・ウォークにしたがう状態空間型で定式化している. Enders (1988), Bleaney (1992), と Pippenger (1993) その他は, 固定レートと弾力的変動レート期間のもとで購買力平価を検定し, 回帰誤差プロセスが単位根をもつことを示した. Baillie and Selover (1987) は, 他のマネタリーモデルに加えて Frankel (1979) の弾力的マネタリーモデルと Dornbusch (1983) の価格硬直的マネタリーモデルを推定し, 多くの回帰係数が有意であるが, 符号と大きさは理論から期待されるものとは異なり, 回帰結果は回帰の変数間の共和分関係を支持していないと結論した. 彼らは拡張 Dickey-Fuller (ADF) 検定を用いた.

共和分と誤差修正モデルは, 最近の外国為替レートの研究に多数応用されている. Baillie and Selover (1987) に加えて, その他のいくつかの論文がある. Copeland (1991) は, Johansen の LRT 検定によって為替レートが互いに共和分しているかどうか検定した. そして為替レートはペア毎に共和分していないが, スポット為替レートはそれ自身のラグ付きのフォワードレートと共和分していることを見い出した. Clifarelli (1992) は ADF と Phillips の修正された ADF 検定 (Phillips and Oulialis, 1990) を行い, スポットレートとラグ付きの外国為替レートが共和分されていることを見い出した. McNown and Wallace (1992) は, M_1 貨幣供給が GNP と利子率に共和分し, 他方 M_2 貨幣供給は GNP, 利子率と為替レートに共和分していることを示した. 彼らは Johansen の LRT を用いている. Booth and Chowdhury (1992) も Johansen の LRT を用いてカナダの為替レートが米国とカナダの利子率と共和分していることを示した.

為替レートの状態空間モデルは Wolff (1987) と Fukao (1988) 等々で行われた. Wolff はマネタリー為替レートモデルの予測パフォーマンスを向上させるた

めにカルマン・フィルターを用いた. Fukao, その他はカルマン・フィルターを用いて金利差と国際収支バランスに伴う回帰パラメータの変化を追跡した. そして, 世界の主要金融市場の国際化を反映して外国為替レートにおける金利差の重要性は増すが, 国際収支バランスの影響は減少すると結論している.

誤差項の分散が自己回帰プロセス, あるいは自己回帰移動平均プロセスによってモデル化されている自己回帰条件付異分散モデル (ARCH あるいは GARCH として知られている) が, 主として外国為替レートのボラティリティあるいはリスクプレミアムを分析するのに用いられている. Lastrapes (1989), Diebold and Nerlove (1989), Baillie and Bollerslev (1990), そして Kroner and Lastrapes (1993) は有名な例である. Tsurumi and Chen (1994) はベイズ法を適用して結合点の数が未知のときで誤差項が自己相関し分散不均一であるとき, スウィチィング回帰モデルを推定した. 彼らは円ドル為替レートは 1973 年から 1992 年の間に 4 つの区間があることを見い出している.

4.3 購買力平価モデルにおける伝統的な共和分分析

購買力平価 (PPP) は, 外国為替レートの動きを説明する理論である. 購買力平価は, 2 つの通貨間の為替レートがそれらの価格水準の比に必ず等しくなる必要がある. この場合に, 為替レートは 1 単位の国内通貨が両国の購買力が等しくなるように決定される.

$$e_t - p_t + p_t^* \equiv u_t$$

ここで e_t は名目の為替レート (外国通貨の国内価格) の対数, p_t は国内価格の対数そして p_t^* は外国価格の対数とする. そして u_t は PPP からの短期の乖離 (実質為替レートの対数) を表すとすると, 長期の PPP は短期では $u_t \neq 0$ を認めるが, 長期では $u_t = 0$ であることが必要条件である. 少なくともこのことが問題になる場合は, u_t が定常プロセスであることが必要条件である. もし u_t が非定常なら, 時間が経るにつれて u_t はどんどん大きくなり, e_t と p_t は際限なく発散する. したがって, もし e_t と p_t が I(1) なら, 長期の PPP は 1 の共和分パラメータで共和分されていることが必要である.

ドルに対する円の為替レートと日米両国の卸売物価の月次データ (1973 年 1

月から 1992 年 12 月) により, 伝統的な共和分分析を用いて PPP を検定する. PPP はこの場合次のように表される.

$$E_t = k\frac{P_t}{P_t^*} \tag{4.1}$$

ここで E_t は t 時点における米国ドルに対する円の名目為替レート, P_t と P_t^* はそれぞれ日本と米国の卸売物価指数である. (4.1) 式の対数をとり, 誤差項を加えると

$$\ln E_t = \ln k + \ln P_t - \ln P_t^* + u_t \tag{4.2}$$

となり, PPP を検定する通常の方法は

$$\ln E_t = \beta_0 + \beta_1 \ln P_t + \beta_2 \ln P_t^* + u_t \tag{4.3}$$

を推定し, 係数制約 $(\beta_1, \beta_2) = (1, -1)$ を検定する.

(4.3) 式の説明変数と従属変数に単位根があると, もし t 検定のような伝統的な統計量を用いると, 見せかけの回帰の問題が生じることがある.

まず, $\ln E_t, \ln P_t, \ln P_t^*$ に単位根があるかどうかを, 拡張 Dickey–Fuller(ADF) 検定と Phillips–Perron の Z_α 検定によって調べる. これらの変数はすべてタイムトレンドがあるので, ドリフト付の ADF と Z_α 検定を用いる. 検定の結果は表 1 にあり, 原系列では $\ln P_t$ の ADF 検定を除いて, すべての ADF と Z_α 検定は単位根という帰無仮説を棄却できなかったが, 他方階差をとった系列は $\ln P_t^*$ の ADF 検定を除いて, すべて 1%有意水準で単位根の仮説を棄却する. そこでこれらの変数はすべてただ一つの単位根をもつという結論になる.

共和分検定に関しては, Engle and Granger の検定と Johansen の検定を用いる.

Engle and Granger 検定

次のステップにしたがって検定を行う.

ステップ 1 (4.3) 式の $\beta_0, \beta_1, \beta_2$ を OLS で推定する.

ステップ 2 OLS 残差に単位根があるかどうかを ADF と Z_αによって検定する.

表 2 にその結果を示した.

ケース I $\ln E_t$ を従属変数として用いる ($\ln E_t$の係数を 1 に正規化).

表 4.1: 単位根検定

	原系列		階差系列	
変数	ADF[a]	Z_α^b	ADF	Z_α
$\ln E_t$	−2.2108	−7.6200	−7.8341	−221.6826
p 値	0.4838	0.6142	1.7284D-10	1.4093D-22
$\ln P_t$	−3.5133	−6.4835	−4.7362	−59.6665
p 値	0.0380	0.7066	0.0006	0.00001
$\ln P_t^*$	−1.5712	−3.7456	−3.2457	−265.8716
p 値	0.8035	0.9010	0.0757	1.7708D-27

a. ラグは $\ln E_t$ は 2, $\ln P_t$ は 3, そして $\ln P_t^*$ は 12 を用いた. また $\Delta \ln E_t$ は 2, $\Delta \ln P_t$ は 3, そして $\Delta \ln P_t^*$ は 12 を用いた.
b. すべての Z_α 検定で打ち切りラグ数は 12 とした.

ケース II $\ln P_t$ を従属変数として用いる.

ケース III $\ln P_t^*$ を従属変数として用いる.

すべての ADF と Z_α 検定は, 共和分がないという帰無仮説を 5% 有意水準ではケース II に対しては ADF 検定と Z_α 検定が, ケース III に対しては Z_α 検定が棄却し, 10% 有意水準ではすべてのケースに対して棄却する.

表 4.2: Engle and Granger の共和分検定

	共和分ベクトル			単位根検定	
ケース	$\ln E_t$	$\ln P_t$	$\ln P_t^*$	ADF[a]	Z_α^b
(I)	1.0	−2.0762	1.6065	−3.1603(0.0926)	−20.1349(0.0687)
(II)	−0.3736	1.0	−0.6902	−3.4626(0.0436)	−24.0902(0.0306)
(III)	0.5550	−1.3251	1.0	−3.3611(0.0568)	−25.0363(0.0251)

a. ADF のラグの長さは AIC によって, それぞれ 2,12,12 が選ばれた. 単位根検定での括弧内の数値は p 値を表す.
b. Z_α のラグの長さは AIC により, それぞれ 2,12,12 が選ばれた.

Johansen の共和分検定

Johansen の提案した方法は, 次のベクトル自己回帰 (VAR) モデルに基ずく正準

相関と考えることができる.

$$X_t = \Pi_1 X_{t-1} + \Pi_2 X_{t-2} + \cdots + \Pi_k X_{t-k} + \epsilon_t \tag{4.4}$$

ここで X_t は $p \times 1$ ベクトルで, その要素はすべて同じ次数の階差定常 (この場合 I(1)) であり, ϵ_t は正規ホワイトノイズである. もし X_t が共和分していれば, $\beta' X_t$ が定常であるような $p \times 1$ ベクトル β が存在する. すると (4.4) は, 誤差修正モデル (ECM) の形に書きなおすことができる.

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi X_{t-k} + \epsilon_t \tag{4.5}$$

ここで, X_t が共和分されていれば $\Pi = \alpha\beta'$ である. $\Pi = \alpha\beta'$ は Π のランクが p より小さいことを意味している. したがって, Johansen の検定は次の仮説に対して行う.

$$\begin{array}{lllll} H_0 & : & \Pi = \alpha\beta' & \text{あるいは} & \operatorname{rank}\Pi \leq r,\ r < p \\ H_1 & : & \Pi \neq \alpha\beta' & \text{あるいは} & \operatorname{rank}\Pi = p \end{array} \tag{4.6}$$

もし $r \neq 0$ なら X_t は共和分されておらず, (4.5) は ΔX_t での VAR モデルになる. もし $r = 0$ なら X_t は定常であり, 定常変数の線形結合はまた定常であるから, 共和分を検定する必要はない. もし $0 < r < p$ なら, X_t は共和分されている. もし帰無仮説 $H_0 : r \leq 0, \cdots, H_0 : r \leq s-1$ (ここで $s < p$) が棄却されるが $H_0 : r \leq s$ は棄却されなければ, Π のランクは s であり, したがって s 個の共和分ベクトルがある (β は $p \times s$ 行列でそのランクは s).

Johansen の検定の結果は次の通りである. (4.6) の仮説検定は帰無仮説 $H_0 : r = 0$ を棄却し, したがって $\ln E_t, \ln P_t, \ln P_t^*$ 間に共和分関係が存在する. $H_0 : r \leq 1$ もまた5%水準で棄却され, $H_0 : r \leq 2$ もまた5%水準で棄却されるので, 3 つの共和分関係が存在するかもしれない. しかし, 1%水準では 1 つだけの共和分間が存在する.

$\ln E_t, \ln P_t, \ln P_t^*$ 間の共和分は Johansen の検定によって支持されるので, 次の制約を検定する

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{4.7}$$

これは帰無仮説のもとで χ^2 分布に収束する Johansen (1991) の尤度比検定によって行うことができる. 検定結果から, 1%有意水準で帰無仮説を棄却することが分かる.

表 4.3: **Johansen の検定**

	$H_0 : \Pi = \alpha\beta'$	
	検定統計量	p 値
$r = 0$	48.750	0.0012
$r \leq 1$	22.215	0.0141
$r \leq 2$	4.222	0.0371

ラグの長さは AIC により 2 とした.

Johansen の検定は為替レートの対数, 日米の卸売物価指数の対数間に長期の市場均衡関係が存在するという結論に導く. しかし, もしもっと厳密な PPP 比例条件 (4.7) を検定すると PPP は円の為替レートに対して成立しない. $\ln E_t, \ln P_t, \ln P_t^*$ 間の共和分関係が棄却されれば, PPP が成立しないという結論になる. しかし実証分析の結果からは, 共和分関係が存在する. 制約 (4.7) は強すぎ, 取引費用, 測定誤差, 貿易障壁等々のような PPP にともなう通常のいくつかの問題を考慮していないためであるということができる.

4.4 単位根のベイズ分析

まず ARMA 誤差の次数が未知の場合のトレンドモデルにおいて単位根がある場合を示し, 系列がトレンド定常プロセス (TSP) と階差定常プロセス (DSP) のどちらであるかを中心に検定する. モデルは, 次のように与えられる.

$$y_t = \mu + \alpha t + \epsilon_t \tag{4.8}$$

$$\epsilon_t = \rho\,\epsilon_{t-1} + \frac{\Theta(B)}{\Phi(B)}\,u_t \tag{4.9}$$

ここで $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ と $\Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ である. $\Theta(B)$ は反転可能で $\Phi(B)$ は定常であると仮定されている.

自己回帰誤差のある回帰モデルで単位根を検定するのに, 一般に 2 つのモデル化の方法がある. Said (1991), その他は次の定式化にしたがう. $\Pi(B) = \Phi(B)/\Theta(B) = 1 + \pi_1 B + \pi_2 B^2 + \cdots$ とおくと, (4.9) の多項式は次のように表される.

$$(1 - \rho B)(1 + \pi_1 B + \pi_2 B^2 + \cdots) \tag{4.10}$$

そして (4.8) は

$$e_t = \pi_1 e_{t-1} + \pi_2 e_{t-2} + \cdots + u_t \tag{4.11}$$

であり, ここで

$$\begin{aligned} e_t &= y_t(\rho) - x_t(\rho)\beta \\ y_t(\rho) &= y_t - \rho\, y_{t-1} \\ x_t(\rho) &= x_t - \rho\, x_{t-1}, \qquad x_t = (1,\, t) \end{aligned}$$

である.

もう一つの表現法は, 単位根や共和分の文献でより多く使用されており, (4.10) を

$$a(B) = \eta\, B + (1 - B)(1 - \psi_1 B - \psi_2 B^2 - \cdots) \tag{4.12}$$

と表し, ここで $\eta = a(1)$ である. この表現法では (4.8) 式は

$$\Delta y_t = \mu^* + \eta\, y_{t-1} + \alpha^*\, t + \sum_{i=1}^{\infty} \psi_i \Delta y_{t-i} + u_t \tag{4.13}$$

となり, ここで μ^* は新たな定数項, そして $\alpha^* = \alpha\eta$ である.

Δy_{t-i} のラグを p 期で打ち切ると, (4.9) は次のようになる.

$$\Delta y_t = \mu^* + \eta\, y_{t-1} + \alpha^* t + \sum_{i=1}^{p} \psi_i \Delta y_{t-i} + u_t \tag{4.14}$$

(4.14) 式を用いると, DST と TSP 仮説はこの場合次のようになる.

$$\mathrm{H}_0 :\ \eta = 0 \quad \text{and} \quad \alpha^* = 0 \qquad \text{vs.} \qquad \mathrm{H}_1 :\ \eta \neq 0 \tag{4.15}$$

$\delta = (\eta,\, \alpha^*)'$ に対する事後密度を導くために, 事前密度を与える必要がある. 多くのベイズ分析では非情報フラット事前分布や Jeffreys 型の不変事前分布が良好な結果を得る. しかし, 単位根のベイズ検定では, フラットな事前分布を用いると $\eta = 0$ の帰無仮説が真であるときに, η の負の領域に大部分の密度がある事後分布になる. そこで DSP が真であるときに, DSP よりも TSP を受け入れやすい. この点を考慮して, Phillips (1991a) は Jeffreys の事前分布を用いて事後分布を 1 の方に引き寄せる望ましい結果を得ているが, その場合事後分布が双峰分布になり, $\rho > 1$ の発散区域内では Jeffreys 事前分布は指数的に増大す

るので裾が厚くなる結果を生む傾向がある. Jeffreys (1961, pp.358–359) は系列相関の分析に用いる場合, 係数が 1 のときにこの Jeffreys 型の不変原理を用いる点についていくつかの問題を指摘している.

ここでフラットな事前分布を使うときの問題を説明しよう. (4.14) の η の最尤推定量 (MLE) $\hat{\eta}$ は, $\eta = 0$ のとき負になる傾向がある. Fuller (1976) は誤差項がホワイトノイズの場合 (すなわち (4.9) で $\Theta(B) = \Phi(B) = 1$) に, このことを示した. Sims and Uhlig (1991) は誤差項がホワイトノイズの場合に, $\hat{\rho}\,(= \hat{\eta}+1)$ の標本分布を描くことによってこのことをグラフ的に示した. すなわち, 分布は定常区域に密度の重心があり左に偏っている.

単位根あるいは DSP を検定するのに, 事後密度を $\eta = 0$ の方向に引っ張る事前密度を選択する必要がある. 以前に Jeffreys 型事前分布は問題があると述べた. DSP 対 TSP を検定するのに, DeJong and Whiteman (1991a) は AR(3) モデルを用いて最大根に対する事後密度を導き, それが 1 であるか否か検定した. 彼らは AR パラメータにフラット事前分布を用い, AR(3) 過程の根に変換した後でこのフラット事前分布は, 1 の近傍にもっとウェイトをおくことによって最大根に対する事前分布を生じ, 事後密度を僅かに単位根の方に引っ張ることになる. 彼らの事前密度は 1 の近傍で切断されている.

本論文では, 次のような自然共役事前分布を用いることを提案する.

$$\mathrm{p}(\mu^*, \eta, \alpha^*, \psi_1, \cdots, \psi_k, \sigma) \propto \sigma^{-1}\,\mathrm{p}(\eta,\, \alpha^*|\sigma) \tag{4.16}$$

ここで

$$\mathrm{p}(\eta,\, \alpha^*|\sigma) \propto \exp\left\{-\frac{1}{2\sigma^2}\,(\delta - \delta_0)'\,V_0\,(\delta - \delta_0)\right\} \tag{4.17}$$

ここで $\delta' = (\eta,\, \alpha^*)$, そしてこの事前分布を $\eta = 0$ と $\alpha^* = 0$ の帰無仮説の値の中心におく. δ の事後密度を導くために, (4.14) を次のように書きなおす.

$$w = Z\delta + X\beta + u$$

ここで

$$w = \begin{bmatrix} \Delta y_1 \\ \vdots \\ \Delta y_n \end{bmatrix} \quad Z = \begin{bmatrix} y_{-1} & 1 \\ \vdots & \vdots \\ y_{n-1} & n \end{bmatrix} \quad X = \begin{bmatrix} 1 & \Delta y_0 & \cdots & \Delta y_{1-k} \\ \vdots & \vdots & & \vdots \\ 1 & \Delta y_{n-1} & \cdots & \Delta y_{n-k} \end{bmatrix}$$

そして $\beta' = (\mu^*, \psi_1, \cdots, \psi_k)$. δ の事後分布は次式によって与えられ

$$\mathrm{p}(\delta|\mathrm{data}) \propto \left[\nu s^2 + d + (\delta - \tilde{\delta})'(V_0 + V)(\delta - \tilde{\delta})\right]^{-(n-k-1)/2} \tag{4.18}$$

ここで $V = Z'M_xZ$, $M_x = I - X(X'X)^{-1}X'$, $\nu s^2 = w'M_xw$,

$$d = \widehat{\delta}'V_0(V_0 + V)^{-1}V\widehat{\delta} \qquad \widehat{\delta} = V^{-1}Z'w, \qquad \tilde{\delta} = (V_0 + V)^{-1}V\widehat{\delta}$$

(4.18) を用いるためには, 2 つのことを決めねばならない. (i) (4.14) 式におけるラグの長さ p と, (ii)(4.17) 式の事前分布における V_0 の値である. 過去の多くの研究では, ラグの長さは AIC のような尺度やホワイト・ノイズの χ^2 検定によって決定される. AIC や χ^2 検定は確かに役に立ち, ラグの長さが分からない場合にはある程度拠り所となるが, AIC は複数個の最小値を持つ場合があることや χ^2 検定はラグの長さが, たとえば $p = 6$ ではホワイトノイズを示すが, $p = 10$ では示さない場合がある.

ここで事前密度の V_0 の値を得るために, ブートストラップ法を用いることにする.

(1) (4.14) 式のパラメータの MLE $\widehat{\theta} = (\widehat{\mu}^*, \widehat{\eta}, \widehat{\psi}_1, \cdots, \widehat{\psi}_p)$ と $\widehat{\sigma}$ を用いる. $\widehat{u}_1, \cdots, \widehat{u}_n$ を $\mathrm{N}(0, \widehat{\sigma}^2)$ から抽出し, ダイナミック・シミュレーションを行う.

$$\Delta y_t = \widehat{\mu}^* + \widehat{\eta}\, y_{t-1} + \widehat{\alpha}^* t + \sum_{i=1}^{p} \widehat{\psi}_i \Delta y_{t-i} + \widehat{u}_t.$$

計算された Δy_t が利用できるようになると, y_{t-i} の実際のデータを置き換える. y_t の $n + r$ 個の値を発生した後, y_t の最初の r 個の計算値を棄て、残りの n 個の値を用いる.

(2) y_t の n 個のシミュレートされた値を用いて, 新しい $\theta = (\mu, \eta, \alpha^*, \psi_1, \cdots, \psi_p)$ を推定する. それらを $\widehat{\theta}^{(1)} = (\widehat{\mu}^{(1)}, \widehat{\eta}^{(1)}, \widehat{\alpha}^{*(1)}, \widehat{\psi}_1^{(1)}, \cdots, \widehat{\psi}_p^{(1)})$ とする.

(3) (1) と (2) を R 回繰り返す. $\widehat{\delta}^{(i)} = (\widehat{\eta}^{(i)}, \widehat{\alpha}^{*(i)})'$ を $i = 1, \cdots, R$ に対して集め

$$C = \mathrm{Cov}(\delta) = \frac{1}{R}\sum_{i=1}^{R}\left(\widehat{\delta}^{(i)} - \bar{\widehat{\delta}}\right)\left(\widehat{\delta}^{(i)} - \bar{\widehat{\delta}}\right)' \tag{4.19}$$

を計算する. ここで $\bar{\widehat{\delta}} = \frac{1}{R}\sum_{i=1}^{R}\widehat{\delta}^{(i)}$. そこで $V_0 = C^{-1}$ とする.

ステップ (3) で求めたものは, δ の無条件共分散である. この代わりに, ψ_i が与えられたときの δ の条件付共分散を用いることができる. これは $\widehat{\theta}^{(i)} = (\widehat{\mu}^{(i)}, \widehat{\eta}^{(i)}, \widehat{\alpha}^{*(i)}, \widehat{\psi}_1^{(i)}, \cdots, \widehat{\psi}_p^{(i)})'$ を集めて, θ の分散共分散行列を計算する:

$$C^* = \mathrm{Cov}(\theta) = \frac{1}{R}\sum_{i=1}^{R}\left(\widehat{\theta}^{(i)} - \bar{\widehat{\theta}}\right)\left(\widehat{\theta}^{(i)} - \bar{\widehat{\theta}}\right)' \tag{4.20}$$

ここで $\bar{\widehat{\theta}} = \frac{1}{R}\sum_{i=1}^{R}\widehat{\theta}^{(i)}$. V_0 を C^{*-1} の2番目と3番目の行と列からなる $C^{*-1}[2:3, 2:3]$ の 2×2 行列として得る. これを V_0^* と表す. V_0^* は $\psi_1, \cdots, \psi_p$ 所与での δ の共分散の逆行列であるから, V_0^* は V_0 よりも大きくなり, したがって $V_0^* - V_0$ は正準定符号になる. したがって, V_0 の代わりに V_0^* を用いると, 事後密度は η と α^* の帰無仮説の値の方向に強く引かれる傾向にある (図1を参照).

θ と σ の MLE, $\widehat{\theta}$ と $\widehat{\sigma}$, を用いることは, もし (4.14) が y_t を適切に表していると信じるなら, 一致性をもつ $\widehat{\theta}$ と $\widehat{\sigma}$ は尤度関数から事前分布のパラメータを推定する最良の推定になる. この方法はある種の経験ベイズ的な考え方に基づいている.

R 回の抽出のそれぞれに対して MLE である $\widehat{\theta}$ と $\widehat{\sigma}$ を用いる代わりに, $i+1$ 回目の抽出に $\widehat{\theta}^{(i)}$ と $\widehat{\sigma}^{(i)}$ を用いることもできる. それはこれらの推定値が θ と σ の確率的な性質を反映しているからである. これは正に Gibbs サンプラーの方法である. t と y_{t-1} のような強いトレンドがある説明変数の場合には, y_t の拡張されたデータが大きく変動するために Gibbs サンプラーが難しい場合がある.

いくつかのベイジアン・ブートストラップ法がある [Rubin (1981, 1984), Hill (1990) そして Bedrick and Hill (1992)]. ブートストラップ法は用いるモデルとデータを反映するように計画されており, 頻度派-ベイズ派の折衷的な弾力性を持っている.

(4.18) から η と α^* の事後分布を評価することができる:

$$\mathrm{p}(\eta|\mathrm{data}) \propto \left[\nu s^2 + d + \frac{(\eta - \tilde{\eta})^2}{v^{11}}\right]^{-(n-p-2)/2} \tag{4.21}$$

$$\mathrm{p}(\alpha^*|\mathrm{data}) \propto \left[\nu s^2 + d + \frac{(\alpha^* - \tilde{\alpha}^*)^2}{v^{22}}\right]^{-(n-p-2)/2} \tag{4.22}$$

ここで v^{ii} は $(V_0 + V)^{-1}$ の i 番目の対角要素である.

η と α^* の周辺事後分布を示そう. $\tilde{\eta}$, $\tilde{\alpha}^*$, V と νs^2 の 2,000 回の反復の平均を用いると, 事後密度は平均的な場合を表す. サンプルサイズを $n = 100$ とし,

MA(1) パラメータの値を $\theta_1 = -.8,\ 0,$ と.8 とする. ブートストラップ法では, $r = 100$ とおき, $n + r = 200$ 個の観測値を発生し, 最初の 100 個の観測値を棄てる.

図 1 は η と α^* の事後密度のグラフを表している. 各図には 3 つの事後密度がある. 直線は $V_0 = C^{-1}$ を用いた報知事前分布の場合の事後分布で *inform.prior 1* とする. 点鎖線は $V_0^* = C^{*-1}[2:3,\ 2:3]$ を用いた報知事前分布の場合の事後分布で *inform.prior 2* とする. 点線は δ に対するフラットな事前分布の場合の事後分布で *diffuse prior* とする. 図 1 から次のことが分かる.

(1) すべての場合で, η の事後分布の図は報知事前分布は事後密度を$\eta = 0$ の帰無仮説の値の方向に引っ張る傾向にあることを示している. 報知事前分布はフラットな事前分布の場合よりはモードの値の回りに大きな密度がある.

(2) η の事後密度は MA(1) パラメータの値 θ_1 に敏感である. $\theta_1 = -.8$ のとき, η の事後密度は $\theta_1 = 0$ あるいは .8 の場合よりもずっと広がっている.

(3) α^* の事後密度は, 事前密度と同様 θ_1 のさまざまな値に対して頑健である.

(4.15) における DSP 対 TSP 仮説を検定するために, 最高事後密度区域 (HPDR) あるいは最高事後密度区間 (HPDI) 検定を行う. もし HPDR 検定なら α_1 の所与の水準 (たとえば $\alpha_1 = 0.05$) に対して, $(1 - \alpha_1)$ HPDR 等高線が点 $(\eta = 0,\ \alpha^* = 0)$ を含むか否かを見る. HPDI 検定の場合には, 個々に η の $(1 - \alpha_1)$ HPDI が $\eta = 0$ を含み, α^* の $(1 - \alpha_1)$ HPDI が $\alpha^* = 0$ を含むかどうか見る. 単に点推定値や検定統計量を見るだけよりも同時事後分布や周辺事後分布をプロットすると, η と α^* についてもっと多くの情報を得ることができる.

4.5 共和分のベイズテスト

DSP 対 TSP 仮説を検定した後, 対応する G 個のそれぞれの変数が $I(1)$ プロセスであると結論すると, 次に共和分検定を行う. 本節では, ベイズ共和分検定を作成する. 最初に, 散漫な (*diffuse*) 事前分布を用いて, 共和分関係を含む誘導型係数行列, Π, に対する事後密度を導く. そして, HPDR 検定が $\Pi = 0$ を検

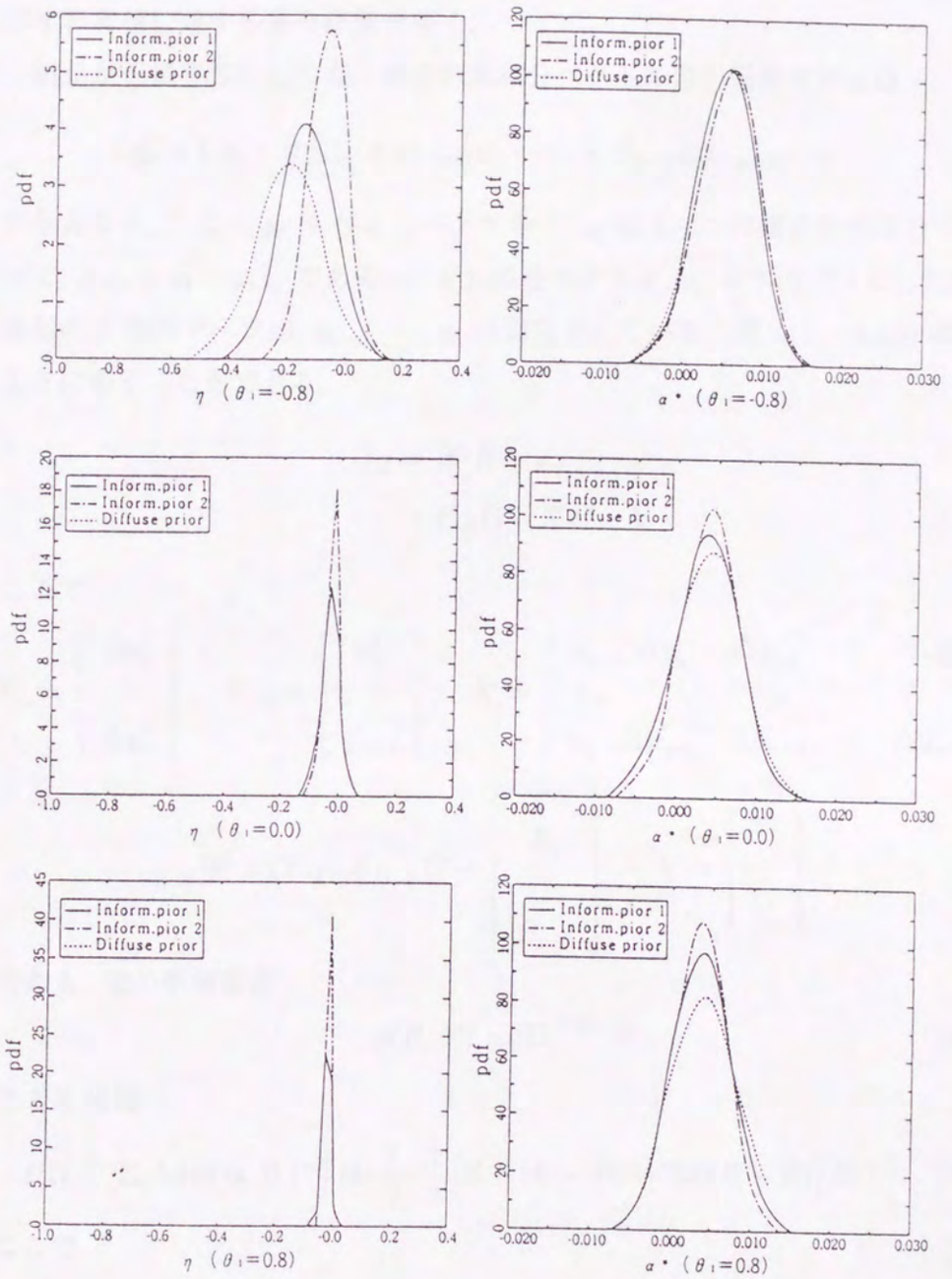


図 4.1: η と α^* の事後密度

定する Johansen の尤度比検定 (LRT) に等しいことを示す. 次に, Π のランクを示す特異値に対する事後密度を導く.

誤差が正規分布にしたがう場合の共和分ベクトル自己回帰モデルは

$$\Delta y_t = \Gamma z_t + \Pi y_{t-1} + B_1 \Delta y_{t-1} + \cdots + B_{p-1} \Delta y_{t-p+1} + v_t \tag{4.23}$$

で与えられ, ここで y_t は $G \times 1$ ベクトルで, z_t は $\ell \times 1$ の確定的要素のベクトルで, $\Delta x_t = x_t - x_{t-1}$ である. $G \times 1$ 誤差ベクトル, v_t は $\mathrm{N}(0, \Sigma)$ にしたがう. 最初の p 個のデータ点, $y_{1-p}, \cdots, y_0$ は固定されていると考える. (4.23) は次のように書くことができる.

$$\begin{aligned} Y_\Delta &= W B + V \\ &= Y_{-1}\Pi + X C + V \end{aligned} \tag{4.24}$$

ここで

$$Y_\Delta = \begin{bmatrix} \Delta y_1' \\ \vdots \\ \Delta y_n' \end{bmatrix}, \quad Y_{-1} = \begin{bmatrix} y_0' \\ \vdots \\ y_{n-1}' \end{bmatrix}, \quad X = \begin{bmatrix} z_1' & \Delta y_0' & \Delta y_{-1}' & \cdots & \Delta y_{2-p}' \\ \vdots & \vdots & \vdots & & \vdots \\ z_n' & \Delta y_{n-1}' & \Delta y_{n-2}' & \cdots & \Delta y_{n-p+1}' \end{bmatrix}$$

$$W = (Y_{-1}, X), \quad C = \begin{bmatrix} \Gamma' \\ B_1 \\ \vdots \\ B_{p-1} \end{bmatrix}, \quad V = \begin{bmatrix} v_1' \\ \vdots \\ v_n' \end{bmatrix}$$

である. 次の事前密度

$$\mathrm{p}(B, \Sigma) \propto |\Sigma|^{-(G+1)} \tag{4.25}$$

と尤度関数

$$\ell(\Pi, C, \Sigma | \mathrm{data}) \propto |\Sigma|^{-\frac{n}{2}} \operatorname{etr}\left\{ -\frac{1}{2} \left[S + (B - \widehat{B})' W' W (B - \widehat{B}) \right] \Sigma^{-1} \right\} \tag{4.26}$$

ここで

$$\begin{aligned} S &= Y_\Delta' M_w Y_\Delta, \quad M_w = I - W(W'W)^{-1}W' \\ \widehat{B} &= (W'W)^{-1} W' Y_\Delta \end{aligned}$$

が与えられていると, Π の周辺事後密度は

$$\mathrm{p}(\Pi | \mathrm{data}) \propto \left| S + (\Pi - \widehat{\Pi})' F (\Pi - \widehat{\Pi}) \right|^{-(n-k_2)/2} \tag{4.27}$$

で与えられる. ここで $k_2 = [\ell + (p-1)]G$ であり, さらに

$$\hat{\Pi} = (Y_{-1}M_xY_{-1})^{-1}Y'_{-1}M_xY_{\Delta}, \quad M_x = I - X(X'X)^{-1}X'$$
$$F = Y'_{-1}M_xY_{-1}$$

である. もし共和分関係が存在しなければ, $\Pi = 0$, であり, 共和分関係がないという検定は

$$\mathrm{H}: \ \Pi = 0 \qquad \text{vs.} \qquad \mathrm{K}: \ \Pi \neq 0 \tag{4.28}$$

で与えられる.

(4.27) から Box and Tiao (1973) にしたがう HPDR 検定を得る. 彼らは, 点 $\Pi = \Pi_0$ が最高事後密度区域 (HPDR) に含まれるか否かを決めるのに, 次の統計量を用いることを示唆した.

$$U(\Pi_0) = \left|I + S^{-1}(\Pi_0 - \hat{\Pi})'F(\Pi_0 - \hat{\Pi})\right|^{-1} \tag{4.29}$$

Box and Tiao は

$$\Pr[U(\Pi) > U(\Pi_0)|\text{data}] \sim \Pr[\chi^2_{G^2} < -n^* \log U(\Pi_0)] \tag{4.30}$$

で近似した. ここで n^* は標本サイズ n の近似関数である [Box and Tiao (1973, p.452)]. ところで今 $\Pi = 0$ を検定したいのであるから, $\log U(0)$ は

$$-n^* \log U(0) = n^* \log \left|I + S^{-1}\hat{\Pi}'F\hat{\Pi}\right| \tag{4.31}$$

で与えられる.

$n^* = n$ とおくと, Π のランクがゼロであることを検定する場合に (4.31) は Johansen の LRT に等しいことを示すことができる. Johansen の LRT は次式によって与えられる.

$$\mathrm{LRT} = -n \log \left|I - S_{11}^{-1}S_{10}S_{00}^{-1}S_{01}\right| \tag{4.32}$$

ここで

$$S_{00} = Y'_{\Delta}M_xY_{\Delta} \quad S_{01} = Y'_{\Delta}M_xY_{-1} \quad S_{11} = Y'_{-1}M_xY_{-1}$$

である. (4.31) と (4.32) は同一であり, このことは次からすぐに示すことができる.

$$S = S_{00} - S_{01}S_{11}^{-1}S_{10}, \qquad \hat{\Pi}'F\hat{\Pi} = S_{01}S_{11}^{-1}S_{10} \quad \text{そして} \quad |I + AB| = |I + BA|.$$

(4.31) から

$$\begin{aligned}&\left|I+S^{-1}\hat{\Pi}'F\hat{\Pi}\right|\\&=\left|I+(S_{00}-S_{01}S_{11}^{-1}S_{10})^{-1}S_{01}S_{11}^{-1}S_{10}\right|=\left|I-S_{00}^{-1}S_{01}S_{11}^{-1}S_{10}\right|^{-1}\\&=\left|I-S_{11}^{-1}S_{10}S_{00}^{-1}S_{01}\right|^{-1}\end{aligned}$$

である.

(4.31) の HPDR の $\chi^2_{G^2}$ 近似は, 多変量 t 分布の性質を用いて得ることができる. もし標本検定統計量として $-n\log U(0)$ を見ると, 極限分布はブラウン運動の m-次元関数のトレースを含む [たとえば, Johansen (1988, 1991) と Kunitomo (1992) を参照]. Johansen (1988) はこの極限分布は大標本の場合 $c\chi^2_f$ によって正確に表されることを示した. ここで (4.23) の Γ がゼロのとき $f=2G^2$ と $c=.85-.58/f$ である.

(4.23) の z_t が定数項であるとき, 新たに修正した定数 c は $c=1.01-.31/f+5.9G/n$ によって与えられることを標本実験によって確かめることができる. 一方自由度は依然として $f=2G^2$ で近似できる. これは LRT の臨界値が (4.23) の確定項 z_t であることを示している.

修正因子 c を別にすれば, 散漫な事前分布に基づいた HPDR 検定と Johansen の LRT の間の主要な違いは χ^2 に対する自由度にある : 前者に対しては G^2 であり, 後者に対しては $2G^2$ である. したがって, HPDR 検定は共和分がないという帰無仮説を棄却する傾向が Johansen の LRT よりもずっと強く, 単位根の場合と同様に, 散漫な事前分布を用いることに伴う問題は共和分検定にも当てはまることが分かる.

したがって, 事後分布を $\Pi=0$ の方向に引き寄せる情報をもつ事前分布 (*informative prior*) を用いる必要がある. さらに, $\Pi=0$ (すなわち, Π のランクがゼロ) だけであるだけでなく, Π のランクが $1,2,\cdots,G-1$ であることを検定できるようなベイズテストを得る必要がある. (Johansen の LRT 検定は Π のランクが高々 r, $r=1,\cdots,G-1$ である検定である.)

行列のランクを調べる場合, 固有値を調べる方が簡単である. 非対称である Π から固有値を求めようとする場合, (4.23) の G 本の方程式の順序によって固有値が変わり, その固有値が複素根になることがあるという問題がある. 複素根

は Stock and Watson (1988) が行ったように, その実根部分だけを考えることによって扱うこともできるが, G 本の方程式が特定の順序にあるという強い先見的な理由がない限り, 固有値が方程式の順序とともに変わるという問題が生じる. これらの問題を避けるために, 特異値分解を利用する.

尤度関数 (4.26) は次のように書くことができる.

$$\begin{aligned}\ell(\Pi, C, \Sigma|\text{data}) &\propto |\Sigma|^{-\frac{n}{2}} \text{etr}\left\{-\frac{1}{2}\left[S + (B-\hat{B})'W'W(B-\hat{B})\right]\Sigma^{-1}\right\} \\ &= |\Sigma|^{-\frac{n}{2}} \text{etr}\left\{-\frac{1}{2}\left[S + (\Pi-\hat{\Pi})'F(\Pi-\hat{\Pi}) + (C-\tilde{C})'X'X(C-\tilde{C})\right]\Sigma^{-1}\right\}\end{aligned} \tag{4.33}$$

ここで

$$\tilde{C} = \hat{C} + X'Y_{-1}(Y_{-1}'Y_{-1})^{-1}(\Pi - \hat{\Pi})$$

である.

Π と $\hat{\Pi}$ を特異値分解を用いて書きなおす.

$$\hat{\Pi} = U'LV \quad \text{と} \quad \Pi = \Psi'\Lambda\Upsilon \tag{4.34}$$

ここで $L = \text{diag}(\ell_1, \ell_2, \cdots, \ell_G)$; ℓ_i^2 は $\hat{\Pi}'\hat{\Pi}$ の i-番目の固有値, そして$\Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_G)$; ここで λ_i^2 は $\Pi'\Pi$ の i-番目の固有値; U, V, Ψ, そして Υ は直交行列である. (4.34) を用いて書きなおし

$$(\Pi-\hat{\Pi})'F(\Pi-\hat{\Pi}) = (\Psi'\Lambda\Upsilon - U'LV)'F(\Psi'\Lambda\Upsilon - U'LV)$$

と尤度関数を用いて

$$\begin{aligned}&\ell(\Lambda, \Psi, \Upsilon, C, \Sigma|\text{data}) \\ &\quad\propto |\Sigma|^{-\frac{n}{2}}\text{etr}\left\{-\frac{1}{2}[S + (\Psi'\Lambda\Upsilon - U'LV)'F(\Psi'\Lambda\Upsilon - U'LV)\right. \\ &\qquad\left.+(C-\tilde{C})'X'X(C-\tilde{C})\right]\Sigma^{-1}\Big\}\end{aligned} \tag{4.35}$$

を得る.

ここで特異値, $\Lambda = (\lambda_1, \cdots, \lambda_G)$, に関心があるので, 尤度関数のその他のパラメータは局外母数 (*nuisance parameters*) とする.

次の事前密度 を用いる.

$$\mathrm{p}(\Lambda, \Psi, \Upsilon, C, \Sigma) \propto \mathrm{p}(\Lambda, \Psi, \Upsilon|\Sigma)\mathrm{p}(\Sigma) \propto |\Sigma|^{-\frac{1}{2}} \mathrm{etr}\left\{-\frac{1}{2}S_0\Sigma^{-1}\right\}$$
$$\mathrm{etr}\left\{-\frac{1}{2}(\Psi'\Lambda\Upsilon - \Pi_0)'F_0(\Psi'\Lambda\Upsilon - \Pi_0)\Sigma^{-1}\right\} \qquad (4.36)$$

ここで Π_0 と F_0 は所与のパラメータで, $\mathrm{etr}(\cdot)$ は $\exp(\mathrm{tr})$ を表す. (4.35) と (4.36) を結びつけ, C と Σ について積分すると, Λ, Ψ, と Υ に対する事後密度を得る:

$$\mathrm{p}(\Lambda, \Psi, \Upsilon|\mathrm{data}) \propto |S + S_0 + (\Psi'\Lambda\Upsilon - \Pi_0)'F_0(\Psi'\Lambda\Upsilon - \Pi_0)$$
$$+ (\Psi'\Lambda\Upsilon - U'LV)'F(\Psi'\Lambda\Upsilon - U'LV)|^{-(n-k_2)/2} \qquad (4.37)$$

ここで $k_2 = \ell + (k-1)G$ である.

関心あるパラメータは λ_i であるので, λ_i の周辺密度を得る必要がある. これはすべてのパラメータ (λ_j $(j \neq i)$, Ψ, と Υ) を, Gibbs サンプラーのような数値積分法を用いて積分することによって行う. もう一つの別の方法として, Ψ と Υ それぞれに対する U と V の標本値を用いて, λ_i の条件付周辺事後密度を得る方法がある. 本論文ではこの条件付周辺密度を採用し, 次のようになる.

$$\mathrm{p}(\lambda_i|\Psi = U, \Upsilon = V, \mathrm{data})$$
$$= \int \mathrm{p}(\Lambda|\Psi = U, \Upsilon = V, \mathrm{data})\mathrm{d}\lambda_1 \cdots \mathrm{d}\lambda_{i-1}\mathrm{d}\lambda_{i+1} \cdots \mathrm{d}\lambda_G \qquad (4.38)$$

主要な帰無仮説は $\Pi = 0$ であるので, $\Pi_0 = 0$ とおき, 事前密度 のパラメータ S_0 と F_0 をブートストラップを拡張して得ることにする. 単一方程式 (14) と (4.23) の $(\widehat{\mu}^*, \widehat{\eta}, \widehat{\alpha}^*, \widehat{\psi}_1, \cdots, \widehat{\psi}_k, \widehat{\sigma})$ の代わりに, (4.23) の G 本の多変量回帰と $(\widehat{\Gamma}, \widehat{\Pi}, \widehat{B}_1, \cdots, \widehat{B}_k, \widehat{\Sigma})$ を用いる. ブートストラップ試行の平均として S_0 と F_0 を得る.

$$F_0 = \frac{1}{R}\sum_{i=1}^{R}(Y'_{-1}M_xY_{-1})_i, \quad S_0 = \frac{1}{R}\sum_{i=1}^{R}S_i$$

ここで添字 i は i-番目のブートストラップ試行を表す.

人工的に生成したデータを用いて, 特異値 λ_i の事後密度を示す. (4.23) のベクトルの次元を 3 $(G = 3)$ とする. 3 変量ベクトル自己回帰モデルは, 次式のよ

ここで添字 i は i-番目のブートストラップ試行を表す.

人工的に生成したデータを用いて, 特異値 λ_i の事後密度を示す. (4.23) のベクトルの次元を 3 ($G = 3$) とする. 3 変量ベクトル自己回帰モデルは, 次式のように表される.

$$y_{t1} = 1.0 + y_{t-1,1} + \epsilon_{t1} \qquad y_{01} = 0 \tag{4.39}$$

$$y_{t2} = 1.0 + y_{t-1,2} + \epsilon_{t2} \qquad y_{02} = 0 \tag{4.40}$$

$$y_{t3} = 6.0 + .53y_{t1} - .11y_{t2} + u_{t3} \tag{4.41}$$

$$u_{t3} = \rho_3\, u_{t-1,3} + (1 + \theta_1 B)\epsilon_{t3}, \qquad u_{03} = 0 \tag{4.42}$$

ここで, $i = 1, 2, 3$ に対して $\epsilon_{ti} \sim \mathrm{NID}(0, 1)$. 方程式 (4.41) は $|\rho_3| < 1$ なら共和分関係があることを示している.

図 4.2 は, サンプルサイズが 100($n = 100$) で (4.42) の MA パラメータを $\theta_1 = -.8$ とした場合に対する 3 つの特異値, λ_1, λ_2 と λ_3 に対する事後密度を示している. いま $\rho_3 = 1.0$ とおく, したがって 3 変量ベクトル自己回帰モデルは共和分関係がない.

図 4.2 では, 各 λ_i に対して, 実線は情報のある事前分布 (4.36) を用いた場合の事後密度, 他方点線は散漫な事前分布 (4.25) を用いた場合の事後分布である. それぞれの λ_i に対して, 情報のある事前分布を用いた事後分布は$\lambda_i = 0$ の方向に引き寄せられている. 最大の特異値 λ_1 に対して散漫な事前分布を用いた事後密度はλ_1 がゼロではないことを示している, 一方情報のある事前分布を用いた事後分布は明らかに λ_1 がゼロであることを示している. λ_i の事後密度の形状は図 4.2 の密度に極めて近いので, $\theta_1 = -.8$ に対する結果だけを示した.

4.6 為替レートモデルの単位根検定と共和分分析

次の回帰モデルを考える.

$$\begin{aligned} s_t = \beta_1 &+ \beta_2(m_t - m_t^*) + \beta_3(y_t - y_t^*) \\ &+ \beta_4(r_t - r_t^*) + \beta_5(p_t - p_t^*) + \beta_6 f_t + u_t \end{aligned} \tag{4.43}$$

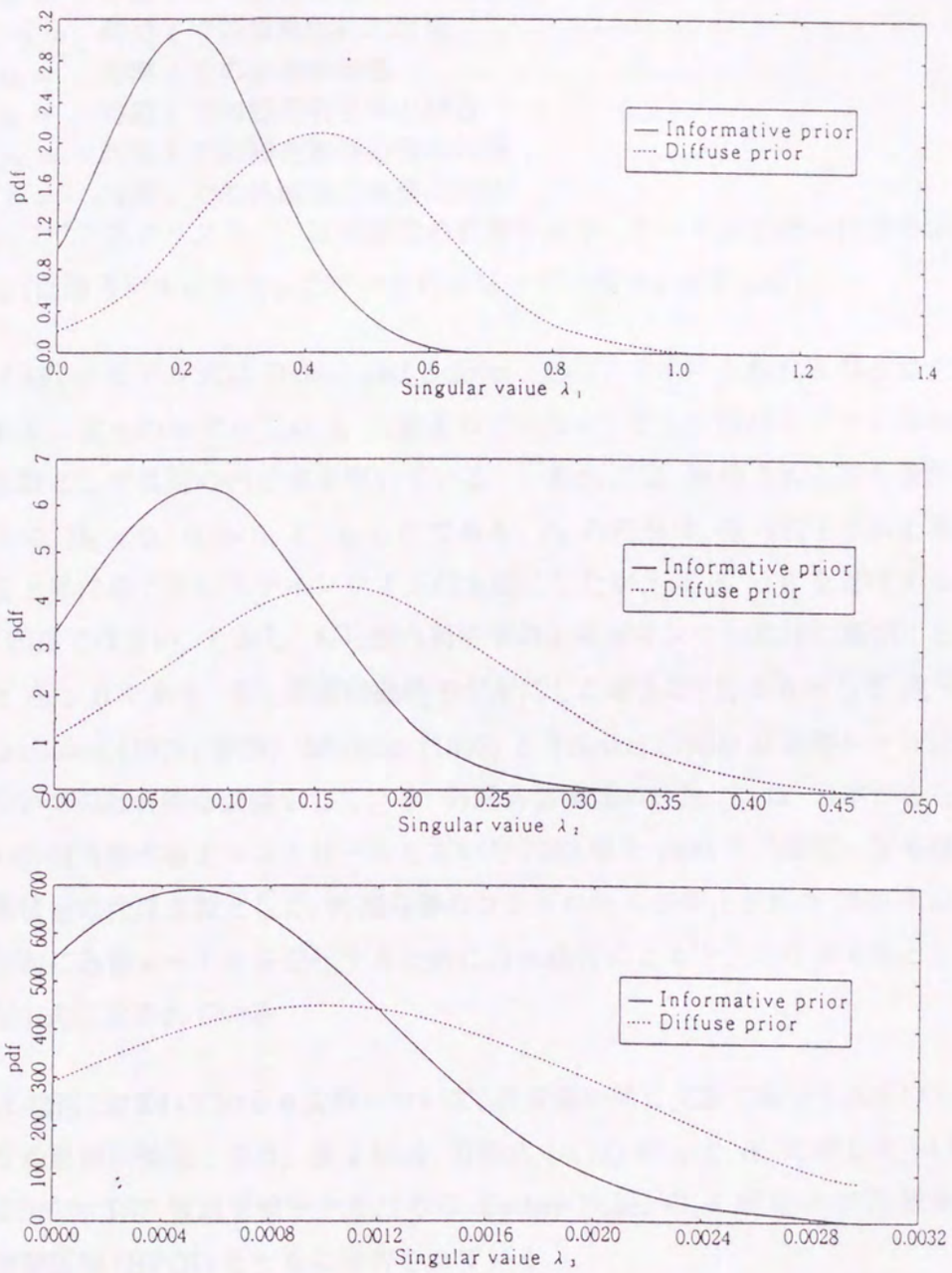


図 4.2: 特異値 $\theta_1 = -.8$ の事後密度

ここで

$s_t =$ 時間 t での名目為替レートの対数
$m_t =$ 時間 t での貨幣供給の対数
$y_t =$ 時間 t での生産の対数
$r_t =$ 時間 t での短期利子率の対数
$p_t =$ 時間 t での卸売物価指数の対数
$f_t =$ 時間 t での外国為替準備の対数

ここでアスタリスク $(*)$ は外国での数量を表す. データの出所は付録で説明した (回帰モデルに使用したデータのプロットは図 4.4 で示した).

(4.43) のモデル式は Baillie and Selover (1987) で用いられたものと似た式である. 彼らのモデルには f_t は含まれていない, そして期待インフレ率の代理変数として長期の利子率を用いている. 一般的には, 期待される符号条件は $\beta_2 > 0$, $\beta_3 < 0$, $\beta_5 > 0$, と $\beta_6 < 0$ である. β_4 の符号は, 国内利子率の上昇が通貨上昇に導くというケインジアンの主張にしたがえば $\beta_4 < 0$ を意味するので, 明確ではない. しかし, もし国内利子率の上昇がインフレ期待に基づくとすると $\beta_4 > 0$ である. もし硬直的価格モデルにしたがえば, $\beta_4 = 0$ そして $\beta_5 > 0$ (Dornbusch (1976) 参照). Mishkin (1992) と Frankel (1993) は為替レート決定についての総合的な議論をしている. 外国為替準備の対数, f_t, は, 日本がある程度の外国為替市場をコントロールしていた 1973 年と 1980 年の期間に貿易収支の累積額の代理変数として, 外国為替のコントロールが中止された 1980 年以降の期間に為替レートを安定化するために日本銀行による介入の代理変数として (4.43) 式に含まれている.

(4.43) に含まれている 6 変数について, 各変数が同じ次数で和分されているかどうか最初に検定しよう. 表 4 には, 方程式 (4.14) の η と α^* に対して (4.15) で DSP 対 TSP 仮説を検定するために, Dickey-Fuller の F 検定が 95% 最高事後密度区間 (HPDI) とともに報告されている.

表 4.4 から, Dickey-Fuller F 検定は 5% 有意水準で $p_t - p_t^*$ 以外のすべての変数は DSP であることが分かる. $p_t - p_t^*$ に対しての 7.15 という F 統計量は 5%水準で有意であるが, 5%水準からそんなに離れていない. 他方, 99%- HPDI

表 4.4: Dickey-Fuller F-検定と最高事後密度区間 (HPDI)

変数	ADF F-統計量	η に対する 99% HPDI	α^* に対する 99% HPDI
s_t	3.98	$(-.07, .025)$	$(-.0004, .0002)$
$m_t - m_t^*$	3.84	$(-.085, .025)$	$(-.0003, .0001)$
$y_t - y_t^*$	3.94	$(-.135, .030)$	$(-.0001, .0003)$
$r_t - r_t^*$	2.64	$(-.050, .021)$	$(-.0002, .0003)$
$p_t - p_t^*$	7.15	$(-.075, .009)$	$(-.0002, .0001)$
f_t	4.15	$(-.050, .015)$	$(-.0001, .0004)$

注: (1) ラグの長さは 12 ($p = 12$) とした.
(2) F 検定の臨界値は 5% 水準で $n = 250$ のとき 6.34.
(3) 99%HPDI = 99% 最高事後密度区間
V_0 に対しては (4.19) の C を用いる.

は $\eta = 0$ と $\alpha^* = 0$ の帰無仮説の値の回りに集中している. V_0 の値を得るのに (4.19) 式を用いた.

6 つの変数がすべて $I(1)$ であったので, 次に Johansen の正準相関検定 [Johansen (1988, 1991)] と特異値に対するベイジアン HPDI 検定によって共和分検定を行う. 多くの研究では, ADF 検定が共和分検定として利用されている.

図 4.3 では 6 つの特異値に対する事後密度を示し, 表 4.5 では Johansen の LRT と特異値に対する 99% HPDI を示した. LRT に対する確率値は χ^2 近似, $c\chi_f^2$ によって計算された. ここで $f = 2G^2$, そして $c = 1.01 - .31/f + 5.9G/n$ である. LRT にしたがうと, 明らかに最高 1 つの共和分関係があり, 他方特異値の 99% HPDI に対しては 3 つの共和分関係がある. λ_3 に対する 99% HPDI は 0 を含んでいないけれども, 区間は (.005, .07) であり, これは数値的にはゼロに近いとして扱うことができる. したがって, ベイズ法によれば 2 つの共和分関係があるといえよう.

誤差修正モデルに関心があれば, ベクトル自己回帰モデル (4.23) を推定する. Π を誘導型係数として扱うと, (4.23) は 1 ステップで推定できる. Π に制約をおくことによって制約付最尤法あるいはベイズ推定を行うことができる. 誤差修正モデルの推定はここでは関心がないので, 最小 2 乗法によって推定した共和分関係だけを報告することにする.

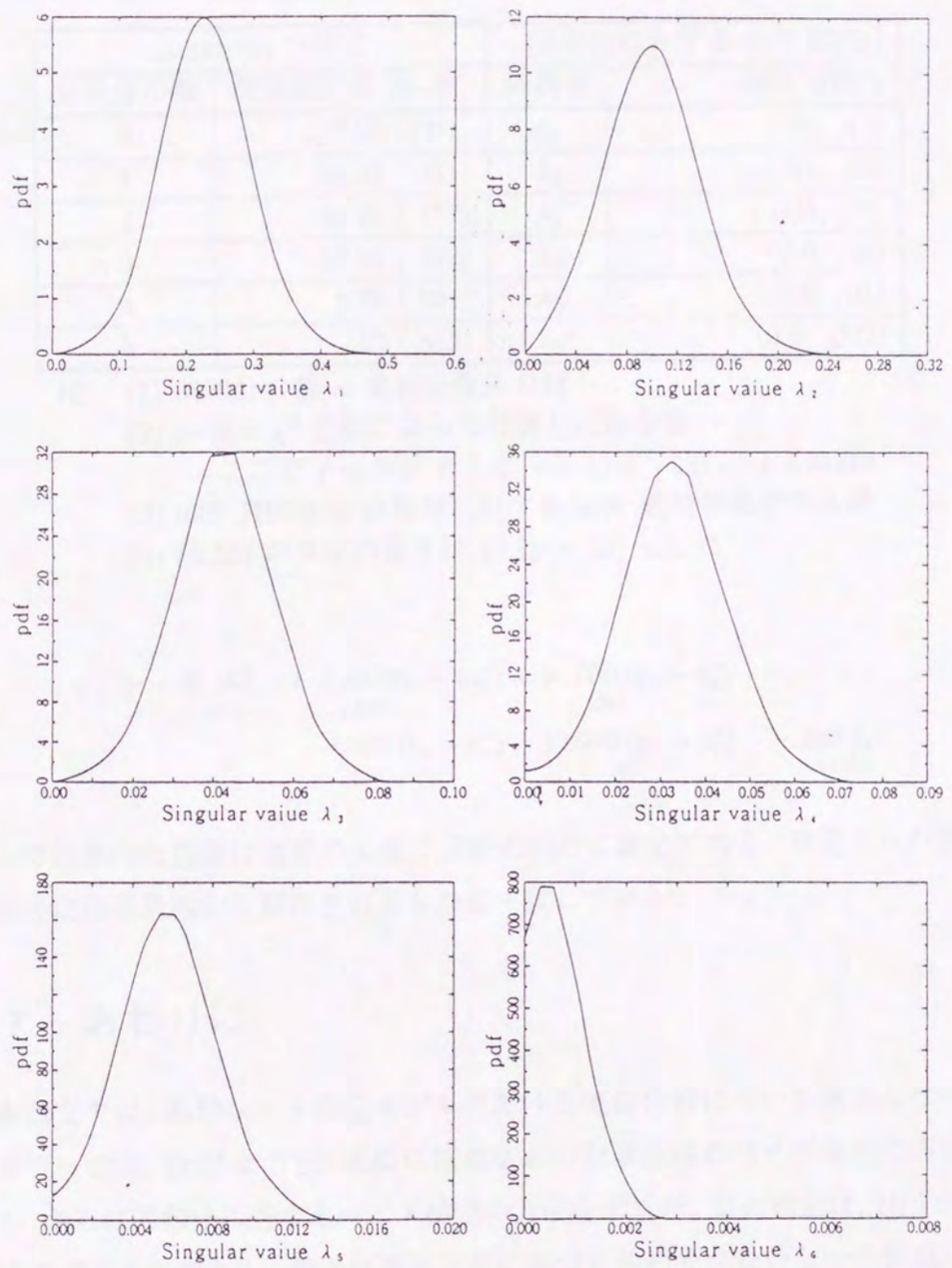


図 4.3: 特異値 (為替レートモデル) の事後密度

表 4.5: 共和分検定

Johansen の検定		特異値に対する 99% HPDI	
共和分の数	検定統計量 (p–値)	特異値	99% HPDI
0	117.00 (.001)	λ_1	(.05, .41)
1	69.95 (.031)	λ_2	(.01, .23)
2	39.25 (.177)	λ_3	(.005, .07)
3	19.99 (.333)	λ_4	(0.0, .06)
4	6.78 (.561)	λ_5	(0.0, .01)
5	.25 (.885)	λ_6	(0.0, .002)

注: (1) 共和分の数 = 共和分関係の数
(2) p–値= χ^2 近似によって計算した確率値: $c\chi^2_f$,
ここで $f = 2G^2$ そして $c = 1.01 - .31/f + 5.9G/n$.
(3) 99% HPDI = 特異値に対する 99% 最高事後密度区間
(4) (4.23) のラグの長さは 12 ($p = 12$) とした.

$$
\begin{aligned}
s_t \;=\; & \underset{(.237)}{9.163} \;+\underset{(.065)}{.160}\,(m_t - m_t^*) \;+\underset{(.139)}{.730}\,(y_t - y_t^*) \\
& -\underset{(.014)}{.082}\,(r_t - r_t^*) \;+\underset{(.116)}{.563}\,(p_t - p_t^*) \;-\underset{(.025)}{.368}\, f_t
\end{aligned}
$$

ここで括弧内の数値は通常の t-値で診断のために載せてある. 推定された係数の符号は経済理論から期待されるものに一致している.

4.7 おわりに

本論文では, 為替レート決定モデルにおける実証分析について簡単なサーベイを行った後, DSP と TSP 仮説に焦点をおいた単位根のベイズ検定の方法を示し, さらに共和分に対するベイズ検定の方法を示した. この検定は, HPDR 原理に基づくものである. 散漫な事前分布に基づく共和分がないという仮説に対する HPDR 区域は Johansen の LRT と同等であることを示した. 次に, ランクすなわち共和分関係の数を表す特異値に対する事後密度を導いた. 円ドル為替レートモデルにこのベイズテストの方法用いて, 1973 年 1 月から 1992 年 12 月までの月次観測値を用いた分析を行った. その結果として回帰モデルに含まれる変数が I(1) であり, 共和分関係が存在することを示した.

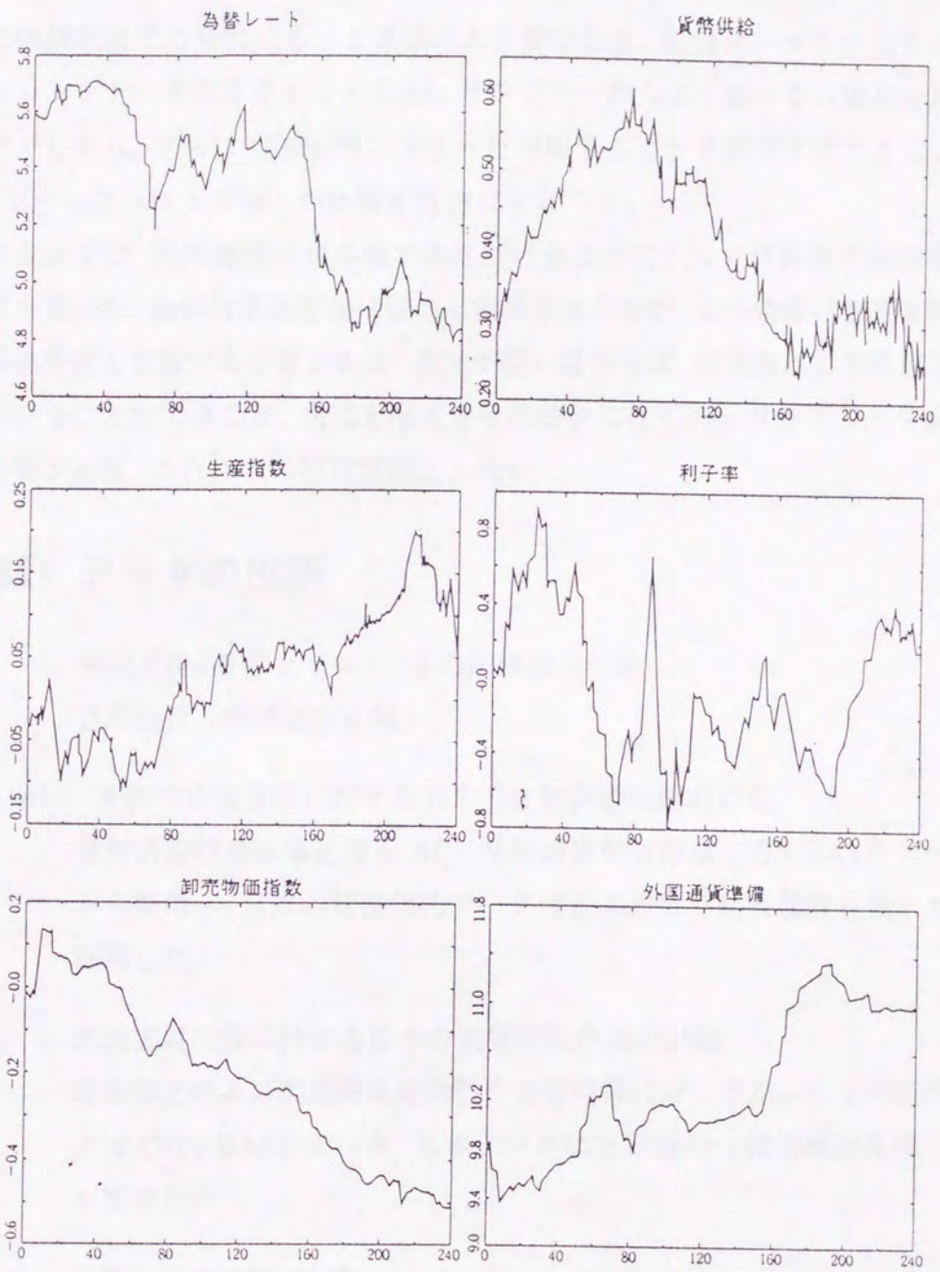


図 4.4: 使用したデータのプロット

トストラップ法を用いたが, データ拡張を伴ったGibbsサンプリングのような別のリサンプリングの方法を用いることも可能である. この件に関するわれわれの準備段階での研究はもっと煮詰める必要がある. 拡張データを生成する際にトレンドデータが存在するとGibbsサンプラーがうまく働かない場合が見られた. しかし, さらにGibbsサンプラーを利用することを研究することによって, ブートストラップ法との比較が可能になる.

本論文では, 局外母数に標本値で条件づけることによって特異値の周辺事後密度を導いた. 条件付事後密度は妥当な結果を生じるが, この条件付密度を無条件事後密度と比較する必要がある. 次元が低い場合には, 球積法による数値積分を用いることができるが, 次元を増大させた場合にはGibbsサンプラーを試みる必要がある. これは次の研究課題としたい.

付録: データの出所

$s =$ 米国ドル (スポットレート) の円価格の対数
日本銀行「経済統計月報」

$m - m^* =$ 米国の貨幣供給に対する日本の貨幣供給の比の対数
貨幣供給は季節修正済み M_1. 米国の貨幣供給はCITI BASEデータから採用し, 日本の貨幣供給データは日本銀行「経済統計月報」から採用した.

$y - y^* =$ 米国実質所得に対する日本の実質所得の比の対数
季節修正済みの製造業生産指数を実質所得に対して用いた：米国データはCITI BASEデータ, 日本データは日本銀行「経済統計月報」から採用した.

$r - r^* =$ 短期利子率の差の対数
r はコールレート, r^* は米国商務省債権利率でそれぞれCITI BASEデータと日本銀行「経済統計月報」から採用した.

$p - p^* =$ 米国卸売物価指数に対する日本の卸売物価指数の比の対数, 季節修正済み系列でそれぞれCITI BASEデータと日本銀行「経済統計月報」

から採用した.

$f =$ 外国通貨準備の対数, 季節修正済み系列で日本銀行「経済統計月報」から採用した.

参考文献

1. Ballie, R.T. and D.D. Selover (1987) "Cointegration and model of exchange rate determination," *International Journal of Forecasting,* 3, 43–51

2. Ballie, R.T. and T. Bollerslev (1990) " A multivariate generalized ARCH approach to modeling risk premia in forward foreign exchange rate markets," *Journal of International Money and Finance*, 9, 309–324

3. Bedrick, E.J. and J.R. Hill (1992) "A generalized bootstrap," In *Exploring the Limits of Bootstrap*, R. LePage and L. Billard (eds.), pp.319–326. John Wiley and Sons, New York

4. Bleaney, M. (1992) "A test of long-run purchasing power parity, using annual data for seven countries, 1900–88", *Economia Internazionale*, 45, 180–196

5. Bilson, J.F.O. (1978) "Rational expectations and the exchange rates," In *The Economics of Exchange Rates*,J. and H. Johnson (eds.), Addison-Wesley, Reading, Mass

6. Bilson, J.F.O. (1979) "The Deutsch mark/dollar rate – a monetary analysis," In *Policies for Employment, Prices and Exchange Rates*, Brunner, K. and A.H. Meltzer (eds.), North-Holland, Amsterdam

7. Booth, G.G. and M. Chowdhury (1992) "Canadian foreign exchange policies: intervention, control, cointegration," *Weltwirtschaftliches Archiv*, b128, h2, 21–33

8. Boothe, P. and d. Glassman (1987) " The statistical distribution of exchange rates," *Journal of International Economics*, 22, 297–319

9. Box, G.E.P. and G.C. Tiao (1973) *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, Mass

10. Cifarelli, G. (1992) "Exchange rate market efficiency tests and cointegration analysis," *Economia Internazionale*, 45, 197–208

11. Cochrane, J.H. (1988) "How big is the random walk in GNP?", *Journal of Political Economy*, 96, 893–920

12. Copeland, L.S. (1991) "Cointegration tests with daily exchange data," *Oxford Bulletin of Economics and Statistics*, 53, 2, 185–198

13. DeJong, D.N. and C.H. Whiteman (1991a) "Reconsidering trends and random walks in macroeconomic time series', " *Journal of Monetary Economics*, 29, 221–254

14. DeJong, D.N. and C.H. Whiteman (1991b) "The case for trend - stationarity is stronger than we thought," *Journal of Applied Econometrics*, 6, 413–421

15. Diebold, F.X. (1988) "Empirical modelling of exchange rate dynamics," Spriger Verlag, New York, NY

16. Diebold, F.X. and M. Nerlove (1989) "The dynamics of exchange rate volatility: a multivariate latent factor ARCH model," *Journal of Applied Econometrics*, 4, 1–21

17. Dickey, D.A. and W.A. Fuller (1981) "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, 74, 427–431

18. Dornbusch, R. (1976) "Expectations and exchange rate dynamics," *Journal of Political Economy*, 84, 1161–1176

19. Dornbusch, R. (1983) "Exchange rate economics: Where do we stand ?" *Economic Interdependence and Flexible Exchange Rates*, Bhandari,J.S. and B.H. Putnam (eds.), MIT Press, Cambridge Mass.

20. Dornbusch, R. and J.A. Frankel (1987) "The flexible exchange rate system: Experience and alternatives," NBER Working Paper 2464

21. Edison, H.J. and E.O. Fisher (1988), "A long-run view of the European Monetary System," International finance discussion paper XXX(Federal Researve Board, Washington, D.C.)

22. Enders, W. (1988) "ARIMA and cointegration tests of PPP under fixed and flexible exchange rate regimes," *Review of Economics and Statistics*, 55, 251–276

23. Frankel, J.A. (1979) "On the mark: a theory of floating exchange rates based on real interest differentials," *American Economic Review*, 69, 610-622

24. Frankel, J.A. (1983) "Monetary and portfolio-balance models of exchange rate determination," *Economic Interdependence and Flexible Exchange Rates*, Bhandari, J.S. and B.H. Putnam (eds.), MIT Press, Cambridge Mass.

25. Frankel, J.A. (1993) *On Exchange Rates*, MIT Press, Cambridge, Mass.

26. Frankel, J.A. and R. Meese (1987) "Are exchange rates excessively variable?", NBER Macroeconomic Annual

27. Fukao, M. (1988) 「金融の国際化が為替レートの変動および国際収支に与える影響について」(Influences of liberalization of financial markets on the foreign exchange rates and on balances of payments), 金融研究, 7, 1-42

28. Fuller, W.A. (1976) *Introduction to Statistical Time Series*, John Wiley and Sons, New York

29. Hakkio, C.S. (1986) "Does the exchange rate follow a random walk ? A Monte Carlo study of four tests for a random walk," *Journal of International Money and Finance* 5, 221–229

30. Hill, J.R. (1990) "A general framework for model-based statistics," *Biometrika*, 77, 115–126

31. Hooper, P. and J. Morton (1982) "Fluctuations in the dollar: a model of nominal and real exchange rate determination," *Journal of International Money and Finance*, 1, 39-56.

32. Huizinga, J. (1987) "An empirical investigation of the long run behavior of real exchange rate," *Carnegie-Rochester Conference Series on Public Policy*, 27, 149–214.

33. Isard, P. (1988) "Exchange rate modeling: an assessment of alternative approaches," *Empirical Macroeconomics for Interdependent Economies*, Bryant, R.C., D.W. Henderson, G. Holthar, P. Hooper, and S.A. Symansky (eds.),1, 183–201 Brookings Institution, Washington D.C.

34. Jeffreys, H. (1961) *Theory of Probability*, 3rd edition, Oxford University Press, London

35. Johansen, J. (1988) "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control*, 12, 231–254

36. Johansen, J. (1991) "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models," *Econometrica*, 59, 1551–1580

37. Kaminsky, G. (1988) "The real exchange rate in the short and in the long run," Working Paper 88–11, University of Calfornia, San Diego, CA

38. Kariya, T. and K. Fukao (1988) **合理的予想形成によるインフレーション–為替分析** (Rational Expectation Analysis of foreign exchange rates), 有斐閣, 東京

39. Kariya, T. and Y. Matsue (1988)「非線形分散変動モデルによる日次と週次の為替レート変動分析」(Analysis of daily and weekly foreign exchange

rates within the framework of nonlinear heteroscedastic variance models), *mimeograph*

40. Kroner, K.F. and W.D. Lastrapes (1993) "The impact of exchange rate volatility on international trade: reduced form estimates using the GARCH-in-mean model," *Journal of International Money and Finance*, 12, 298–318

41. Kunitomo, N. (1992) "Tests of unit roots hypotheses in econometric models," *mimeograph*

42. Lastrapes, W.D. (1989) " Exchange rate volatility and U.S. monetary policy: An ARCH application," *Journal of Money, Credit, and Banking*, 31. 66–77

43. Levich, R.M. (1985) "Empirical studies of foreign exchange rates: price behavior, rate determination and market efficiency," *Handbook of International Economics II*, Jones, R.W. and P.B. Kenen (eds.), North-Holland, Amsterdam

44. Maddala, G.S. (1992) *Introduction to Econometrics*, 2nd edition, Macmillan, New York

45. McNown, R. and M.S. Wallace (1992) "Cointegration tests of a long-run relation between money demand and the effective exchange rate," *Journal of International Money and Finance*, 11, 107–114

46. Meese, R.A. and K. Rogoff (1983a), "Empirical exchange rate models for the seventies: Do they fit out-of-sample?," *Journal of International Economics*, 14,3–24

47. Meese, R.A. and K. Rogoff (1983b), "The out-of-sample failure of empirical exchange rate models: Sampling error or misspecification?", in J.A. , ed., *Exchange Rates and International Macroeconomics*, University of Chicago Press, Chicago, IL.

48. Meese, R.A. and K. Rogoff (1988), "Was it real? The exchange rate-interest differencial relation over the modern floating rate period," *Journal of Finance,* 43, 933–948

49. Mishkin, F. (1992) *The Economics of Money, Banking and Financial Markets,* 3rd edition, Harper Collins Publication

50. Morris, C. (1986) "Discussion of paper by B. Efron," *American Statistician*, 40, 7–8

51. Mussa, A. (1979) "Empirical regularities in the behavior of exchange rates and theories of the foreign exchange rates," *Carnegie-Rochester Conference Series on Public Policy*, Autumn, 11, 9a–57

52. Nelson, C.R. and C.I. Plosser (1982) "Trends and random walks in macroeconomic time series: some evidence and implications," *Journal of Monetary Economics*, 10, 139–162

53. Phillips, P.C.B. (1987) "Time series regression with a unit root," *Econometrica*, 55, 277–302

54. Phillips, P.C.B. (1991a) "To criticize the critics: an objective Bayesian analysis of stochastic trends," *Journal of Applied Econometrics*, 6, 333–364

55. Phillips, P.C.B. (1991b) "Bayesian routes and unit roots: De rebus prioribus semper est disputandum," *Journal of Applied Econometrics*, 6, 413–421

56. Phillips, P.C.B. and S. Ouliaris (1990) "Asymptotic properties of residual based tests for cointegration," *Econometrica*, 58, 165–193

57. Phillips, P.C.B. and P. Perron (1988) "Testing for a unit root in time series regression," *Biometrika*, 75, 335–346

58. Pippenger, M.K. (1993) "Cointegration tests of purchasing power parity: the case of Swiss exchange rates," *Journal of International Money and Finance,* 12, 46–61

59. Rubin, D.B. (1981) "The Bayesian bootstrap," *Annals of Statistics*, 9, 130–134

60. Rubin, D.B. (1984) "Bayesianly justifiable and relevant freqency calculations for the applied statistician," *Annals of Statistics*, 12, 1151–1172

61. Said, E.S. (1991) "Unit root test for time-series data with a linear time trend," *Journal of Econometrics*, 47, 285–304

62. Sims, C.A. (1988) "Bayesian skepticism on unit root econometrics," *Journal of Economic Dynamics and Control,* 12, 463–474

63. Sims, C.A. and H. Uhlig (1991) "Understanding unit rooters: a helicopter tour," *Econometrica*, 59, 1591–1599

64. Stock, J.H. and M.W. Watson (1988) "Testing for common trends," *Journal of American Statistical Association*, 83, 1097–1107

65. So, J.C. (1987) "The sub-Gaussian distribution of currency futures: stable Paretian or nonstationary ? " *Review of Economics and Statistics*, 69, 100–107

66. Taylor, S.J. (1986) *Modeling Financial Time Series,* John Wiley and Sons, New York

67. Tsurumi, H. and C.L. Chen (1994) "Estimating unknown join points: determination of the yen-dollar exchange rate," *Financial Engineering and the Japanese Markets*, forthcoming

68. Tsurumi, H. and H. Wago (1996) "A Bayesian analysis of unit root and cointegration with an application to a Yen-Dollar exchange rate model," *Advances in Econometrics,* Vol 11-B, 51–86.

69. Westerfield, J.M. (1977) "An examination of foreign exchange risk under fixed and floating rate regimes," *Journal of International Economics*, 7, 181–200

70. Wolff, C.C.P. (1987) "Time-varying parameters and the out-of-sample forecasting performance of structural exchange models," *Journal of Business and Economic Statistics*, 5, 87-97

71. Zangari, P. and H. Tsurumi (1994) "A Bayesian analysis of censored autocorrelated data on exports of Japanese passenger cars to the U.S.," mimeograph

72. 和合 肇 (1996) 「経済時系列における単位根のベイズ分析」貯蓄経済理論研究会年報第 11 巻, 101–119

73. 和合 肇 (1997) 「経済時系列における共和分のベイズ検定」貯蓄経済理論研究会年報第 12 巻, 91–107

第5章　エネルギー価格の変動と産業調整

5.1　はじめに

1973年以降の2度にわたる石油危機において, 原油価格の大幅な高騰を契機として, 我が国に限らず世界的に産業の調整が行われた. マクロ的には, 日本経済はこの局面に対してはうまく調整を行ったといえる. しかし, 個々の産業がこの状況に対してとった調整過程を見ると, 産業ごとに異なっており, 必ずしもすべての産業でスムーズに調整がなされたわけではない. そこで, 特にエネルギー多消費産業において, 石油危機に対してどのような調整がなされたか, そしてエネルギー相対価格の大幅な変化に対して, 生産要素間の代替がどのように行われたかを分析することは, 今後の政策決定のためにも有用であると考える.

石油危機以前においては, エネルギー, 輸入原材料等の資源価格は安定的に推移しており, また生産要素の相対価格の変化は, もっぱら賃金上昇に起因していたために, 生産要素間の代替は資本と労働の間で顕著に観察されていた. これに対して, 石油危機における石油を中心としてエネルギー価格の大幅な上昇は, エネルギーの相対価格を大幅に変化させ, 省エネルギー投資に見られるように, エネルギーと資本の代替を促進させた. このような関係を計量的に分析するための生産関数として従来用いられてきた, 付加価値をアウトプットとするコブ・ダグラス型のものは, 石油危機以後当てはまりが悪く, また得られた係数の推定値も不適切であるような欠陥を示すようになった. この理由として, 付加価値をアウトプットとした場合, 資本と労働以外の生産要素を無視している点, さらにエネルギー等の投入要素を組み入れた関数においても, 生産要素の代替弾力性を1とする関数上の制約が現実に当てはまらなくなった点が指摘できる.

石油危機以降の生産構造を描く生産関数として, 最近実証分析で多く適用される関数型にトランスログ型生産関数がある. トランスログ型の生産関数は, 資

本や労働以外の生産要素を組み込むことが可能であり, またそれらの要素間の偏代替弾力性を容易に計測しうること, そしてコブ・ダグラス型をその特殊なケースとして含むことから, このような分析に適していると考えられる.

トランスログ型の生産(コスト)関数を用いてエネルギーの相対価格の上昇による要素間代替を分析した例は, Berndt and Wood (1975) をはじめ, アメリカ, カナダを中心として数多くなされている. わが国における分析例としては, 伊藤・松井 (1978), 伊藤 (1983), 和合・鈴木 (1980, 1981), Kuroda and Wago (1982), 和合 (1983), 中島 (1983), 大山 (1983) 等があげられる. これらの分析に対して, 次のような問題点を指摘することができる.

1. 特に全産業あるいは製造業全体を分析対象とした場合の, 集計の問題
2. 年データを使用した場合の, 情報のロスと精度の低下
3. コスト関数に基づく推定の場合の, 価格調整が即座に行われるという仮定
4. 多くの分析で, 原材料は残差として計算されている点

もう一つの重要な問題点として構造変化の取り扱いがあげられる. 分析対象としている期間は, 石油危機をはさんで生産構造に大きな変化があったということを, 多くの分析が示唆している. このようなデータを従来のような固定パラメータに基づく推定法で分析するのには問題がある. ところで, 構造変化については最近いろいろな提案がなされている. たとえば, ある時点で変化があったことが分かっている場合には, 観測期間を分割して推定したり, ダミー変数によって両期間におけるパラメータの違いを分析することができる. もう一つの考え方は, 変化時点も他のパラメータと同時に推定する方法である. われわれがここで考えるは後者の考え方によるものであって, 特にある期間の生産構造から次の期間のそれへなめらかに変化すると仮定するベイズモデルである. ベイズモデルを用いる利点の一つは, 限られた標本データしか得られないときに, 従来の点推定値が十分な情報をを含んでいない場合があるのに対して, 事後分布全体を見て分析できる点にある.

本章では, 以上の点を考慮してエネルギー多消費産業の中から次節に述べる理由により, 鉄鋼業と紙パルプ産業を採りあげ, , エネルギー価格の変動に対す

る両産業の調整の違いと生産構造の変化について分析することにする.

本章は次のように構成される. 第2節では, 産業レベルの省エネルギーについて述べ, 分析に用いたデータの作成方法を示す. 第3節では, モデルに用いたトランスログ型の生産・コスト関数について簡単に紹介し, 方程式間の確率制約のもとで多変量回帰モデルにおける構造変化点とその調整スピードを見付けるためのベイズ推定法について述べる. 第4節では, 鉄鋼業と紙・パルプ産業に対して適用した計測結果について分析し, 最後に, 第5節でまとめを述べる.

5.2 エネルギー多消費産業とデータ

5.2.1 エネルギー多消費産業における省エネルギー

表4.1は, 生産額に占める燃料使用額の割合を示したものである. この表からも明らかなように, 紙・パルプ, 化学, 窯業・土石(セメント), 鉄鋼の各産業においてエネルギーシェアが高い. これら4つの代表的なエネルギー多消費産業において, エネルギー価格の上昇に対する対応を各生産要素の変化として計測するとともに, 産業の形態の相違を比較検討することは重要である. この4つの産

表 5.1: 製造業における生産額に占める燃料使用額(%)

産業	1974年	1979年
食料品	0.92	1.06
繊維	1.18	1.69
衣服その他繊維製品	0.55	0.65
木材・木製品	0.93	0.91
家具・装備品	0.52	0.48
紙・パルプ	3.12	4.71
出版・印刷	0.34	0.34
化学工業	3.70	3.44
石油・石炭	3.77	3.07
ゴム製品	1.04	1.15
皮革製品	0.49	0.44
窯業・土石製品	6.64	8.42
(セメント)	7.41	10.15
鉄鋼	4.57	3.82
非鉄金属	1.39	1.85
金属製品	0.77	0.75
一般機械	0.44	0.42
電気機械器具	0.33	0.27
輸送用機械器具	0.46	0.41
精密機械	0.29	0.24
武器	0.42	0.57
その他	0.53	0.63
平均	1.82	1.77

資料:工業統計表・産業編

業の中で, セメント工業が最も代表的なエネルギー多消費産業といえる. 事実, この産業では石油危機以降エネルギー・コストと生産設備の代替という形で省エネルギーが進展した. セメント工業の省エネルギーは, そのエネルギーの約8

割を消費している焼成行程の設備の改良と新鋭化に観察される．石油危機以降，NSPキルンへの転換が進み，それとともにエネルギー原単位は低下している．ところで，NSPキルンは脱硝効果等の環境対策の点でも優れており，第3次環境規制により，1977年6月以降の新設キルンはNSPキルンに限られていることから，エネルギー・コストの上昇に伴うNSP化による省エネルギー分を計数的に把握するのは難しい．

また，化学工業では最終製品が多岐にわたり，製品構成のいかんによってプロセス全体のエネルギー消費量が大きく異なるので，最終需要の変化がデータに反映されることになり，望ましい結果が期待できない．

以上の理由から，エネルギー多消費産業のうち上記の2つの産業を分析対象から除き，鉄鋼業と紙・パルプ産業を選択した．鉄鋼業は第1次石油危機以降，石炭から石油への転換，廃熱利用による発電，副生ガスの利用等の省エネルギーが進展したのに対して，紙・パルプ産業においては，第2次石油危機以降まで顕著なエネルギー節約が図られなかったという好対照を示している．

5.2.2 鉄鋼業と紙パルプ産業における省エネルギー

第1次石油危機以降，省エネルギーと脱石油の効果的な達成を成し遂げた産業として，鉄鋼業は代表的な産業といえる．鉄鋼の製造工程は，銑鉄部門，製鋼部門，圧延部門の3つに大きく分けられるが，大半のエネルギーが銑鉄部門で消費されている．鉄鋼業における省エネルギーは，歩留まりの向上や投下エネルギーの減少等の使用エネルギー原単位の改善，および投下エネルギーの排熱や副生ガスの回収という形で実施されている．

石油危機以前の省エネルギー対策は，エネルギー投下量の減少を主眼としたことから，主としてエネルギー消費構成の高い銑鉄部門において講じられていた．その対策は主に労働者，管理者の熟練に依存したもので，いわばエネルギーと労働の質の代替の進展が省エネルギーに寄与したといえる．

一方，石油危機以後では，既存設備，工程改良のための付加設備，および新鋭設備の導入が活発になり，省エネルギーに結びつく技術が積極的に採用され，いわば高額な設備とエネルギーとの代替を促進させたといえる．その他，目立ったエネルギー対策として，高炉への吹き込み重油のコークスへの転換があげられる．

石油危機以後の省エネルギー効果に関して, 紙・パルプ産業は鉄鋼業と著しい対照を示している. 73年以降78年頃までエネルギー原単位(燃料, 購入電力)はほとんど改善していない. 紙・パルプの製造工程で消費されるエネルギーとしては, 重油, 石炭等の燃料および電力が外部から購入されるが, そのほかに製造工程の副産物であるパルプ廃液やバーク等が燃料として利用される. 紙・パルプ産業における省エネルギー対策をあげれば, パルプ廃液, バーク, および木材くずを燃料として有効利用, ボイラーやタービンの効率的な運転, 発生蒸気による自家発電など設備投資額の少ないものが中心だったといえる. 蒸解工程における技術開発には多額の設備投資を要するが, 省エネルギー効果の高い対策が業界全体で進展しなかったことが, 鉄鋼業における省エネルギーとの大きな差異をもたらしたといえる.

5.2.3 データの作成方法

今回の分析に用いたデータは, 資本以外は原則として鉄鋼業については「鉄鋼統計年報」から, また紙・パルプ産業については「紙パルプ統計年報」から, それぞれ月次ベースで原データを取り, 四半期データに集計した後, センサス局のX-11により季節調整を行った. しかし, いくつかの系列については上記の資料からはとれないこと, およびシェアデータとしての整合性の問題から, 上記以外の資料からとったり, 計算式から求めたものがある. 以下この2つの産業のデータの作成方法について, その概要を記す. なお, 原データの出所および単位については付録の表を, また推計方法の詳細についてはデータ作成マニュアルを参照されたい[1].

資本のデータについては, 公表されている資本ストックのデータが十分とはいえないことから, 新たに推計し直した. 方法については付録で説明した.

モデルの計測は生産関数とコスト関数の両方について行ったので, 資本, 労働, エネルギー, 原材料, その他生産要素に関する数量, 価格および価額のデータが必要になる. 資本については, 別に付録で説明を行ったので, その他のデータについての概要を述べる. 原則として数量と価格を原データとして作成し, 価額は

[1] 文部省科学研究費エネルギー特別研究, ディスカッション・ペーパー(1982,1983)参照.

両者の積として求めた. ただし, その他生産要素の価額は

(生産額 – 資本費用 – 労働費用 – エネルギー費用 – 主要原材料費用)

として求め, 数量データは上記の計算式にそれぞれの項目の75年基準の実質値を代入して計算した実質値系列として求めた. 価格データは, このようにして求めた価額データを数量データで割って, インプリシットに求めた. 両産業において, その他生産要素のシェアは30~40%を占めているが, 通常よく行われているように, これを資本や原材料の項目に入れてしまうのは適当でないと考える.

(A) 数量データ

鉄鋼産業 生産量は鉄鋼一次製品の段階で把握した. 普通鋼圧延鋼材生産量と特殊鋼圧延鋼材生産量をDivisia集計して求めた. 労働投入量は生産動態統計の月間延べ人員に毎月勤労統計の実労働時間数を乗じて人・時間として算出した. 原材料投入量は, 鉄鋼統計年報よりとった鉄鉱石消費量, 砂鉄消費量, 溶解用鉄くずの純消費量 (屑消費量–屑発生量) をDivisia集計して求めた. なお, コークスは近年の消費実態からエネルギー部門で処理した. エネルギー投入量は鉄鋼統計年報よりとった電力買電量, 重油消費量, コークス消費量をカロリー換算し, Divisia集計して求めた.

紙・パルプ産業 生産量は, 紙と板紙の生産量である. 労働投入量は鉄鋼業と同様の方法で求めた. 原材料投入量は, 紙パルプ統計年鑑よりとった原料製紙パルプ輸入量, 故紙, 故紙パルプ, 原木, チップをDivisia集計して求めた. エネルギー投入量は, 買電消費量, 重油消費量, 石炭消費量をエネルギー・マトリックスを用いてカロリー換算し, Divisia集計した.

(B) 価格データ

価格は原則として, 物価指数年報によった. ただし, 賃金は毎月勤労統計の賃金指数を利用した. また, 紙・板紙価格は紙・パルプ統計年報の出荷金額を出荷量で割って求めた. なお, 重油価格はそれぞれの産業で, 油種別に加重平均して求めている.

図4.1は鉄鋼業における資本, 労働, エネルギー, 原材料, その他生産要素について, 1961年第1四半期から1989年第4四半期までの四半期データを, 1961年

第2期を100とする(a)価格指数と(b)数量指数としてプロットしたものである. データを見ると, 数量指数では, 資本がほぼ一定の割合で上昇し, 労働は僅かながら低下傾向にあるもののほぼ横這い状態である. エネルギーと原材料の動きは似ており, 石油危機以降は低下の以降にある. また価格指数は, 労働だけがトレンド的に上昇し, 石油危機以後は変動が大きく, 上昇率は低下している. 資本は平均が低下し, エネルギーは上昇しているのが見られる.

図4.2は紙・パルプ産業における(a)価格指数と(b)数量指数を同様にプロットしたものである. 価格指数には, 資本と労働に石油危機の一時的影響がはっきり表れており, エネルギー価格指数はそれ以後上昇しているのが分かる. 数量指数は労働以外は石油危機まで同じ強い上昇傾向を示すが, それ以後エネルギーは横這い, また原材料は大きな落ち込みを見せている.

5.3 モデルの定式化と推定

5.3.1 トランスログ型生産・コスト関数

本節では, モデルに用いたトランスログ型の関数について簡単に説明する. トランスログ型の生産(コスト)関数のいろいろな性質については, 既に他の文献[2)]でも述べられているので, ここでは定式化のみを示すにとどめる.

トランスログ型の生産関数は, Christensen, Jorgenson and Lau (1973) で最初に提案されたが, それはi) Allen の偏代替弾力性に a priori な制約をおかない, ii) 任意の生産関数への2次までの近似として考えられるもの, という2つの条件を満足する一般的なフレキシブル関数の中の1つであると考えることができる.

投入要素として資本(X_K), 労働(X_L), エネルギー(X_E), 原材料(X_M), その他生産要素(X_Z)を考え, 効率的な生産活動により達成できる産出量を Y としたときに, 両者を結びつける次のような生産関数を仮定する.

$$Y = f(X_K, X_L, X_E, X_M, X_Z) \tag{5.1}$$

(5.1)の生産関数に対応して, ある産出量を達成するための最小コスト G と, 各要素価格 W_K, W_L, W_E, W_M, W_Z を関係づける双対なコスト関数が得られる. さ

[2)] たとえば, 鈴木 (1982), 和合 (1983a,b) を参照.

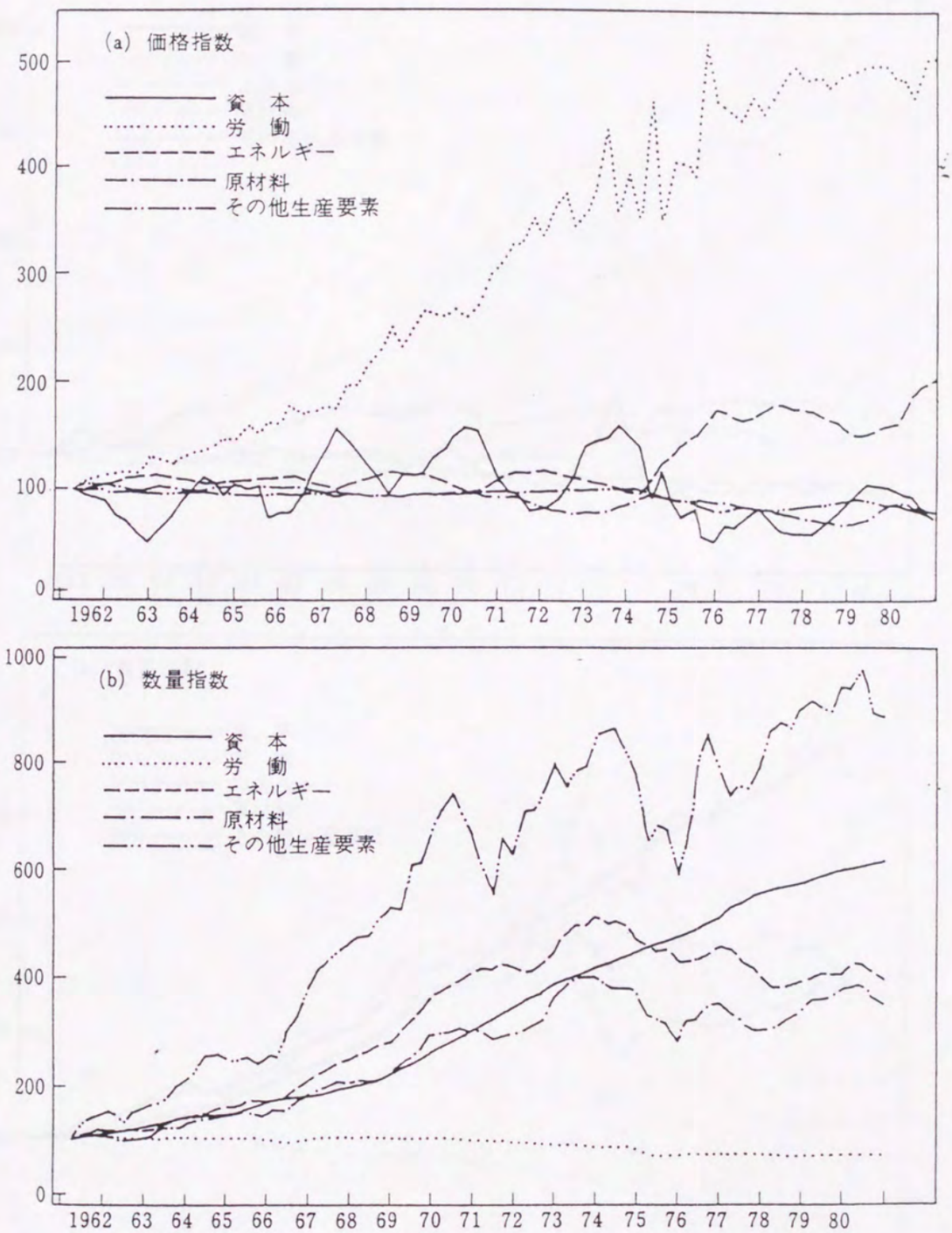


図 5.1: 鉄鋼産業のデータ：価格指数と数量指数

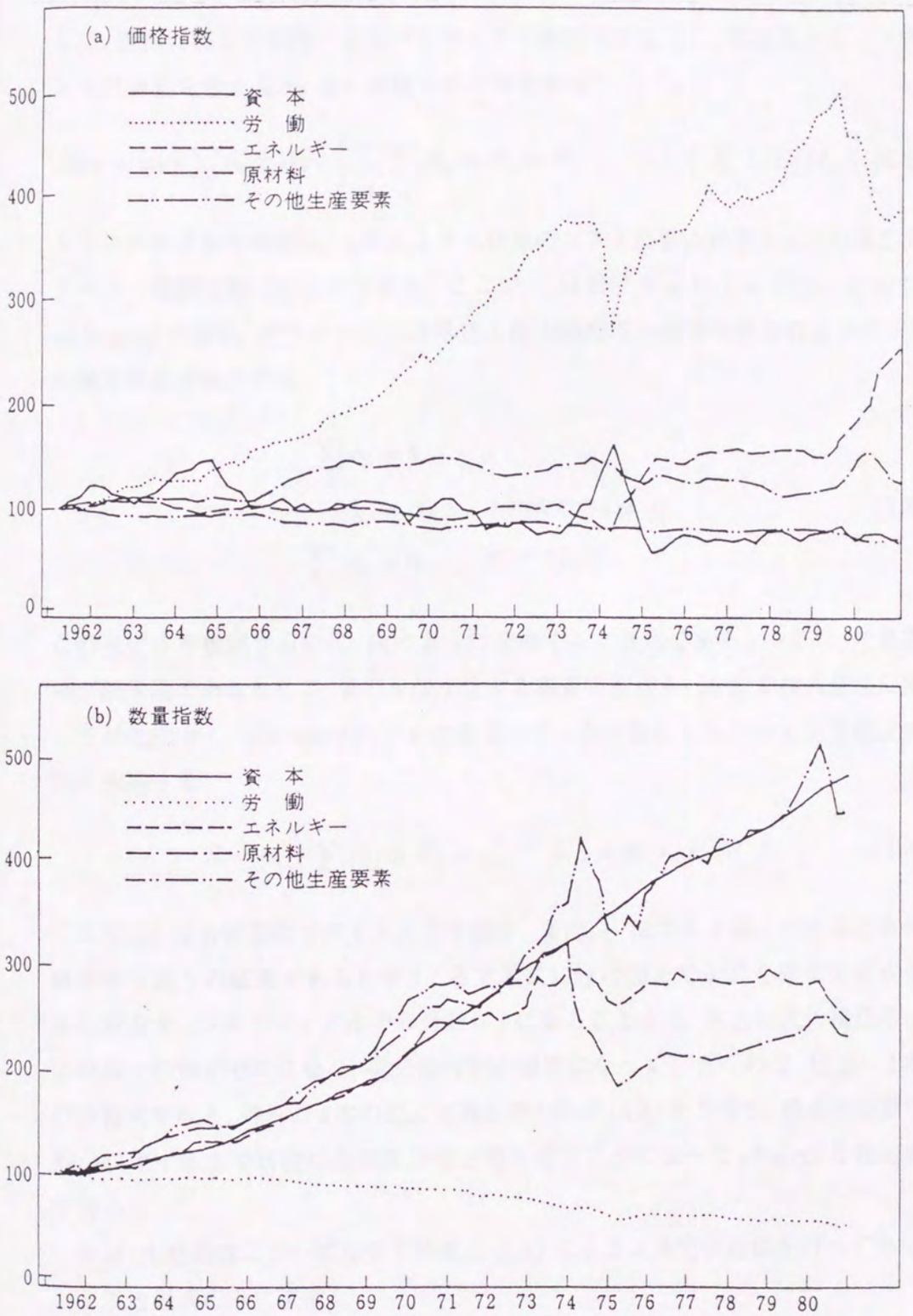


図 5.2: 紙・パルプ産業のデータ：価格指数と数量指数

らに, 規模に関して収穫一定とパラメータの双対性を仮定し, 関数型としてトランスログ型を考えると, 次の単位コスト関数を得る.

$$\ln c = \alpha_0 + \sum_i \alpha_i \ln W_i + \frac{1}{2} \sum_i \sum_j \beta_{ij} \ln W_i \ln W_j, \qquad i, j = K, L, E, M, Z \tag{5.2}$$

トランスログ型の関数は, 上記のように任意のコスト関数の対数を2次の項までテーラー展開近似したものである. ここで, c は単位コスト, $c = G/Y$, を表す. adding-upの制約, パラメータの対称性と投入価格の一次同次性の仮定から, 次の線形制約がおかれる.

$$\begin{aligned} &\sum_i \alpha_i = 1 \\ &\beta_{ij} = \beta_{ji} \qquad (\text{すべての} i \neq j) \\ &\sum_j \beta_{ij} = 0 \qquad (\text{すべての} i) \end{aligned} \tag{5.3}$$

このモデルを推定するのに, 次のように変換すると便利である. つまり, 要素市場が競争的であるとして, 費用を最小化する需要方程式を, (5.2) を投入価格に関して対数微分し, Shephardのレンマを用いて, 次の要素コストシェア方程式の形に変換する.

$$S_i = \alpha_i + \sum_i \beta_{ij} \ln W_j + \epsilon_i, \qquad i, j = K, L, E, M, Z \tag{5.4}$$

ここで, S_i は各要素のコストシェアを表す. また, ϵ_i はコスト最小化するときの確率的な誤りの結果であると考え, 各方程式に付け加えて上式を確率方程式体系と見なす. コストシェアはその合計が1になることから, 各方程式の攪乱項は各時点での和がゼロになり, 共分散行列が特異になってしまうので, 任意に1本の方程式を除き, 残りの4本の式に方程式間の制約 (5.3) をつけて, 残差共分散行列が収束するまで制約付多変量回帰を繰り返すことによってefficientな推定値を得る.

なお, 生産関数についても全く同様に (5.1) にトランスログ近似を行ってから, シェア方程式の形に変換し

$$S_i = \gamma_i + \sum_i \delta_{ij} \ln W_j + \epsilon_i, \qquad i, j = K, L, E, M, Z \tag{5.5}$$

次の制約をつけて推定を行う.

$$\begin{aligned} &\sum_i \gamma_i = 1 \\ &\delta_{ij} = \delta_{ji} \qquad (\text{すべての} i \neq j) \\ &\sum_j \delta_{ij} = 0 \qquad (\text{すべての} i) \end{aligned} \tag{5.6}$$

この関数型の問題点として, トランスログ型の関数は真の生産関数あるいはコスト関数の近似であるために, 一般には自己双対にはならない. したがって, 生産関数とコスト関数のどちらを選択するかに関して明確な基準があるわけではなくて, 分析目的によることになる. また, トランスログ型の関数は以前に指摘したような性質を持つ唯一のフレキシブル関数ではない. 他に, 一般化レオンティエフ, 一般化ボックス・コックスなどの関数型も含まれるので, モデル間の選択が必要になる. これらの関数型はいずれも真の関数に対するTaylor近似であるが, 最近この他にBarnett (1983) によるLaurent近似や, これらの2つの関数型の欠点を改良したFourierフレキシブル関数がGallant(1981, 1982) によって提案されている[3].

もう一つの問題点は誤差項の確率分布についての仮定である. シェアは必ず0と1の間になければならないことから, 通常仮定されるような多変量正規分布の場合には, この制約が満たされないことが起こりうる. このためにシェア方程式におけるこの条件を考慮した定式化がいくつか提案されている[4].

5.3.2 確率制約下における構造変化モデル

次の多変量回帰モデルを考える.

$$Y_i = X_i\beta_i + \epsilon_i, \qquad i = 1, \ldots, m \tag{5.7}$$

ここで, Y_i は第 i 方程式の $n \times 1$ の従属変数の観測値ベクトル, X_i は $n \times k_i$ の説明変数行列, ϵ_i は第 i 方程式の $n \times 1$ 誤差項ベクトル, そして β_i は $k_i \times 1$ の

[3] Taylor級数やLaurent級数展開は, ある解析的な関数を一様に近似することはできるが, その導関数を近似することはできない. 関数近似の近似誤差をある範囲内に押さえても, 導関数に対しては近似できない. ところが, 要素需要体系や代替弾力性は, 生産関数やコスト関数の導関数に基づいている. そこで, この欠点を補うために Gallant はフーリエ級数展開を使って, Sobolov ノルムの意味で一様に近似する方法を提案した.

[4] たとえば, Dirichlet分布を用いたWoodland(1979) や, 対数正規分布を用いたRossi (1984) などがその例である. これらの問題についてのサーベイが和合(1983b) にある.

未知係数ベクトルである. (5.7) では, 係数 $\beta_i = (\beta_{1i}, \ldots, \beta_{k_i i})'$ は, ある状態から次の状態へ t^* という時点ですぐに移行するのではなく, ある一定時間の間に徐々に移行すると仮定する. ここで, 変化点 t^* と調整スピード η を2つの未知パラメータとし, Bacon and Watts (1971) と Tsurumi (1980) で提案された遷移関数を用いて, (5.7) を次のように書き直す.

$$Y_{ti} = X_{ti}\beta_i + \mathrm{trn}(S_t/\eta)X_{ti}\delta_i + \epsilon_{ti}, \qquad i = 1, \ldots, m, \qquad t = 1, \ldots, n \tag{5.8}$$

ここで, Y_{ti} は Y_i の t 番目の要素, X_{ti} は X_i の t 番目の行, そして δ_i は $k_i \times 1$ の未知係数ベクトルである. この遷移関数 $\mathrm{trn}(S_t/\eta)$ は, 次の性質を満足する.

$$\begin{aligned} \lim_{S_t \to \infty} \mathrm{trn}(S_t/\eta) &= 1 \\ \mathrm{trn}(0) &= 0 \\ \lim_{\eta \to 0} \mathrm{trn}(S_t/\eta) &= 1 \end{aligned} \tag{5.9}$$

そして S_t は

$$S_t = \begin{cases} 0 & t \le t^* \text{のとき} \\ t - t^* & t > t^* \text{のとき} \end{cases}$$

である. 遷移関数の2つの未知パラメータ t^* と η は, すべての回帰係数 δ_i に対して同一であると仮定する. この多少きつい仮定は, 回帰係数 β_i と δ_i が方程式間の線形制約に従っているためであり, そのためにたぶん同じ時点でシフトを始め, 同じ調整スピードでシフトするであろうと考えるからである.

(5.8) を行列の形で次のように書き直す.

$$Y = X\beta + \epsilon \tag{5.10}$$

ここで

$$\underset{(mn\times 1)}{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{bmatrix}, \qquad \underset{\substack{(mn\times K) \\ K=2\sum_{i=1}^m k_i}}{X} = \begin{bmatrix} (X_1 \cdot \mathrm{trn}X_1) & 0 & \cdots & 0 \\ 0 & (X_2 \cdot \mathrm{trn}X_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (X_m \cdot \mathrm{trn}X_m) \end{bmatrix}$$

$$\underset{(K\times 1)}{\theta} = \begin{bmatrix} \beta_1 \\ \delta_1 \\ \vdots \\ \beta_m \\ \delta_m \end{bmatrix}, \qquad \underset{(nm\times 1)}{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{bmatrix}$$

trn は $\mathrm{trn}(S_t/\eta)$ の略称である.

次に θ に対する確率線形制約は, 階層的 (hierarchical) な情報表現または混合的 (mixed) な情報表現の形で表すことができる. Lindley and Smith(1972) と Smith(1973) による階層的な情報表現は, 次のように表すことができる.

$$\theta = Q\gamma + u \tag{5.11}$$

ここで, γ は $q\times 1$ ベクトル, Q は $k\times q$ 行列でそのランクは q, そして u は $K\times 1$ の確率ベクトルである. また, Theil and Goldberger (1961) と Theil(1963) による混合的な情報表現は次のように表す.

$$r = R\theta + v \tag{5.12}$$

ここで, r は $p\times 1$ ベクトル, R は $p\times K$ 行列でそのランクは p, そして v は $p\times 1$ の確率ベクトルである.

最初に階層構造の場合を考える. (5.11) の u は $u\sim N(0,\sigma^2 I)$ であり, (5.10) の ϵ は $\epsilon\sim N(0,\Sigma)$, ここで $\Sigma=\Omega\otimes I_n$ で, Ω は $m\times m$ の正値定符号行列である. 次の事前密度

$$P(\theta,\gamma,t^*,\eta,\Omega)\propto\sigma^{-1}|\Sigma|^{-\frac{1}{2}} \tag{5.13}$$

と尤度関数

$$\ell(\theta,\gamma,t^*,\eta,\sigma,\Omega|data)\propto\sigma^{-(K/2+1)}|\Sigma|^{-(n+1)/2}$$
$$\exp\left[-\frac{1}{2\sigma^2}(\theta-Q\gamma)'(\theta-Q\gamma)-\frac{1}{2}(Y-X\theta)'\Sigma^{-1}(Y-X\theta)\right] \tag{5.14}$$

から, γ と θ に関して積分し, σ と Ω の条件の下で t^* と η に関する周辺事後密度を得る.

$$P(t^*,\eta|data,\sigma,\Omega)\propto|\sigma^{-2}M_Q+X'\Sigma^{-1}X|^{-\frac{1}{2}}\exp\left[\frac{1}{2}(Y'M_XY+d_1)\right] \tag{5.15}$$

ここで

$$M_X=\Sigma^{-1}-\Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}$$
$$d_1=\hat{\theta}X'\Sigma^{-1}X\hat{\theta}-c_1'(\sigma^{-2}M_Q+X'\Sigma^{-1}X)c_1$$
$$c_1=(\sigma^{-2}M_Q+X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$$
$$M_Q=I-Q(Q'Q)^{-1}Q'$$
$$\hat{\theta}=(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$$

次に, 混合表現の場合は $v \sim N(0, \sigma^2 I)$ と仮定して, (5.12) の事前情報を用いると, σ と Ω の条件の下での t^* と η に関する次の周辺事後密度を得る.

$$P(t^*, \eta | data, \sigma, \Omega) \propto |\sigma^{-2} R'R + X'\Sigma^{-1}X|^{-\frac{1}{2}} \exp\left[\frac{1}{2}(Y'M_X Y + d_2)\right] \quad (5.16)$$

ここで

$$d_2 = \widehat{\theta} X'\Sigma^{-1} X\widehat{\theta} - c_2'(\sigma^{-2}R'R + X'\Sigma^{-1}X)c_2$$
$$c_2 = (\sigma^{-2}R'R + X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$$

(5.16) は (5.12) の r をゼロベクトルとおいて導かれる.

(5.15) と (5.16) を比べると, 階層構造の場合も混合構造の場合も t^* と η に対する似たような周辺事後密度になり, したがって, 似かよった事後推定値になる. このことは θ に対して何の先験的情報ももっておらず, また (5.13) と同じ事前分布を考えた場合には正しい. しかし, もし θ に対して先見情報を持っている場合には, 階層的な表現の方がそのような情報を取り入れるのにずっと楽である. たとえば, (5.11) のパラメータ γ がわかっている場合には, (5.15) を導いたときのように γ について積分する必要はない. あるいは, $\theta > 0$ なら, これは単にトランケートした正規分布と考えられるので, 尤度関数に含めた方がよい. この応用として, たとえば, 関数型が凹性という条件についての事前情報を用いて推定することも可能である. しかし, 混合的な情報表現では $\theta > 0$ のような情報を取り入れようと思っても簡単ではない.

このベイズの方法を先のトランスログ型の関数に適用する. adding-up の制約は, その他生産要素のシェア式を除くことで含める. (5.8) の様に遷移関数を導入して, 各生産要素のシェアを表す (5.4) を次のように書き直す.

$$S_{ti} = \alpha_i + \sum_{j=1}^{5} \beta_{ji} \ln W_{tj} + \alpha_i' \mathrm{trn}(S_t/\eta) + \sum_{j=1}^{5} \beta_{5+j,i} \mathrm{trn}(S_t/\eta) \ln W_{tj} + \epsilon_{ti} \quad (5.17)$$

$$i = K, L, E, M, \quad t = 1, 2, \ldots, n$$

行列表示では

$$Y = X\theta + \epsilon \quad (5.18)$$

ここで

$$Y = (Y_K', Y_L', Y_E', Y_M'), \quad Y_i = (S_{1i}, \cdots, S_{ni})'$$

$$X = [I_4 \otimes (Z, \mathrm{trn} \cdot Z)]$$

$$Z = \begin{bmatrix} 1 & \ln W_{1K} & \cdots & \ln W_{1Z} \\ \vdots & \cdots & & \vdots \\ 1 & \ln W_{nK} & \cdots & \ln W_{nZ} \end{bmatrix}, \quad \mathrm{trn} \cdot Z = \begin{bmatrix} \mathrm{trn} & \mathrm{trn} \cdot \ln W_{1K} & \cdots & \mathrm{trn} \cdot \ln W_{1Z} \\ \vdots & \cdots & & \vdots \\ \mathrm{trn} & \mathrm{trn} \cdot \ln W_{nK} & \cdots & \mathrm{trn} \cdot \ln W_{nZ} \end{bmatrix}$$

$$\theta = \begin{bmatrix} \theta_K \\ \theta_L \\ \theta_E \\ \theta_M \end{bmatrix}, \qquad \theta_i = \begin{bmatrix} \alpha_i \\ \beta_{1i} \\ \vdots \\ \beta_{5i} \\ \alpha_i^* \\ \beta_{6i} \\ \vdots \\ \beta_{10i} \end{bmatrix}, i = K, L, E, M,$$

$$\epsilon = \begin{bmatrix} \epsilon_K \\ \epsilon_L \\ \epsilon_E \\ \epsilon_M \end{bmatrix}, \qquad \epsilon_i = \begin{bmatrix} \epsilon_{1i} \\ \epsilon_{2i} \\ \vdots \\ \epsilon_{ni} \end{bmatrix}, i = K, L, E, M,$$

となる. そして, 対称性と一次同次の制約 (5.3) を確率的に課すことを考える. これは次のように考えることができる. コスト関数の理論的性質は個々の企業について想定されたものであるが, 多くの利用できるデータは集計されたものである. そのために, これらの制約はある誤差の幅を持って考えることができることから, このような確率制約を用いる方が現実的である.

我々の5要素コスト (生産) 関数の場合, $m = 4$ で $K^* = 6$ であるから, 各方程式の Q は $2K^*m \times m(m+3)$ の既知行列である. K^* は最初の状態における (5.17) 方程式の係数の数であり, Q は $Q = (I_2 \otimes Q_1', I_2 \otimes Q_2', I_2 \otimes Q_3', I_2 \otimes Q_4')'$ で与えられる. ここで, 具体的には

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$Q_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$Q_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & -1 & 0 & 0 \end{bmatrix}$$

$$Q_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & -1 \end{bmatrix}$$

で, γ は $m(m+3) \times 1$ の未知定数ベクトルである.

5.4 計測結果の評価

前節で述べた生産(コスト)関数の構造変化モデルを, 2つのエネルギー多消費産業である, 鉄鋼業と紙・パルプ産業に適用した[5]. データは3節で説明した1961年第II期から1980年第IV期までの四半期データ(季節調整済み)である. これらの産業は, 2節で述べたように1973年に始まる石油危機に対する省エネルギー対策の対照的な例として選ばれた. ある時期から次の時期へ構造変化があったとすると, 産業は硬直的な価格で調整するよりも, 数量での調整の方を早く行うことが考えられる. また, 鉄鋼業ではパラメータのシフトが期待される一方で, 紙・パルプ産業でははっきりとしたシフトは見いだされないと思われる.

モデルの遷移関数としては, 双曲線正接(hyperbolic tangent) を用いたが, これは単に使いやすいためと, 事後密度の推定が遷移関数の選択によって敏感には変化しないように見えるためである[6]. さらに, 制約については, 階層的な表

[5] 本節での計算はすべてSAS 82.4のPROC MATRIX を用いて行った.

[6] 遷移関数としては他に $1-\exp(-x)$ や $x^2/(1+x^2), x = S_t/\eta$ なども用いたが, 事後密度に

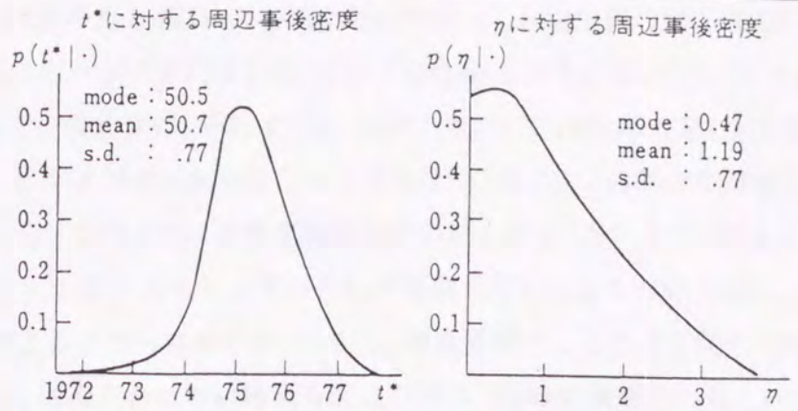


図 5.3: 鉄鋼業(コスト関数)における周辺事後分布

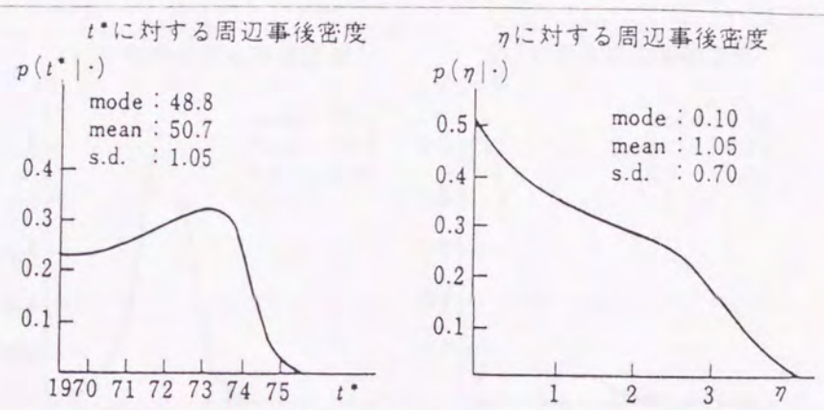


図 5.4: 鉄鋼業(生産関数)における周辺事後分布

現と混合的な表現とは同じ事後密度になるので, ここでは前者の場合による結果を用いて分析する.

図5.3と図5.4は, 鉄鋼業における変化点 t^* と調整スピード η に対する周辺事後密度を, コスト関数と生産関数について表したものである. この周辺事後密度は $\sigma^2 = 1 - 10^{-9}$ と確定的な制約の下での反復多変量回帰の残差から推定した共分散行列 Ω [7], という条件の下で得られたものである.

σ^2 の値は, 推定する方程式での先験的な制約の重要性を表すもので, この 10^{-9} では対称性について5桁まで等しくなっている. また 10^{-3} ぐらいでは, この対称性はほとんど満たされない. 変化点 t^* と調整スピード η の事後平均あるいはモードを見ていくと, コスト関数(価格調整)での変化点は50.68で, これは1973

大きな変化はなかった.

[7] この計算は, 確定的な $R\theta = r$ の制約を用いて得られた残差の共分散行列で, 全サンプルを用いて求めた.

年の第III四半期と第IV四半期の間に対応し，石油危機の頃の時期になる．また，調整スピードの平均は1.2，モードは0.46と非常に早い．一方，生産関数(数量調整)での変化点は，平均47.78，モード48.75で1973年の第I四半期頃に見られるが，それほどはっきりしたものではない．また，この場合の調整も即時に行われている．このように調整が即時に行われるのは，次のように考えることができる．データはコストシェアの合計が毎期毎期1になるという会計上の恒等式を満たすように作られている．しかし，調整期間ではこの恒等関係は対称性と同様，しばしば満たされない場合も起こりうる．即時に調整されるというのは，調整期間であってもこの条件を満足するように制約を与えた結果と考えられる．

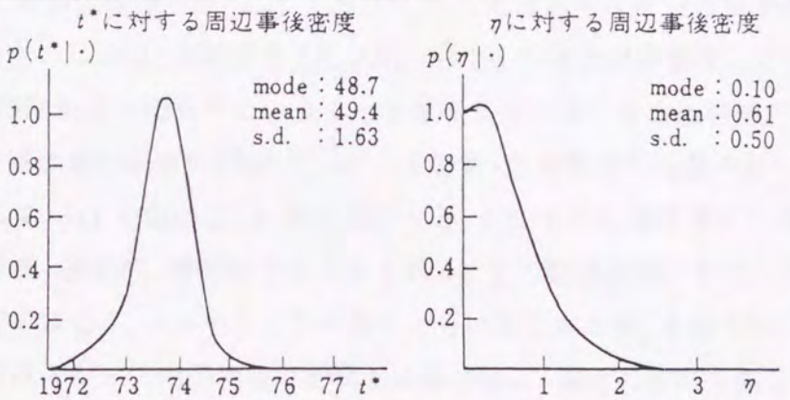


図 5.5: 紙パルプ産業(コスト関数)における周辺事後分布

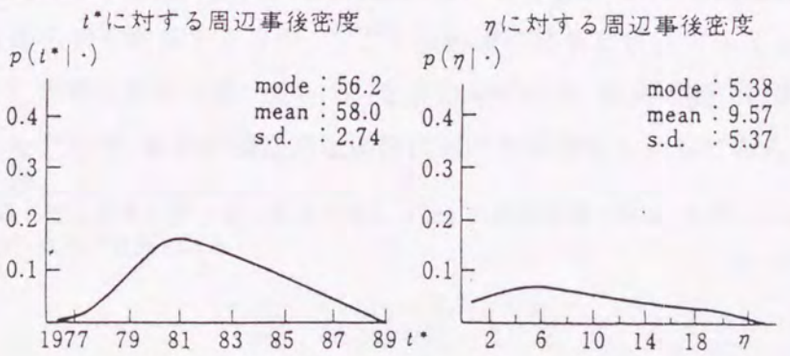


図 5.6: 紙パルプ産業(生産関数)における周辺事後分布

上の図5.5と図5.6は，紙・パルプ産業における変化点と調整スピードの周辺事後密度であり，σ^2 と Ω については鉄鋼業の場合と同じである．コスト関数での t^* のモードは48.75で1973年の第I四半期と第II四半期の間にあり，この場

合も調整スピードは速い. この時期は為替レートの自由化により卸売物価が高騰した頃に当たり, この影響が大きいと思われる. 一方, 生産関数の場合には, $t*$ も η もはっきりしたピークは見られない. 両産業において, コスト関数による方が, 変化が明瞭に読みとれる傾向にある.

次に要素間の代替弾力性と自己価格弾力性の変化について見てみよう. 石油危機の前後で両産業の成長がどのように変化したかを考えるとき, 問題となるのは代替弾力性や価格弾力性の大きさである. エネルギーと生産量との関係を考えるとき, エネルギーと他の生産要素間の代替関係の大きさが問題となってくる. エネルギー経済分析で最もよく利用される代替弾力性の概念は, Allen の定義による偏代替弾力性 (σ_{ij}) であり[8], たとえばエネルギーと資本の弾力性 (σ_{KE}) が1というのは, 相対価格 (X_E/X_K) の1% の増加が要素投入比率 (X_K/X_E) の1%の増加をもたらすということを意味している. また価格弾力性 (η_{ij}) は, 第 j 生産要素の価格が1%変化したときの第i生産要素の数量の変化率で表される. そして, σij の値が正, 0, 負の値のとき, i と j の生産要素間の関係はそれぞれ代替的, 独立的, 補完的であるとされる. また観測期間にわたって σ_{ij} と η_{ij} は一定ではなく, コストシェアの値とともに変化するが, 本稿では石油危機の前後におけるシェアの平均値で計算した弾力性と, 最近の傾向を見るために80年における弾力性を示した. 表5.2は鉄鋼業についての Allen の偏代替弾力性 (σ_{ij}) と自己価格弾力性 (η_{ij}) である. これらの弾力性は, コスト関数と生産関数のどちらからでも計算できるが, ここでは前述の結果と計算の容易さからコスト関数から求めたものを用いている. なお参考のため, 最後の欄には構造変化を考えないモデルで, 制約が確定的な場合における弾力性を示してある.

---

[8] 第 i 投入要素と第 j 投入要素の間の Allen の偏代替弾力性は, トランスログ・コスト関数の場合, 次式で計算される.

$$\begin{aligned} \sigma_{ij} &= (\beta_{ij} + S_i S_j)/S_i S_j \qquad i \neq j \\ \sigma_{ii} &= (\beta_{ii} + S_i^2 - S_i)/S_i^2 \end{aligned}$$

ここで, S_i は第 i 投入要素のコストシェアである. β_{ij} は第 j 方程式の第 i 投入要素の係数で, この場合(5.17)式の

$$c_1 = (M_Q/\sigma^{-2} + X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$$

で得られる. 自己価格弾力性は

$$\eta_{ii} = \sigma_{ii} \cdot S_i$$

で計算される.

表 5.2: 鉄鋼業における Allen の偏代替弾力性 (σ_{ij}) と自己価格弾力性 (η_{ii})

	石油危機前	石油危機後	1980年(平均)	全期間(確定制約)
σ_{KL}	-0.808	-0.204	-0.244	0.527
σ_{KE}	-0.676	0.546	0.609	1.023
σ_{KM}	-0.309	1.233	1.222	0.025
σ_{KZ}	1.855	0.147	0.153	0.729
σ_{LE}	1.843	-0.033	-0.046	1.520
σ_{LM}	-0.262	-0.615	-0.803	-0.540
σ_{LZ}	2.184	0.850	0.825	2.197
σ_{EM}	-2.235	0.119	0.179	0.106
σ_{EZ}	2.582	-0.194	-0.162	-0.631
σ_{MZ}	2.067	-0.523	1.000	0.313
η_{KK}	-0.337	-0.342	-0.366	-0.560
η_{LL}	-0.975	-0.174	-0.113	-0.083
η_{EE}	-0.670	-0.022	-1.365	-0.138
η_{MM}	-0.200	0.053	0.064	-0.066
η_{ZZ}	-1.441	-0.001	0.019	-0.299

表5.2の結果を見ると, 資本と労働は両期間を通じて補完的であったが, 石油危機後は補完性は弱くなっている[9]. エネルギーに対する関係ではすべてに大きな変化が見られる. エネルギーと資本は, 石油危機前では補完的であったのが, 危機後は代替的になっている. この結果は, 最近のエネルギーと資本の補完性についての議論と比べてみるとおもしろい. Berndt and Wood (1979) による説明が, この場合よく当てはまるように思える. この仮説は, エネルギーと資本は, それが利用される資本についての部分関数を考えたときには代替的であるというものである. 石油危機以前では, エネルギーと資本の代替関係は, 利用される資本を一定とすると, 資本利用の変化により大きな拡張効果を持っており, それが両投入要素の需要を増加させることになった. したがって, 推定されたネットでの効果は補完的ということになる. 第1次石油危機以後状況は変わり, 鉄鋼業は省エネルギーのための投資を積極的に行ったが, それは低い資本利用がもたらす

[9]本稿では技術変化は仮定しなかった. 当然, 技術変化を考慮すると, ここでの結論は変わってくる. 技術変化を含んだモデルで同様の分析を行った例は Tsurumi, Wago and Ilmakunas (1985) を参照のこと.

マイナスの効果を重要視しすぎたためかも知れない. したがって, エネルギーと資本はネットでは代替的であると考えてよいであろう. 同じことは原材料に対しても見られるが, これもエネルギー節約のための投資が, 同時に効率的に原材料を使用するものでもあったと考えられる (σ_{KM} と σ_{EM}). もう一つのパターンはエネルギーと労働との間に見られる. 石油危機以前では代替的であったが (2.2節参照), その後補完的または独立的になっている. この理由は, 資本投資はエネルギー節約のためだけでなく, 僅かだが労働節約的な方向にも働いたということかもしれない.

価格弾力性については, 石油危機以前は経済理論から期待されるようにすべて負であったが, その後原材料は独立的になっている. エネルギーの価格弾力性は, 石油危機以前では労働に次いで高かったが, その後平均的に低くはなっているが, 最近では極めて高くなっている. これは後述するように, エネルギーのコストシェア (80年で平均29%) と比べるとかなり大きい.

次に, 紙パルプ産業についての結果が表5.3である. 資本と労働については両期間を通じて代替的であったが, 石油危機以後は代替性は弱くなっており, 最近では補完的になっている. 鉄鋼業と比べてみるとエネルギーに対する関係では大分異なったパターンが見られる. エネルギーと資本は両期間とも代替的であり, エネルギーと労働は補完的である. しかし, その値を見ると, 石油危機によっても前者の関係は余り変化しなかったのに対して, 後者の関係から労働投入は大幅に減ったことが分かる. 原材料に対しては, 石油危機以後, 製品構成の変化にともなうと思える理由により補完的になっているが, 最近では独立的である.

価格弾力性については原材料を除いてすべて負であり, 石油危機以前では労働が一番高い. 石油危機以後には一般に弾力性の大きさは小さくなっているが, 鉄鋼業の場合と比べるとその変化は小さい. 特にエネルギーの自己価格弾力性は変化が少なく, 80年では鉄鋼業の −1.36 に対して −0.26 の大きさである.

次に, エネルギー価格とエネルギー生産性の関連について考えてみる. 佐和 (1983) によれば, エネルギー価格の上昇がエネルギー生産性を上昇させるための必要十分条件は, 代替の弾力性, または価格弾力性の絶対値が, エネルギー・コストシェアより大きいことである. 第1次石油危機のようなエネルギー供給制約, あるいは価格高騰に対して産業が頑健であるか否かは, 今弾力性の大きさを構造

表 5.3: 紙・パルプ産業における Allen の偏代替弾力性 (σ_{ij}) と自己価格弾力性 (η_{ii})

	石油危機前	石油危機後	1980年(平均)	全期間(確定制約)
σ_{KL}	0.497	0.122	-0.163	1.562
σ_{KE}	0.342	0.351	0.517	0.399
σ_{KM}	0.016	0.393	0.409	0.297
σ_{KZ}	0.800	0.350	0.280	0.289
σ_{LE}	-0.103	-1.217	-1.109	0.297
σ_{LM}	0.069	-1.391	-1.976	-1.524
σ_{LZ}	1.751	1.780	1.110	2.278
σ_{EM}	0.065	-0.529	-0.070	-0.531
σ_{EZ}	0.672	0.621	0.698	0.230
σ_{MZ}	-0.550	-0.451	-0.511	-0.441
η_{KK}	-0.421	-0.296	-0.288	-0.390
η_{LL}	-0.798	-0.349	-0.198	-0.849
η_{EE}	-0.327	-0.108	-0.264	-0.056
η_{MM}	-0.209	0.332	0.293	0.342
η_{ZZ}	-0.219	-0.183	-0.185	-0.187

的なパラメータと見なすと, コストシェアの時間的な変動によって変わってくる. 第1次石油危機以前では, エネルギーのコストシェアは平均で, 鉄鋼業で19.8%, 紙・パルプ産業で9.7%であった. その後の期間では平均で26.6%と13.1%にまでなり, 80年では29%と18%になっている. この数値を代替の弾力性や価格の弾力性と比べてみると, 生産性を上昇させる条件は満たされている可能性は高い. しかし, 鉄鋼業と比べると紙・パルプ産業はその差があまり大きくなく, エネルギー価格上昇に対する抵抗力は, このままエネルギーのコストシェアが上昇するとかなり弱くなってくると考えられる. そのためにも, 省エネルギーの進展, あるいは産業構造の転換等によってこの弾力性を高くし, それを維持する体質にする努力が必要になる.

5.5 おわりに

本稿では, エネルギー価格の上昇が産業の生産性の調整にどのような意味を持っていたかについて, 代表的なエネルギー多消費産業である鉄鋼業と紙・パルプ産業を例にとって, その省エネルギー対策の相違を対照させて分析した. 石油危機を境に構造変化があったとの想定から, ベイジアンの手法を用いて, 変化点と調整スピードについての周辺事後密度を求めた. 生産構造を表すモデルとして, 強い制約をおかないトランスログ型の関数を用い, 理論から要請される制約を方程式間の確率制約と考えて推定した. 両産業についての推定の結果, いずれも石油危機の前後に構造変化があったことが確かめられた. コスト関数では明らかに変化が認められるのに対し, 生産関数では明瞭には認められなかった. これは数量データでは, 構造変化がなかったということになる. 調整スピードに関しては, パラメータ制約のためもあり, すべて即時に調整されるという結果を得た.

代替弾力性について, 鉄鋼業ではエネルギーに対する他の生産要素との関係は大きな変化を見せるのに対し, 紙・パルプ産業ではあまり大きな変化はない. また, 両産業の弾力性の大きさとエネルギー・コストシェアの関係を見ると, 今後のエネルギー価格上昇を考えると, 鉄鋼業に比べて紙・パルプ産業の抵抗力が弱まっていくと考えられる. 構造変化モデルによる推定は, このような分析には必要であり, これは確定制約の場合の推定値と比べてみると, 確定制約の場合には必要な弾力性について誤って解釈をする可能性があることを示唆している.

参考文献

1. Bacon, D.W. and D.G. Watts (1971), "Estimating the Transition between Two Intersecting Straight Lines" *Biometrika,* vol.58, pp.525–34.

2. Barnett, W. (1983),"New Indices of the Money Supply and the Flexible Laurent Demand System" *Journal of the Business and Economic Statistics,* vol.1, pp.7-23.

3. Berndt, E.R. and D. Wood (1975),"Technology, Prices, and Derived Demand for Energy" *Review of Economics and Statistics,* vol.57, pp.259–68.

4. ——— (1979), "Engineering and Econometric Interpretation of Energy-Capital Complemetarity" *American Economic Review,* vol.69, pp.342–54.

5. Christensen, L.R., S,W. Jorgenson, and L.J. Lau (1973), " Transcendental Logarithmic Production Frontiers," *Review of Economics and Statistics,* vol.55, pp.28–45.

6. Gallant, A.R. (1981), "On the Bias in Flexible Functional Forms and an essentially Unbiased Forum: The Fourier Flexible Form" *Journal of Econometrics,* vol.15, pp. 211–45.

7. ———— (1982), "Unbiased Determination of Production Technologies" *Journal of Econometrics,* vol.20, pp.285–324.

8. 林　周二他 (1983),「トランスログ型関数とエネルギー多消費産業における要素間代替」昭和57年度文部省科学研究費補助金, エネルギー研究(1), ディスカッションペーパー

9. 伊藤 浩吉 (1983),「トランスログ生産関数による代替の弾力性, 価格の弾力性の計測」エネルギー経済研究, 第1巻, pp.233–50.

10. 伊藤・松井 (1978),「安定成長軌道における経済活動とエネルギー」日本エネルギー経済研究所, pp.27–44.

11. Kuroda, M. and H. Wago (1982), "Relative Price Changes and Biases of Technical Change in Japan–KLEM Production Model," Discussion Paper 142, Institute of Socio-Economic Planning, University of Tsukuba.

12. Lindley, D.V. and A.F.M. Smith (1972), "Bayes Estimates for Linear Model, " *Journal of Royal Statistical Society, Ser. B*, vol.34, pp.1–18.

13. 中島 清三 (1983),「トランスログ関数によるマクロ経済の実証分析」エネルギー需給の計量分析, 経済企画庁経済研究所研究史シリーズ, 第40号, pp.25–50.

14. 大山 達雄 (1983)「トランスログモデルによるわが国の1次エネルギー消費分析」エネルギー需給の計量分析, 経済企画庁経済研究所研究史シリーズ, 第40号, pp.232–66.

15. Rossi, P. (1984), "Specification and Analysis of Econometric Production Function Models," Unpublished Doctoral Dissertation, University of Chicago.

16. 佐和 隆光 (1983),「エネルギー価格変動と潜在GNP」エネルギー需給の計量分析, 経済企画庁経済研究所研究史シリーズ, 第40号, pp.8–24.

17. Smith, A.F.M. (1973), "A General Bayesian Linear Model," *Journal of Royal Statistical Society, Ser. B.* vol.35, pp.67–75.

18. 鈴木 武 (1982),「トランスログ型関数とエネルギー多消費産業における要素間代替」昭和556年度文部省科学研究費補助金, エネルギー研究(1), ディスカッションペーパー

19. Theil, H. (1963), "On the Use of Incomplete Prior Information in Regression Analysis," *Journal of Americal Statistical Association,* vol.58, pp.65–78.

20. Theil, H. and A.S. Goldberger (1961), "On Pure and Mixed Statistical Estimation in Economics," *International Economic Review,* vol.2, pp.65–78.

21. Tsurumi, H. (1980), "A Bayesian Estimation of Structural Shifts by Gradual Switching Regressions with an Application to the U.S. Gasoline Market," A. Zellner ed., *Bayesian Analysis in Econometrics and Statistics,* in honor of Harold Jeffreys, North-Holland, Amsterdam.

22. Tsurumi, H. and H. Wago, and P. Ilmakunnus (1985), "Gradual Switching Multivariate Regression Models with Stochastic Cross-Equational Constraints and an application to the KLEM translog production model," paper presented at the 5th World Congress of the Econometric Society and *Journal of Econometrics,* vol.31, pp.235–253, 1986.

23. 和合 肇 (1983a),「エネルギー価格, 代替弾力性と技術変化」日本統計学会誌, vol.13, pp.73–88.

24. 和合 肇 (1983b),「シェアモデルにおける推定と検定」竹内 啓編『計量経済学の新展開』東京大学出版会, pp.203–36.

25. 和合・鈴木 (1980),「日本経済における要素間代替と技術変化」第47回 日本統計学会報告, mimeo.

26. 和合・鈴木 (1981),「わが国製造業における要素間代替」第48回 日本統計学会報告, mimeo.

27. Woodland, A.D. (1979), "Stochastic Specification and Estimation of Share Equations," *Journal of Econometrics,* vol.10, pp.361–83.

付録1：資本ストックデータの作成について

われわれのモデルに必要な四半期の資本ストックデータは，これまで利用に十分耐えうるものは作成されていない．そこで，今回，鉄鋼業と紙・パルプ産業について実質資本ストックを新たに推計し直すことにした．また，これと関連して資本の価格系列の推計も行った．いずれの手法も，他の産業に対しても適用可能なものである．

(i) ベンチマーク

1970年国富調査における「産業大分類，資産中分類別，有形固定資産額–全国(粗資産額)」を75年価格にデフレートしたものをベンチマークとして採用した．

(ii) 資本の増分

投資の時系列データは「法人企業統計年報」および「季報」のデータから作成した．

(a) 法人企業統計年報
KV = その他の有形固定資産 − 内建設仮勘定
D = 減価償却 + 特別減価償却
から
$Y = KV - KV_{-1} + D$
を年度ごとに作成する．

(b) 年　　報
「その他の有形固定資産(建設仮勘定を含まず)」における，INC(増加額)，DEC(減少額)，DEP(内減価償却額)から，$X = INC - (DEC - DEP)$ を各四半期について作成する．

(c) 四半期分割
年報の Y を分割する際，各期の除却率を用いて除却部分の調整を行う．

$$Y_t = Y \times \frac{INC_t}{\sum_1^4 INC_t} - KV_{t-1} \times R'_t$$

ただし，R_t(除却率) $= (DEC_t - DEP_t)$/季報の KV_{t-1}
$R'_1 = R_1, R'_2 = R_2(1-R_1), R_3 = R_3(1-R_2)(1-R_1), R'_4 = R_4(1-R_3)(1-R_2)(1-R_1)$

(d) 実質化と資本ストック
(c)で求められた投資を，次に述べるデフレータによって75年価格へと実質化した．次に，70年10-12月期における資本ストックがベンチマークに一致するように基準化をして，投資額の累積データストックを作成し，期種データへ変換した．

(iii) デフレータ

国富調査と法人企業統計の実質化のためのデフレータは以下のように作成した．

(a) 国富データ
各資産分類ごとにデフレートして，積み上げた．その際使用したデフレータは次の通り．

「住宅」「非住宅」「構築物」については建設省の「建設デフレータ」(建設統計月報)の「住宅」「非住宅」「土木」を用いた．なお，「建物」については，住宅・非住宅のウェイトで合成したデフレータを使用した．「機械及び装置」については，法人企業統計の

投資デフレータのうち, 自動車, 自動二輪車, 住宅, 非住宅などを除いた機械及び装置に対応するデフレータを作成して利用した.

「船舶」については適当な価格が入手できず, また金額もきわめて小さいため, 卸売物価指数の「輸送用機器」で代用した. 「車両及び運搬具」については, 「鉄道車両」と「自動車」とを卸売り物価指数で合成して作成した.

「工具・器具・備品」及び「大動植物」については, それぞれ卸売物価指数の「機械工具類」及び「総合」を利用した.

(b) 法人企業データ

鉄鋼業, 紙パルプ産業については, 「昭和45年産業連関表」における「資本機能 × 資本財マトリックス」から投資財のウェイトを求め, これを用いた卸売物価指数と建設デフレータを合成した.

以上が資本ストックデータの作成方法の概要である. この作成方法に関して, 次の問題点が指摘できる. 法人企業年報における, D = 減価償却 + 特別減価償却 は会計上の概念であって, 生産力を表すべき資本ストックの減少分とは見なせない. したがって, 設備等が除却されたときの除却額は, 生産力という意味からは過小評価されている. われわれの採用した方法は, このデータ処理に当たって, 過大推定と過小推定の2つの系列を作成し, 生産力に対応する資本ストックを不等式ではさまれる区間で捉えようとしたものである. いま, (c) の四半期分割において, 年報の Y を季報の X によって単純に分割する. $Y_t = Y \times X_t / \sum_{j=1}^{4} X_j$ とすれば, これは減価償却を一切認めないことによって明らかに過大推計になる. 一方, (c) において, R_t を DEC/年報の KV_{t-1} で置き換えれば過小推定になる. われわれの採った方法でも, まだ除却率 R_t が過小推定されているため, 資本ストックは過大推定であると思われる. すなわち, その分子 $DEC - DEP$ は会計上の除却を表しており, 減価償却が除かれている分だけ過小に推計されている.

次に, この方法を国富調査と比較することによって, その整合性をチェックしてみる. 国富調査における純資産と法人企業統計におけるその他の有形固定資産 (建設仮勘定を除く) を比較すると, 全企業では昭和35年で14%, 45年で11% の誤差である. そこで, ここでの方法と同じ形式のものを年度単位にして, 国富調査の粗資産と比較すると約7.5% の誤差になる (詳細は林他 (1983) 参照). 過大推計が疑われる数値が, 国富調査と比較すれば僅かに小さいということで, この方法にも問題があるが, 当面これ以上の改善は困難であると考えられる.

付録2：データの出所等

＜鉄鋼業＞

数量データ	単位	出所
普通鋼圧延鋼材生産量	(万 t)	鉄鋼統計年報より作成
特殊鋼 〃	〃	〃
鉄鉱石消費量	〃	鉄鋼統計年報
砂鉄 〃	〃	〃
溶解用鋼くず消費量	〃	〃
発生量	〃	〃
月間延べ人員合計	(100万人)	〃
実労働時間指数	(50年=100)	毎月勤労統計
電力買電量	(10万Kwh)	鉄鋼統計年報
コークス消費量	(万 t)	〃
重油消費量計	(1000Kl)	〃

価格データ	指数の出所	絶対水準の出所
普通鋼圧延鋼材価格	(50年=100) 物価指数年報	鉄鋼統計年報
特殊鋼 〃	〃 〃	〃
鉄鉱石 価格	〃 〃	〃
砂鉄 価格	〃 〃	〃
鉄くず 価格	〃 〃	日本経済新聞
賃金	〃 毎月勤労統計	毎月勤労統計
電力(大口用) 価格	〃 物価指数年報	電気事業便覧
コークス 価格	〃 〃	鉄鋼統計年報
重油 価格	〃 〃	石油統計年報

＜紙パルプ産業＞

数量データ	単位	出所
紙, 板紙生産量	(万 t)	紙パルプ統計年報より作成
原材料紙パルプ輸入量	(万 t)	〃
故紙　消費量	(万 t)	紙パルプ統計年報
故紙パルプ　〃	〃	〃
原木　〃	(万 m^3)	〃
チップ　〃	〃	〃
月間実働延人員計	(万人)	〃
実労働時間指数	(50年=100)	毎月勤労統計
買電消費量	(1000万Kwh)	紙パルプ統計年報
重油消費量	(1000Kl)	〃
石炭消費量計	(1000t)	〃
カロリー換算表		日本エネルギー経済研究所 エネルギーマトリックス
紙・板紙出荷金額	(億円)	紙パルプ統計年報
〃　数量	(万 t)	紙パルプ統計年報

価格データ	指数の出所		絶対水準の出所
パルプ　価格	(50年=100)	物価指数年報	紙パルプ統計年報
故紙　〃	〃	〃	日本経済新聞
原木　〃	〃	〃	紙パルプ統計年報
チップ　〃	〃	〃	〃
賃金	〃	毎月勤労統計	毎月勤労統計
電力(大口用)価格	〃	物価指数年報	電気事業便覧
重油　〃	〃	〃	石油統計年報
石炭　〃	〃	〃	物価指数年報

第6章　ベイズ予測

6.1　はじめに

予測とは, 不確実な将来についての一つの見方であり, 予測を行う主な理由の1つは, 賢明な行動がとれることにある. 経済予測の場合を考えてみよう. 適切な経済計画の遂行のためには, 来年の経済成長はどのくらいになるだろうか, あるいはインフレ率は何パーセントぐらいになるだろうかということを知りたい. 企業ではそれによって新規の工場建設に対する設備投資を行うべきか否かを決定し, 将来の顧客の需要, 販売コスト, 利益について計画を立てる. また, 多品種少量生産の下での在庫管理では, ここの商品の毎月の販売高の予測を的確に行う必要がある. このような予測を行う方法としては, 単純な指数平滑法から, 時系列モデル, 回帰モデル, そして複雑な計量経済モデルによる方法などが知られている. しかし, これらの方法による予測では, 実際はただ一つの予測値だけしか考慮の対象にしていない. これは非常に断定的で危険が多い結果をもたらす場合が多い. われわれが必要とするのは, むしろ可能性や確率を表す確率分布による予測であろう. ベイ時予測は, このような予測を可能にする方法であるといえよう.

予測を行う場合に利用可能な情報源は, 大ざっぱにいえば, 過去の歴史的な出来事と, 将来のことに関する知識とに分けられる. 過去の歴史は単にGNPや物価の動きとかある製品の動きとかある製品の販売高や価格のような記録されたデータだけでなく, すべての過去の経験が含まれる. これらはすべて一緒になって予測システムを形成する. すなわち, これは経験と分析に基づいてルーティン的に処理したデータと情報を取り入れた統計モデルである. このようなモデルによる予測は, 次のような例外的な事態がない限り用いられる. 1つは将来起きるかもしれない出来事について, その専門知識に基づいた見通しに関するもの

である. たとえば, 半導体の輸出規制の問題とか, 突然の円高や, 株暴落の可能性, 第3次オイルショックによるエネルギー危機の可能性, 等々についての予想に関しては, 分析者がそのインパクトをどのように考えるかによっている. もう1つは, フィードバック情報に関するもので, 予測パフォーマンスが不適切であるために生ずる. これは通常, 統計的モニタリング, すなわち最近時点でのモデルの予測能力を連続的に評価することによって発見することができる. 完全な予測システムは, このように統計モデルによって処理した情報と, 主観的な情報を結びつけたものでなくてはならない. そして, 例外的な事態が生じたときには, その原因を調べてモデルの関連した部分を, 新しい情報をもっと早く取り入れることができるようなシステムに改良する必要がある. あるいは, その変化について特定の対立仮説を想定する多段階モデルの考え方 (Harrison and Stevens [1971, 76] 参照) をシステムに含める方法もある. 多段階モデルやベイズ予測の考え方は, それが完全に自動的な予測システムを提案するものであると思われている. しかし, ベイズ予測は全く新しい考え方というものではなく, モデル化したデータ処理と主観的な情報を結びつけることの重要性を強調した予測のシステムである.

以下で述べるベイズ予測の方法は, 動的線形モデルとその要素についてのモデル・ビルディングに基づいて行われるもので, Harrison and Stevens (1976) によって提唱され, West, Harrison and Migon (1985) や West (1986) でさらに展開されている. システムを状態空間 (*state space*) 表現を用いて表し, 状態空間変数を未知の係数と考えたとき, この係数は時間で変化するので, このモデルのことをDLM (*Dynamic Linear Model*) という. この係数とその共分散行列の推定には, カルマン・フィルター方程式が用いられる. ベイズ予測の目的は, 将来観測値に対する予測分布を導くことである. 今までに状態空間表現を用いたいろいろなモデルが提案されている. たとえば, Harvey (1984) やHarvey and Todd (1983) の1変量状態空間構造モデル (*univariate state space structure model*), Kitagawa (1981), Gersch and Kitagawa (1983), Kitagawa and Garsch (1983), Kitagawa and Garsch (1984, 1985) などはこのモデルに非常に似通ったモデルである. ここで用いられ状態空間モデルによる予測は, ベイズ予測モデルをはじめ, 伝統的な回帰モデルやARIMA時系列モデル, そして時変係数モデルなどをその特別な

場合として含んでいる. このモデルについては, その特徴は一部の専門家を除いてあまり知られていないので, Harrison (1967) や Harrison and Stevens (1971, 1976) に沿って, 第2節で状態空間モデルについて簡単に紹介し, 第3節で予測に用いる構造時系列モデルの各成分のモデル化の方法を述べ, 第4節では主要結果と情報割引の考え方を示し, 外生的な出来事 (*event*) によるモデルの影響と修正について説明する. 第5節で典型的な分析れいについて述べた後, 最後に第6節で多変量データを扱うモデルへの一般化を考える.

6.2 ベイズ予測と動的線形モデル (DLM)

A. 動的線形モデル

ベイズ予測システムの基本要素は, 次のようにまとめることができる.

(1) 構造モデルによって分析対象をはっきりと定義し, パラメータによってモデルの性質を意味付けし, 解釈し, 他の分析者とコミュニケートすることを可能にする.
(2) 任意の時間 t で利用できる情報と知識 D_t の下で, パラメータについての確率的な情報を用いる.
(3) どのようにシステムを考え, どのようにパラメータが確率的に変化するのかについて記述する逐次的モデルを用いる.
(4) 予測は確率分布の形を取る.

動的線形モデル (DLM) とは, (1) ある確率過程 (以下, 簡単化のためプロセスと呼ぶ) の観測値がパラメータにどのように確率的に依存しているか, (2) プロセス・パラメータが, その固有の動学プロセスと確率ショックまたは攪乱項の両方から, 時間が経過するにつれてどのように変動するか, を表す方程式体系である. モデルは, 離散型, 等時間間隔で表される.

このモデルで用いられる状態空間モデルとは, 観測される T 個のデータ y_t を, 時間とともに変化する n 個の状態変数 θ_t と結びつける方程式体系である. このモデルは各時間 t に対して4つのパラメータ $\{F, G, V, W\}_t$ で定義される. n 個の状態を記述した $n \times l$ プロセスベクトル θ_t と事後情報の下で, 連続的な確

率関係は次の2本の確率逐次方程式の形で書くことができる.

$$\text{観測方程式}: y_t = F_t'\theta_t + v_t \quad v_t \sim N(0, V_t) \tag{6.1}$$

$$\text{システム方程式}: \theta_t = G_t\theta_{t-1} + w_t \quad w_t \sim N(0, W_t) \tag{6.2}$$

y_t は将来観測値を表し, F_t は n 個の独立変数ベクトルで t 期では既知, G_t は既知の $n \times n$ システム行列, v_t は独立な正規確率変数ベクトルで平均ゼロ, 分散は未知の V_t, そして w_t は独立な正規確率変数ベクトルで平均はゼロ, 分散共分散行列は $W_t(= V_tB, B$ は任意の正値定符号行列) と仮定する. D_{t-1} は $t-1$ 期までに用いられたすべての知識を表す.

古典的な線形回帰モデルの場合は, $\theta_t = \theta$ であり, V_t が時間に依存しない単純な静学的な場合のDLMになり, システム方程式 $\theta_t = \theta_{t-1}$ は必要なくなる. このプロセスパラメータ θ_t は従属変数と独立変数との関係を表し, 静学的な線形回帰モデルのパラメータ的な解釈ができる. F_t が固定されている場合はモデルは時系列の状態空間表現になり, このときはパラメータはプロセス水準とかプロセス成長として解釈することもできる. このDLMはいくつかのモデルの線形結合として表すこともでき, Lindley and Smith (1972) の階層 (*hierarchical*) モデルもまたDLMとして表すことができる.

B. 推定

初期時点 $t = 0$ では, パラメータベクトル θ_0 に関する情報は, 平均 m_0 で分散共分散行列 c_0 の正規分布で表される. 最初の観測値以前の θ_0 の分布を $N(m_0, c_0)$ とすると, t 期における θ_t の事後分布も正規分布をする.

$$(\theta_t|D_t) \sim N(m_t, c_t) \tag{6.3}$$

ここで, $m_t = E(\theta_t|D_t]$ と $c_t = V(\theta_t|D_t)$ の値は逐次的に得られる. y_t のプロセスをDLMの形で表すと, Kalman (1960) の結果を用いてパラメータベクトルの情報を更新する. この結果を予測への応用に便利な形で表すと, 次のようになる (Anderson and Moore (1979) またはHarvey (1981) 第4章を参照).

$$E(\theta_t|D_{t-1}) = G_tE(\theta_{t-1}|D_{t-1}) \tag{6.4a}$$

$$V(\theta_t|D_{t-1}) = G_tV(\theta_{t-1}|D_{t-1})G_t' + W_t \tag{6.4b}$$

$$E(y_t|D_{t-1}) = F_t' E(\theta_t|D_{t-1}) \tag{6.4c}$$

$$V(y_t|D_{t-1}) = F_t' V(\theta_t|D_{t-1}) F_t + V_t \tag{6.4d}$$

$$A_t = V(\theta_t|D_{t-1}) F_t V(y_t|D_{t-1})^{-1} \tag{6.4e}$$

$$e_t = y_t - E(y_t|D_{t-1}) \tag{6.4f}$$

$$E(\theta_t|D_t) = E(\theta_t|D_{t-1}) + A_t e_t \tag{6.4g}$$

$$V(\theta_t|D_t) = (I - A_t F_t') V(\theta_t|D_{t-1}) \tag{6.4h}$$

このアルゴリズムでは, この逐次性が重要である. 実際の予測では, 現在の事後分布は最も直近の値 (y_t, F_t), 事後分布 $(\theta_{t-1}|D_{t-1})$, 現在の観測値ノイズ分散 V_t と撹乱項分散 W_t から計算される. ここで $E(y_t|D_{t-1})$ と $V(y_t|D_{t-1})$ は $t-1$ 期までの情報の下での y_t の1期先予測の期待値と分散であり, したがって, e_t は条件付きの1期先予測誤差になる. A_t はカルマンゲインと呼ばれ, 多くのシステムでの平滑化定数ときわめて似たものであるが, A_t は一般には一定ではなく, $n \times T$ 行列である.

カルマンフィルターでは各時点において利用可能な情報の下で, θ_t の最小二乗線形推定量(MMSE)を得る. しかし, ベイズ分析ではカルマンフィルターはパラメータの事前分布を更新して, 事後分布を得る方法として解釈される. こう考えることによって, 標本期間の最初で事前情報が利用できる場合にはそれを予測に組み入れることが可能になる. そのような情報が利用できない場合には, 散漫な事前情報(*diffuse prior*)を用いればよい. これは線形回帰モデルで *non-informative* な事前分布を使う場合と同じである.

C. 予測

$t = 1, 2, \ldots$ のすべての時点におけるパラメータベクトル θ_t の事後分布を得た後で, 将来の観測値 $y_{t+k}(k = 1, 2, \ldots)$ の分布を求める. 予測は分布で表され, 現在のパラメータの不確実性や, 将来の観測値のノイズ分散 V_{t+k}, そして撹乱項分散 W_{t+k} から導かれる. 分布情報から, どんな予測がこの状態では適当かが導かれる. したがって, もしある方向での誤差の結果が, 逆方向での同程度の誤差より重大なら, 予測はこれらを考慮に入れて偏ったものとすることができる.

(6.1) と (6.2) のDLMの式から, t 期での将来観測値は

$$y_{t+k} = F'_{t+k}\theta_{t+k} + v_{t+k} \tag{6.5}$$

$$\theta_{t+k} = G_t\theta_{t+k-1} + w_{t+k} \tag{6.6}$$

と書ける. この式より確率ベクトル v_{t+k}, w_{t+k} の分散を既知とすると, y_{t+k} の将来値の予測にはパラメータ θ_{t+k} と独立変数ベクトル F_{t+k} の予測が必要になる.

$$m_{k|t} = E(\theta_{t+k}|D_t) \tag{6.7}$$

$$C_{k|t} = V(\theta_{t+k}|D_t) \tag{6.8}$$

ここで, $m_{0|t} = m_t = E(\theta_t|D_t)$ と $c_{0|t} = c_t = V(\theta_t|D_t)$ は, カルマンフィルター・アルゴリズムより既知である. (6.6) より

$$m_{k|t} = G_t m_{k-1|t} \tag{6.9}$$

$$C_{k|t} = G_t C_{k-1|t} G'_t + W_{t+k} \tag{6.10}$$

そこで, 将来パラメータの平均と分散は, 既知の初期値 $m_{0|t}$ と $C_{0|t}$ を用いて $k = 1, 2, \cdots$ と順次に計算できる. k 期先の予測値の平均と分散は

$$y_{k\cdot t} = E(y_{t+k}|D_t) \tag{6.11}$$

$$C_{k\cdot t} = V(y_{t+k}|D_t) \tag{6.12}$$

であるが, これを計算するには2つの場合に分けて考える必要がある. F_{t+k} が t 期で既知の場合とそうでない場合である. 後者の場合 (予測するときの外生変数が確率的であるような場合), (6.5) は未知パラメータに関して非線形になる.

D. 平滑化

すべての観測値が得られた後で, t 期以後に得られた観測値を考慮するともっとよい推定量が得られる. このような推定量を計算する方法を平滑化という. これにはいくつかの方法があるが, 固定区間平滑化 (*fixed interval smoothing*) の方法は, 最後に得られたカルマンフィルター推定量 m_T と C_T から出発し, 後ろ向きに平滑化方程式を作用させる (詳しくはAnderson and Moore, 1979 または

Harvey,1981 第4章を参照). $m_{t|T}$ と $C_{t|T}$ をそれぞれ平滑化推定量とその共分散行列を表すと, 平滑化は次のように行われる.

$$m_{t|T} = m_t + C_t^*(m_{t+1|T} - G_{t+1}m_t) \tag{6.13}$$

$$C_{t|T} = C_t + C_t^*(C_{t+1|T} - C_{t+1|t})C_t^{*'} \tag{6.14}$$

ここで

$$C_t^* = C_t G'_{t+1} C_{t+1|t} \qquad t = T-1, \ldots, 1 \tag{6.15}$$

また $m_{T|T} = m_T$ で $C_{T|T} = C_T$ である. このとき得られた平滑化推定量 $m_{t|T}$ は, 最後の観測値を含む T 期までのすべての情報に基づく最適推定量になる. 経済学での応用では, それらは観測されない要素 (*unobserved component*) を表しており, 平滑化とは信号除去 (*signal extraction*) を行う方法と考えられる. Engle and watson (1981, 1987) などの研究で, この平滑化された変数についてさまざまな経済学的解釈がなされている. また, 時系列データがいくつかの要素から構成されているとき, その1つである季節要素を除いて季節調整済み系列を作ることにも使われており, 最近その応用がさまざまな分野で行われている.

6.3 トレンド, 季節性, 回帰成分の DLM

大きな線形モデルは, いくつかの単純な線形モデルの線形結合と考えることができる. モデル・ビルディングにおいては, モデルを各構成要素ごとに作ることは実用的と理論的観点の両面から通常利点が多い. このときには, m 個の成分からなるパラメータ・ベクトル $\theta_t = (\theta'_{1\cdot t}, \theta'_{2\cdot t}, \cdots, \theta'_{m\cdot t})'$ において $(\theta_{i\cdot t}|\theta_{t-1}) = (\theta_{i\cdot t}|\theta_{t-1})$ となる. m 個の成分モデルを考えると, 各モデルはそれぞれ $\{F_{it}, G_{it}, 0, W_{it}\}$ とノイズ部分のモデル $\{0, 0, V_t, 0\}$ で表現される. 各成分DLMは次のような $\{F.G.V.W\}_t$ で表される.

$$F_t = (F'_{1t}, F'_{2t}, \cdots, F'_{mt})'$$

$$G_t = \text{ブロック対角}\,(G_{1t}, G_{2t}, \cdots, G_{mt})$$

$$W_t = \text{ブロック対角}\,(W_{1t}, W_{2t}, \cdots, W_{mt})$$

このような構造にすると, モデルビルディング, 意味づけ, 操作性が簡単になるという利点がある. 大きなモデルは, 最初に解釈しやすい個々の成分を別々にモ

デル化し, 最後にそれらを組み合わせて作ることがしばしば行われる. この定式化によって, 他の成分に対する影響を考えずに, 特定の成分の影響だけを取り入れることができる. 個別の成分モデルに関しては, (1) 全体のモデル式は個々の独立な成分式として表されるが, $\theta_t|D_t$ は各周辺分布 $(\theta_{it}|D_t)$ からは決められない. これは, 観測値系列 y_t を通じて全ての構成部分に関する情報が得られるためである. そして(2) 分解の程度に関して注意が必要である. 1つの成分をもっと細かく分解することができる場合がある. たとえば, 13期の季節要素は理論的には6つの別々の直交する周期部分に分解することができる. しかしながら, これはあまり望ましいことではない. むしろ, 全ての季節性については単一の成分モデルとして扱ったほうがよい. 逆に, ある回帰関係が固定的部分と変動的部分からなっている場合には, 回帰成分についてもさらに分解した方がよい場合がある.

実際の時系列分析や予測でよく用いられるのは, 多項式トレンドモデル, 季節モデル, 回帰モデルを含んだダイナミックモデルである. これら3つの成分モデルの例を以下で挙げておこう. 次のようなスカラーの系列, $y_t, t = 1, 2, \cdots$ を考える.

$$y_t = \mu_{1t} + \mu_{2t} + \mu_{3t} + v_t \tag{6.16}$$

ここで μ_{1t} はトレンド項, μ_{2t} は季節項, そして μ_{3t} は回帰項で, v_t の観測ノイズがある. これらの各項を以下でモデル化する.

A. 線形成長モデル (多項式トレンドモデル) 成分

このモデルは, 単純なトレンドモデルに勾配項を加えたもので, 水準 μ と勾配 α のパラメータがランダム・ウォークにしたがって時間とともにゆっくりと変化し, 次のように表される.

$$y_t = \mu_t + \epsilon_t \qquad \epsilon \sim NID(0, \sigma^2) \tag{6.17}$$

$$\mu_t = \mu_{t-1} + \alpha_t + \eta_t \qquad \eta \sim NID(0, \sigma_\eta^2) \tag{6.18}$$

$$\alpha_t = \alpha_{t-1} + \xi_t \qquad \xi \sim NID(0, \sigma_\xi^2) \tag{6.19}$$

このとき, μ_{1t} に対する線形成長の成分DLMは, 次のように表される.

$$\{F_1, G_1, 0, W_1\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, 0, W_{1t} \right\} \tag{6.20}$$

ここで, パラメータ・ベクトル $\theta_{1t} = (\mu_t, \alpha_t)$ は

$$\mu_{1t} = \mu_t = F_1' - \theta_{1t} \tag{6.21}$$

$$\theta_{1t} = G_{1t}\theta_{1\cdot t-1} + w_{1t} \tag{6.22}$$

と書ける. μ_t は t 期でのトレンド水準を表し, α_t は $t-1$ 期と t 期の間での成長の増分を表す.

B. 加法的季節モデル成分

季節性を示すプロセスは経済, 産業, 社会などのデータではごく普通に見られるものであり, 予測には重要な因子である. 季節性を表すモデルにはいくつかある.

(a) 単純な季節効果成分

季節性の影響は, 1年間での季節サイクル P の各期間に対して加法的な季節項ベクトルを対応させたモデルで表される. s_t をある特定の t に対する期(月)とすると, モデルは

$$y_t = \gamma_{s(t)\cdot t} + \epsilon_t \tag{6.23}$$

$$\gamma_{i\cdot t} = \gamma_{i\cdot t-1} + w_t \quad i = 1, \ldots, P \tag{6.24}$$

となる. 確率的な季節モデルでは, 季節効果の和はゼロになるとの制約は

$$\sum_{i=0}^{P} \gamma_{i\cdot t} = 0, \quad t = 1, \cdots, P \tag{6.25}$$

と表される. 単純な形の季節効果モデルをDLMでは

$$y_t = (0, \cdots, 1, \cdots, 0)\gamma_t + \epsilon_t \tag{6.26}$$

$$\gamma_t = \gamma_{t-1} + w_t \qquad w_t \sim NID(0, \sigma_w^2) \tag{6.27}$$

と表す. これは多項式モデルと比較すると, F 行列は s_t 番目の成分だけが1でほかは0, そして時間とともに変化する.

(b) フーリエ型季節性

季節性を扱うためには季節効果パラメータ $(\gamma_{1t}, \cdots, \gamma_{Pt})$ への変換が必要であるが, 操作性とモデル化という観点からは季節効果をいくつかの調和波, あるいはコサイン波を組み合わせた項でモデル化するのが望ましい. このようにモデル化することの利点は, (1) 簡単に季節制約を扱える, (2) 季節要素をさらにいくつかの副要素に分解することが可能, (3) 一定数の調和系列で簡潔に季節性を表現することができる, という点にある. この場合, もしそれが適切なものなら予測精度が上がるが, 不適切なら季節パターンの形は誤ったものになる. 奇数区間, $p = 2q + 1$(q は整数), の任意の季節効果パターンの DLM $\{F_2, G_2, 0, W_2\}$ 表現は, $\lambda = 2\pi/p$ として q の部分調和 DLM $\{(1,0)', H_i, 0, W_i^*\}, j = 1, \cdots, q$ を組み合わせて表すこともできる. ここで

$$H_j = \begin{pmatrix} \cos(j\lambda) & \sin(j\lambda) \\ -\sin(j\lambda) & \cos(j\lambda) \end{pmatrix} \tag{6.28}$$

したがって, $F_2 = [(1,0), (1,0), \cdots, (1,0)]'$
$G_2 =$ ブロック対角 $[H_1, H_2, \cdots, H_\alpha]$
$W_2 =$ ブロック対角 $[W_1, W_2, \cdots, W_\alpha]$

フーリエ表現では, パラメータ・ベクトル θ_2 は $P-1$ より少ない項でモデル化することもできる.

C. 回帰モデル成分

説明変数が X_t の値をとる1つの時系列を考えると, その成分 DLM $\{F_{3t}, G_{3t}, 0, W_{3t}\}$ は $\{X_{3t}, 1, 0, W_{3t}\}$ となり, パラメータ θ_{3t} を使うと

$$\mu_{3t} = X_{3t}\theta_{3t} \tag{6.29}$$

$$\theta_{3t} = \theta_{3t-1} + w_{3t} \tag{6.30}$$

となる. 一般に r 個の説明変数, $X_{it}(i = 1, \cdots, r)$ があるとき, 各々は1つの成分DLMと, r 個の成分を組み合わせて作られている全体の回帰DLMによって表される. 計量経済分析では, 線形回帰モデルの係数を過去のデータに当てはめて推定する. このモデルからたとえある関係が得られたとしても, その時点での係数の値が時間とともに将来も変化しないと保証できる理由はない. もし短い

期間から推定された関係ならば, 何年か後には全く変化してしまうかもしれない. 従来のモデルでは, この状況は捉えることが難しいが, 動学モデルでは「動く関係」をある程度表すことができる. ここで用いられるカルマン・フィルターの計算方法は, 新しいデータが利用可能になったときに最初から全部の計算をやり直したり, 逆行列の計算をする必要がないので, 回帰係数の推定値の修正を経済的に行えるという利点がある. この方法は, 以前 Plackett (1950) によって指摘されていたが, その実用上の重要性はほとんど見過ごされている.

6.4 DLMの予測と更新

本節では, 実際の予測に必要な主要結果と更新関係を, DLM$\{F_t, G, V, W_t\}$ の場合, $(G_t = G, V_t = \sigma^2)$ について述べる. より詳細な説明や, 多変量, 非正規, 非線形, その他の時間に依存するモデルへの一般化について参考文献を参照してほしい. 実用上の重要な概念は, 情報割引の考え方である. これは各成分DLMと一緒に用いると特に便利である.

以下で用いるHarrison-Stevensのモデルでは, 構造方程式の攪乱項の分散は事前に既知であると仮定されている. またB項で述べる情報割引という定式化が, ちょうどHolt-Wintersのモデルにおける平滑仮定数と同じ形で導入されている. これはかなり恣意的な方法であるが, 非常に短いデータを用いる場合には有効な方法である. また, Harrison-Stevensの推定方法はあまり効率的ではないので, 以下では, 未知の平均と分散を持つ正規分布に対する同時共役事前分布に基づく, より効率的な方法 (DeGroot[1970] を参照) を用いている.

A. 主要結果

2.Aのモデル (6.1) と (6.2) の仮定の下で, DLM$\{F_t, G, V, W_t\}$ は次のように表される.

$$y_t = F_t'\theta_t + v_t \tag{6.31}$$

$$\theta_t = G\theta_{t-1} + w_t \tag{6.32}$$

ここで, $\mathrm{V}(v_t) = V_t = \sigma^2, \mathrm{V}(w_t) = \sigma^2 H_w$ とすると, 初期時点において σ^2 が既知で

$$\theta_0 \sim N(m_0, \sigma^2 C_0) \tag{6.33}$$

と仮定すると, $t-1$ 時点ではその時点までのデータ D_{t-1} が与えられたときの θ_{t-1} の分布は

$$(\theta_{t-1}|D_{t-1}) \sim N(m_{t-1}, \sigma^2 C_{t-1}) \tag{6.34}$$

と表される. (6.4)の8ステップからなるカルマン・フィルターをベイズ的に解釈した結果は, 次のようにまとめることができる.

y_t が観測される以前では

$$a_t = Gm_{t-1} \tag{6.35a}$$

$$R_t = GC_{t-1}G' + W_t \tag{6.35b}$$

パラメータ・ベクトル θ_t の事前分布は $(y_t|\sigma^2, D_{t-1}) \sim N(a_t, \sigma^2 R_t)$. ここで

$$f_t = F_t' a_t \tag{6.35c}$$

$$Q_t = F_t' R F_t + 1 \tag{6.35d}$$

とおけば, y_t に対する1期先予測分布は $(y_t|\sigma^2, D_{t-1}) \sim N(f_t, \sigma^2 Q_t)$. また

$$A_t = R_t F_t Q_t^{-1} \tag{6.35e}$$

とする. y_t が観測された以後では, 1期先予測誤差は

$$e_t = a_t - f_t \tag{6.35f}$$

となるので, θ_t の事後パラメータ・ベクトルは, 次のように逐次的に更新される.

$$m_t = a_t + A_t e_t \tag{6.35g}$$

$$C_t = (I_t - A_t F_t')R = R_t - A_t Q_t A_t' \tag{6.35h}$$

また事後分布は, $(\theta|\sigma^2 D_t) \sim N(m_t, \sigma^2 C_t)$ となる.

さらに σ^2 が未知の場合の σ^2 の事前分布 $(\sigma^2|D_0)$ は, 次のような逆ガンマ分布 $IG(n_0/2, 2/n_0 S_0)$ で

$$p(\sigma^2|D_0) \sim \sigma^{-(n_0+1)} \exp(-n_0 S_0/2\sigma^2)$$

と仮定しよう. ここで $X \sim IG(\alpha, \beta)$ とは, $Y = X^{-1}$ がパラメータ α と β のガンマ分布に従うという意味である (鈴木[1987]参照). この場合には逐次更新は, 標準的なベイズ分析を利用して計算すれば, 次のように表すことができる.

$$n_t = n_{t-1} + 1 = n_0 + t \tag{6.35i}$$

$$n_t S_t = n_{t-1} S_{t-1} + e_t^2 / Q_t \tag{6.35j}$$

したがって, 分散は $(\sigma | D_t) \sim IG(n_t/2, 2/n_t S_t)$ と分布する. ここで, $w_t = 0$ で $G = I$ とすると, この関係は重回帰モデルにおける自然共役事前分布を逐次的に表したものと考えることができる (たとえばZellner[1971]を参照).

(6.35b)(6.35d)(6.35h) において σ^2 が未知の場合には, 周辺事後分布は標準的なベイズ分析によって導くことができ, 多変量 t の形をしていることがわかる. なお. 将来予測分布と回顧的予測に対する平滑化された分布は次のようになる.

$k = 1, 2, \ldots$ に対する予測分布は, 次式で与えられる. パラメータ分布は

$$(\theta_t | D_t) \sim T_{n_t}[a_t(k), R_t(k)] \tag{6.36}$$

y_t の分布は

$$(y_{t+k} | D_t) \sim T_{n_t}[f_t(k), Q_t(k)] \tag{6.37}$$

ここで $a_t(k) = G a_t(k-1)$, ただし $a_t(0) = m_t$,
$R_t(k) = G R_t(k-1) G' + W_{t+h}$, $R_t(0) = C_t$,
$f_t(k) = F'_{t+1} a_t(k)$, $Q_t(k) = F'_{t+h} R'_t(k) F_{t+h} + S_t$
である. 回顧的予測に対する平滑化された分布は, $k = 1, 2, \cdots$ に対して

$$(\theta_{t-k} | D_t) \sim T_{n_t}[a_t(-k), R_t(-k)] \tag{6.38}$$

ここで $a_k(-k)$ と $R_t(-k)$ は次式のように逐次的に求められる.

$$a_t(-k) = m_{t-k} + C_{t-k} + C_{t-k} G' R^{-1}_{t-k+1} [a_t(-k+1) - a_{t-k+1}] \tag{6.39}$$

$$R_t(-k) = \{C_{t-k} - C_{t-k} G' R^{-1}_{t-k+1} \cdot [R_{t-k+1} = R(-k+1)] \cdot R^{-1}_{t-k+1} G C_{t-k}\} S_t / S_{t-k} \tag{6.40}$$

B. 情報割引

m 個の部分DLMに対しては, W_t =ブロック対角 $(W_{1t}, W_{2t}, \cdots, W_{mt})$ であり, 各 W_{it} は, Ameen and Harison[1985] による割引法を用いて決められる. 各パラメータ θ_{it} で表される第 i 成分に1つの割引率 $0\ leq\delta_i \leq 1$ が対応する. 各成分の事後周辺分布は

$$(\theta_{i,t-1}|D_{t-1}) \sim T_{n_{t-1}}[m_{i,t-1}, C_{i,t-1}] \tag{6.41}$$

であり, 対応する周辺事前分布は

$$(\theta_{it}|D_{t-1}) \sim T_{n_{t-1}}[a_{it}, R_{it}] \tag{6.42}$$

ここで $a_t = G_i m_{t-1}, R_{it} = G_i C_{t-1} G_i' + W_{it}, W_{it} = (\delta^{-1} - 1) G_i C_{t-1} G_t'$ とおいた. したがって, $R_{iy} = G_i C_{t-1} G_i'/\delta_i$ となる. そこで, θ_t に対する事前分布は

$$(\theta_t|D_{t-1}) \sim T_{n_{t-1}}[a_t, R_t] \tag{6.43}$$

となり, ここで $a_t = (a_{1t}', \ldots, a_{mt}')'$, そして $R_t = GC_{t-1}G' + W_t$ である. 簡単には, 将来何期か先の予測では, システム分散 W_{t+k} の将来値は, $W_{t+k} = W_{t+1}$ とおかれる場合が多い. この方法に関する一般的な考え方は, D_{t-1} の下でパラメータ・ベクトル θ_{t+k} に伴う精度は k とともに減少し, θ_t の精度は θ_{t-1} の δ の割合になるというものである. この意味で精度は δ の割合で割引される. $0 \leq \delta_t \leq 1$ と $G_i = 1$ での回帰成分を考えると, $\delta_i = 1$ ならば $W_{it} = 0$ となり, 将来の $t+k$ に対する$\theta_{i,t+k}$ についての情報は $\theta_{i,t-1}$ と同じになる. したがったて, θ_{it} の値は期間中一定であり, i 番目の成分に関するシステム方程式は決定論的になる. 他方, $\delta_i = 0$ の場合は $\theta_{i,t-1}$ についての知識は将来値 $\theta_{i,t+k}$ についての情報を全く含んでいないので, それに伴う成分システムは必要なくなる. これらの極端な値の中間では, D_{t-1} は δ_i が小さいほど i 番目の成分ベクトルについての情報を含んでいないことを意味する. したがって, 割引率 δ_i は i 番目の成分ベクトルに関する値 $[m_{it}, C_{it}]$ が各機関でどのくらい意味を持ち, 永続性があるかということを意味している. モデル化に際して割引率を1に近くするということは, 決定論的なモデルの法を不確実なモデルより好むという意味から一般的である. しかしながら, 適切でない決定論的モデルは, モデルがデータから乖離してもそれに適応しないので, 予測は惨めな結果になることがある.

DLMにおいて, モデル $\{F_t, G, V, W_t\}$ がシステム分散行列 W_t に関して定式化される場合, 何か目安がないと実際に適用するときに問題が生じる. W_t は説明変数の計測単位に対して不変ではないので, この値の決め方は明確ではない. そこでこのような割引という概念を導入し, それを成分モデルと結びつけることによって, 単純な割引・指数加重回帰概念に基づいたモデル作成が可能になる. この考え方は, 最近の観測値をより重要視するという意味からも直感的に受け入れやすい方法に見える. モデルがトレンド・季節・回帰の各DLMという3つの成分からなる場合では, 別々の割引率を3つの各成分に関係づけることができる. 通常, 観測値系列の変動は, 制御可能であるような独立変数にすべて帰することが出来る方がよい. したがって, 大きく変動する回帰成分よりも安定的な回帰成分のあるモデルモデルを使う方が望ましい. また, トレンド季節モデルでは, 季節成分はトレンドモデルよりもずっと安定的であるから, 季節に対する割引係数はトレンドに対するものよりも高い値をとる場合が多く見られる.

6.4.1 インターベンション分析

政策変更等の出来事についての外部情報に基づくインターベンションがあると, 系列にはそれがはっきりと表れ, モデル式が実際と大きく乖離してしまう. 実際には重要であるがモデルの構造に含まれていない事態が発生したとき, それをうまく取り込んで連続的に予測できることが必要である. インターベンションは, それが前もって分かっている場合でも分からない場合でも次の形をとる. 第1は, はずれ値(outlier)で, ある観測値がモデルの1期先予測値から著しく離れているが, 他はモデルとそう違っていないような場合である. もう1つは構造変化で, モデルのあるパラメータが突然変化したために起こる. たとえば, 水準が落ちたり, 季節パターンの山と谷がずれたり, 観測誤差の分散が増大したためにデータに表れる形である. インターベンションを処理するには, 次の2つの方法がある. 第1にモデルを拡張して θ_t にいくらか加えるか, あるいは標準的なDLMにその乖離を説明する部分モデルを加え, F と G をそれに従って修正する. 第2に, はずれ値や特定の変化を現在のモデルで説明できるものとして考える. 原因は何であれ, 大きくはずれた観測値は標準的な方法で処理すると, 分析に悪影響を及ぼす場合がある. 時間 t を越えて予測するとき, はずれ値を扱う最

も良い方法はそれを無視することである．したがって，更新分析は D_{t-1} の下での現在の事前分布を，y_t を無視して D_t の下での事後分布として考えるように修正される．すなわち，$m_t = a_t, C_t = R_t, n_t = n_{t-1}, S_t = S_{t-1}$ とする．構造変化に関しては，季節パターンや回帰項は，外生的な影響に対して安定的であるので，突然変動した水準や成長についての情報をさまざまな方法で直接モデルに含めるようにする．ベイズ予測の特徴は，このような例外的な事態に対する主観的情報をごく自然な方法で取り入れることが出来る点にある．

はずれ値を除いたり分散が大きくなる場合をモデルに含めるという考え方は，それを発見し，識別するモデル監視の方法と結びつけると，例外的な事態を自動的に調整する方法として使うことが出来る．この方法は，1) 主観的な予想インターベンションが適切な外部情報がなかったために作動しなかったときは自動的に無視し，2) 短期予測への応用では，多数の系列を並行的に分析できる単純で頑健なモデルの採用が可能で，3) はずれ値や構造変化点がモデル推定に影響を及ぼすような時系列分析に役立つものである．もう1つの方法は，多段階モデルに基づいたもので同じ様な能力があるが，比較的小さなモデルでさえもかなり計算費用がかかるという欠点がある．ルーティン的なモデル推定，データ平滑化，短期予測には単純なものの方がよい．自動化の方法には，はずれ値や構造変化を識別することによって予期せざる事態を発見する，はずれ値を捨て変化に対応することによって予期せざる事態を処理する，という2つの要素がある．前者に対しては，モデルの予測パフォーマンスを比較的一般的な対立モデルと比べて一連のベイズファクターを使って評価する．しかし，乖離が生じたときそれは構造変化の始まりによるものかもしれない．その場合に，その時点で小さな割引ファクターを適用することによってはずれ値を落とすと，トレンド，季節性，そして回帰に対する不確実性が増大してしまう．大きなベイズファクターは構造変化が最近起こったことを示しているが，それは大きな予測誤差によるためかもしれない．そこで，そのような変化は，次のような累積ベイズファクターによって調べることが出来る．

$$L_t(k) = H_t H_{t-1} \cdots H_{t-k+1} = H_t L_{t-1}(k-1) \quad (k = 1, 2, \cdots, t) \tag{6.44}$$

これは，モデルの予測パフォーマンスを対立モデルのそれと，それ以前の連続する k 期の観測値について比べてみたものである (詳細は West[1986) を参照).

6.5 DLM分析の応用例

前節までに述べたDLMモデルによるベイズ予測の方法を, いくつかの例によって示すことにしよう. 実際の分析では, モデルの定式化と $t=0$ における初期事前分布, そして割引率の値を決める必要がある. $t=1,\cdots,T$ と表される等時間間隔で観測された1変量時系列データをモデル化する. 各時点 t で, y_t はレベル (μ_t) と観測誤差あるいはノイズ項 (v_t) という2つの成分の和として表される.

$$y_t = \mu_t + v_t \tag{6.45}$$

レベルはさらに, 局所的になめらかな多項式トレンド (p_t), 季節性 (s_t), 回帰項 (r_t) の3つの成分からなっている. したがって, 加法モデルでは

$$\mu_t = p_t + s_t + r_t \tag{6.46}$$

乗法モデルでは

$$\mu_t = (p_t + r_t)s_t \tag{6.47}$$

の形をとる. 時系列は観測誤差, トレンド, 季節, 回帰の各成分に分解される. 誤差項は平均ゼロの正規確率変量で, レベルに関して純粋に確率的で予測できない変動を表している. この誤差項は過去の情報 D_t とは独立である. レベルに対する変動の程度は誤差分散によってコントロールできる.

$$V[v_t = c\mu_t^{\alpha} \tag{6.48}$$

ここで α は観測誤差分散のスケールパラメータで, α は分散のべき乗部分を表す指標である. $\alpha=0$ なら一定分散であり, $\alpha>0$ なら系列のレベルとともに確率誤差の分散が増加する. トレンド項は多項式の次数で表し, 常に存在し通常1次あるいは2次の多項式が用いられる. 周期が P である季節性をモデル化すると, s_t は t 期での季節因子を表す. 加法モデルでは, P 個の季節因子はレベルからの季節的偏差を表し, すべての t で P 個の季節因子は識別のために和がゼロとされる. 乗法モデルでは, 季節因子は季節乗数であり, 和が P となるような制約が課される. 回帰項については, r_t が独立変数に対する線形回帰になり, β_t を t 期での回帰係数とすると $r_t = \beta_t X_t$ となる. 回帰係数にはきわめて小さな

確率変動を認め, そのために添え字tがつけられている. したがって, モデル関係は固定的でなく, モデルに定期化の誤りを認めることになる.

t期における過去の情報D_tはモデルの各成分を表すダイナミックなパラメータに対する事後分布で要約される. これらはモデルの各部分に対しては, ステューデントt分布にしたがい, cに対しては基準化された逆カイ2乗分布(より一般的には逆ガンマ分布)にしたがうとする. 時間が$t+1, t+2, \cdots$と経るにつれて各要素は変化し, 新しいデータが得られると情報セットを更新しながらこれらの事後分布を更新する. これらの次々と得られる各成分の事後点推定値(各時点tに対する事後モード)は, パラメータに対する推定された時間経路を表している. 事後分散はこれらの推定値における不確実性の尺度になる. 同じ様な軌跡は, 多項式トレンド, 回帰効果, 季節性の逐次推定値があれば, 時系列の各成分に対して同様に得られる. たとえば, 回帰係数β_tの軌跡は$(\beta_t|D_t), t = 1, \cdots, T$の連続する事後$t$分布で決められ, このモード, スケール, 自由度は時間がたつにつれて連続的に更新される. モデル評価にとって重要なのは, 1期先予測誤差が作られる1期先予測分布の系列と, それに伴うt予測確率密度である.

回顧的な評価は, 系列の最後$t = T$ですべての利用できるデータを使ってから, モデルのパラメータの平滑化された推定値とすべての情報D_Tのもとで, 各時点tでの平滑化分布から導かれる成分をもとに行う. したがって, 回帰係数の軌跡の回顧的評価はt分布$(\beta_t|D_t)$をもとにしている. 最後に, 予測値はいつの時点でも求められる. t期での将来観測値$y_{t+k}(k > 0)$の予測分布は, D_tで表されるモード, スケール, 自由度を持つ標準のt分布の形をしている.

データを処理する以前のD_0という条件の下で, 事前分布の初期値をモデルの各成分に与える. 事前分布は, 事前モードあるいは事前平均と不確実性の尺度の推定値を各パラメータに対して与えればよい. これらの事前分布は, 自由度n_0のt分布と, 分散スケールパラメータcに対する逆カイ2乗分布あるいは一般的には逆ガンマ事前分布である. 不確実性の尺度はモデルパラメータの標準誤差とする. ただし, これは事前分布が基本的に正規分布の時はこれらの値はn_0が非常に大きくない限り実際の標準偏差ではない. その他の場合はt事前分布のスケール・パラメータを与えることになり, 実際の標準偏差は$\sqrt{n_0/(n_0-2)}$を掛けて得られる. 観測スケールcに対する逆カイ2乗分布の初期値は, 点推定

値と自由度 n_0 を与えると決まる.

モデルパラメータが時間とともに確率的に変化する率が割引率で, モデルの各要素ごとに決められる. レベル, 成長, 季節, 回帰パラメータにおける確率変化の程度は, 各時点でのこれらの値について得られる情報に直接関係している. すなわち, 観測値間での確率的な動きは各パラメータに対する事後分散の増加, すなわち情報の減少に等しくなる. したがって, 0.95という割引率は約5%の情報の減少, あるいは分散で測った不確実性の5%増大を意味する. この割引率が1というのは, 対応する部分に時間を通じて確率変動がない, あるいはその部分のパラメータがスタティックであることを意味している. 以下で, このような数値の例とその計算結果をグラフ表示したものを示す.

ケースI

データは, 食料品の輸入数量指数(外国貿易概況)と円ベースでの食料品輸入価格指数(物価指数月報)の系列である. これらのデータは標準的なベイズ推定をどのように行うかを示すための例で, 1973年第1四半期から1987年第4四半期までの四半期データである. 散漫な事前分布を用いた分析で特定の事前分布を考える前の参考となるような予測を行う. 正則(*proper*)な情報量のある(*informative*)事前分布を用いると, 短い時系列データからも予測可能であり, 短期予測パフォーマンスが向上する.

最初に各系列のデータ分析を行う. 各系列について, 平均, 分散, モードなどの要約統計量ならびにヒストグラムなどからデータの分布状況を見る. さらに, 自己相関関係をチェックすることも必要である. 必要ならデータ変換を行う. データがどのような動きをしているかを見るために, 系列をプロットする(図1). データのプロットを見ると, 食料品(数量)にはっきりとした季節性が認められる. 両変数間の散布図(図2)を見ると負の関係が見られる. そこで, この例での食料品に対するモデルは次のような構造をしていると考える. 一定の(ステディな)トレンド成分, 加法的季節要素の成分, そして価格変数に対する回帰成分である. このモデルの加法的季節モデル要素は周期が4で, 調和項は2つである. 最後にモデルの各要素についての割引率を決める. この例では, レベル(水準) 0.90; 季節0.97; 回帰0.95; 分散0.99 としている. モデルのパラメータに対して情報がない(*non-informative*), すなわち散漫な事前分布を基に行われた分析は図3にあ

る．観測値は順に処理され，各時点で1期先予測分布が計算される．モデルに関する事後分布の要約とその要約統計量は表1にあり，最後の53個の観測値を基に計算される．



図1 原系列のプロット

次に，このモデルのパラメータに適切な事前分布として，次の値を与えてみる．

レベル(水準):平均70，標準偏差10
季節ファクター：平均-10, 2, 4, 4 で，標準偏差はすべて2
価格変数の回帰係数：平均 -0.20 で，標準偏差は0.2
分散：推定値0.15，自由度 5

ここで，標準偏差を大きくとると散漫な事前分布の場合に近いものになり，不確実性の範囲が比較的広いものになる．標準偏差は名目的な正規事前分布で定義されるにもかかわらず，事前分布は実際は t 分布であり，したがって正規性で考えるよりいくらか散漫になる．この事前分布に基づいて当てはめられたモデルの事後分布の要約とその要約統計量は表2の通りであり，その1期先予測分布は図4にある．散漫な事前分布によるものより予測パフォーマンスはいくらか改良されていることが分かる．次の図5から図6はモデルの各要素の1期ずつのデータを順に処理したオンライン推定値を表しており，その時点までに処理し

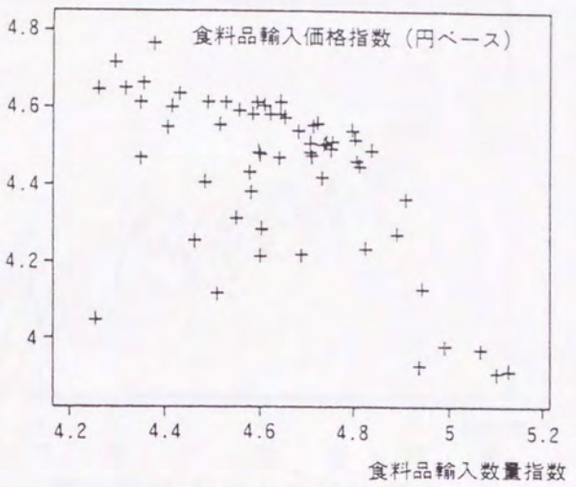


図2 散 布 図

成 分	平 均	標準偏差
レベル（水準）	169.326	6.497
季節要素	-10.964	2.161
	2.766	2.112
	4.697	2.083
	3.501	2.252
価格変数	-0.211	0.117
平均2乗誤差		46.40
平均絶対誤差		5.430
対数予測密度		-2173

表1 事後分布の要約表(散漫な事前分布)

成 分	平 均	標準偏差
レベル（水準）	176.673	8.829
季節要素	-12.797	1.433
	2.973	1.361
	4.285	1.416
	5.540	1.524
価格変数	-0.248	0.173
平均2乗誤差		38.05
平均絶対誤差		4.899
対数予測密度		-2982

表2 事後分布の要約表(報知的事前分布)

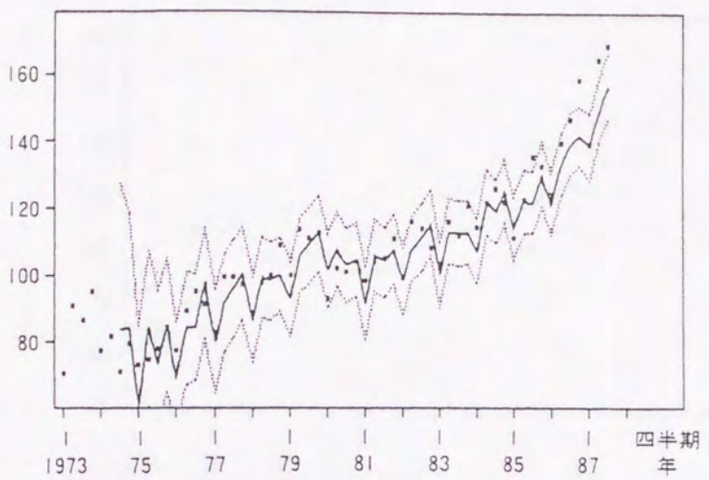


図3 1期先予測(散漫な事前分布)

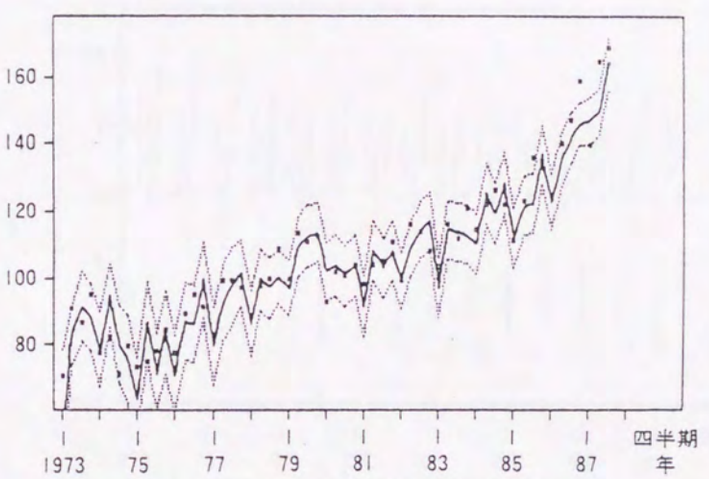


図4 1期先予測(報知的事前分布)

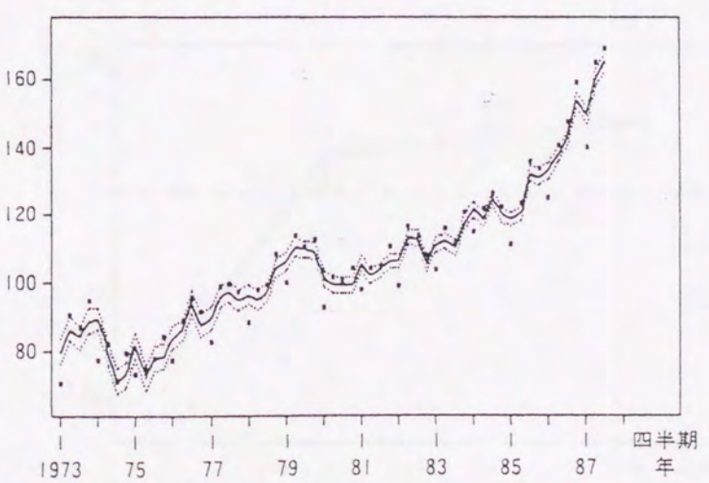


図5 推定されたオンライン・レベル

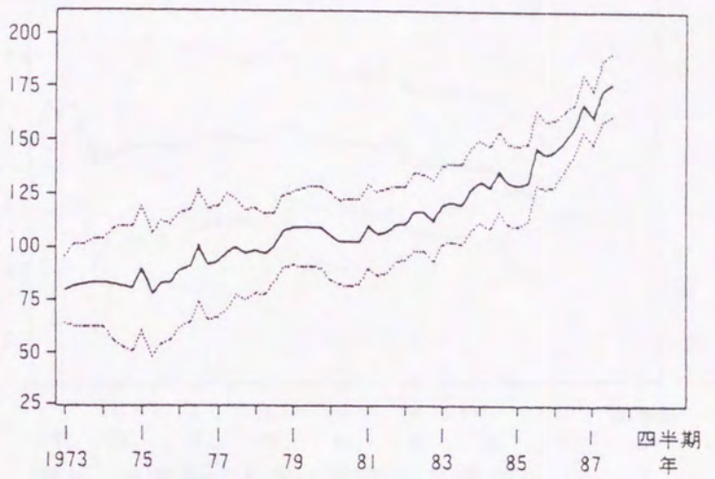


図6 推定されたオンライン・回帰項とレベル

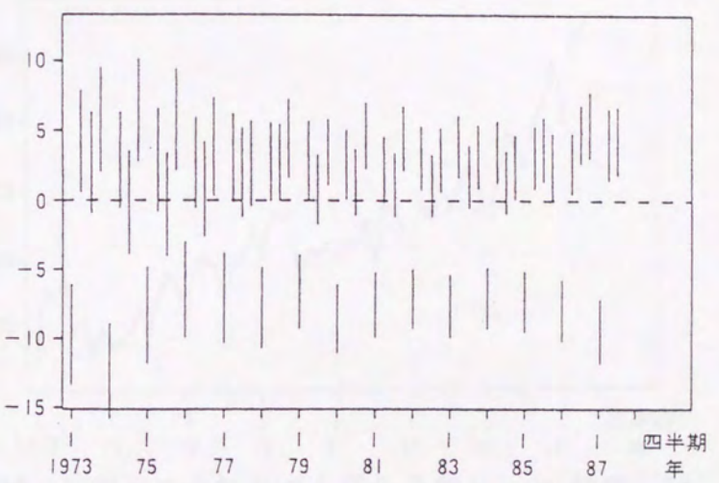


図7 推定されたオンライン・季節項

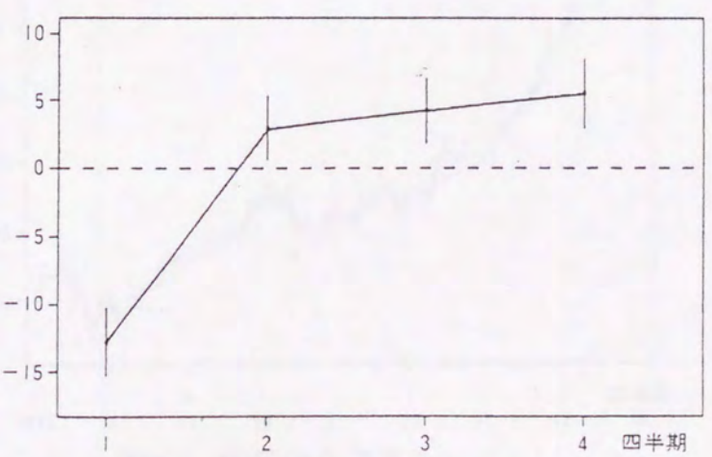


図8 推定された季節ファクター

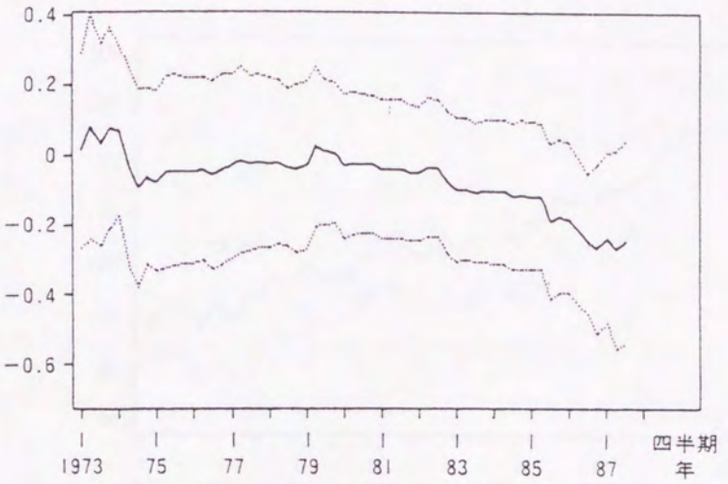


図9 推定された価格変数の係数プロット

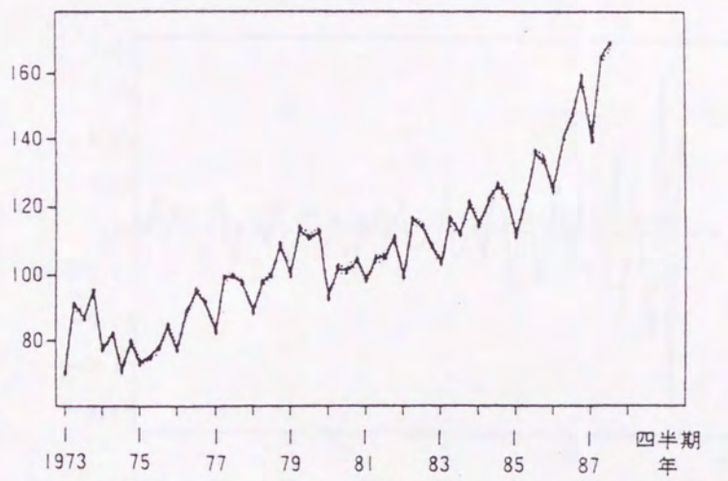


図10 回顧的な食料品の1期先予測と90%信頼区間

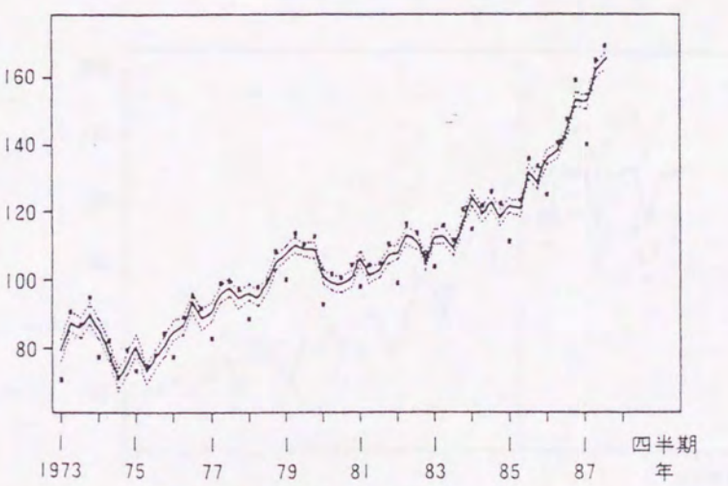


図11 回顧的な平滑化レベル

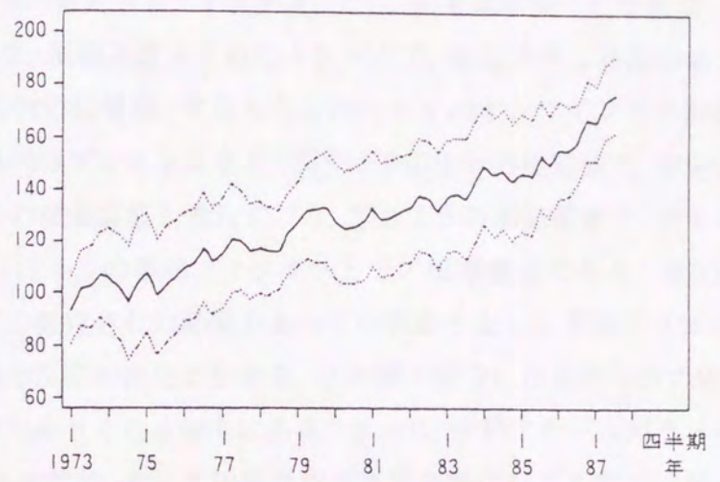


図 12　回顧的な平滑化回帰項とレベル

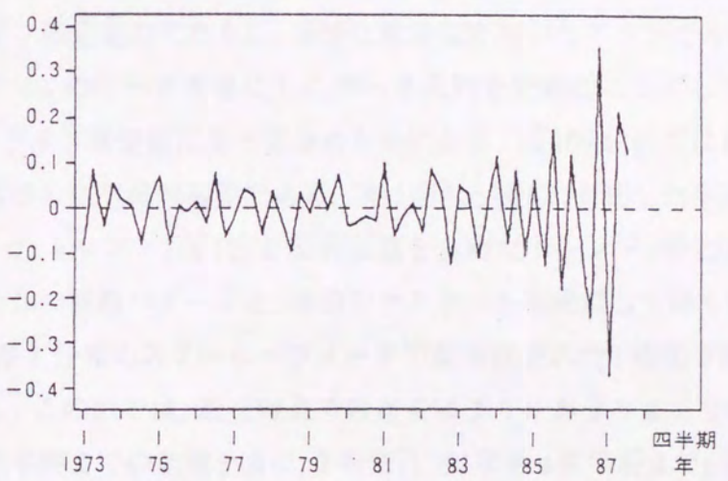


図 13　回顧的な平滑化標準化残差

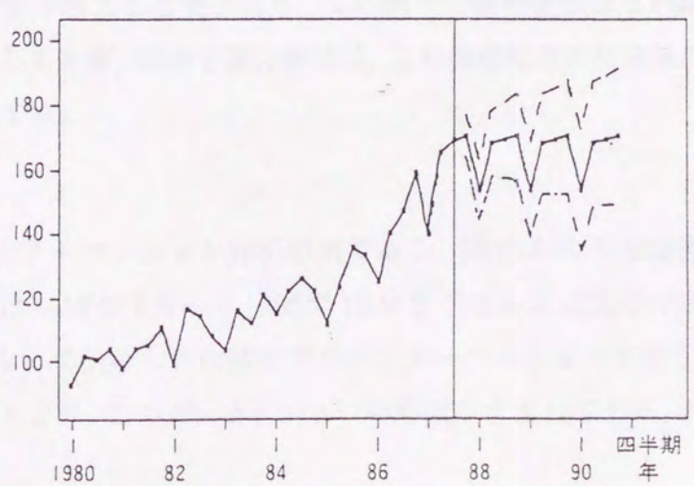


図 14　食料品データの将来予測

たデータを基に計算した各時点での推定値である．図5は説明変数を除いたトレンドで，モデルの推定された1次多項式の成分(各観測値での定数項)のプロットである．図6は，説明変数を含めたトレンドで，推定された各観測値での定数項プラス価格変数の回帰項，すなわちこのモデルのオンラインでの計算された季節調整済み系列のプロットになる．図7は季節成分の推定値で，推定された季節パターンとその標準偏差を表している．図8はその季節要素で，データから推定された最後の12ヶ月の季節ファクターとその標準偏差である．図9は価格変数に対する係数の推定された時間をおっての軌跡を表した係数プロットであり，各係数の安定性や反応の変化が分かる．この例の場合には価格係数の係数が，最近では僅かながら小さくなる傾向にある．さらに，分散スケールパラメータの推定値や，1期先予測誤差，そして回帰係数の従属変数に対する毎期毎期の影響を評価することによってモデルを改良することができる．

次はオンライン推定値の代わりに，平滑化推定値を用いたグラフである．平滑化推定値は，すべてのデータを基にしたデータ系列を回顧的に改訂した点推定値であり，オンライン推定値に比べてなめらかになる．図10は，当てはめられた値と t 分布に基づく90%確率限界である．次は図5と図6に対応した平滑化した説明変数を除いたトレンド(図11)と説明変数を含めたトレンド(図12)である．平滑化された全体の季節パターンと，季節ファクターも同様にして得られる．図13は，1期先予測 t 分布のスケールパラメータで標準化された1期先予測誤差のプロットである．この例では，最近時点での当てはまりがあまりよくない．

1987年第4四半期までの数値を基に，3年先(1980年第4四半期まで)を予測して得られた予測値が図14である．モデルは回帰項を含んでいるので，将来3年間の価格変数の値を与える必要がある．この例では価格変数はこの期間変化しないとして計算してある．将来予測分析では，この価格変数の将来値の値をいろいろ変えて分析する．

ケースII

次の例は，インターベンション分析の例である．図15のような販売データを用いる．データは1955年1月から1960年12月までのある商品の月次販売高である．最初は前もって決められた時点でのインターベンションを行う例である．データのプロットより，このデータには3つの局面が含まれており，各局面では

はっきりとしたレベルの変化があることが分かる. この例では, 予測する人にはさらに次の付加的な事前情報があったと仮定しよう.

- 1955年12月に, 記録ミスがあった.
- 1957年1月にはレベルは上昇したが成長率は変わっていないように見える.
- 1958年1月には1年前と同様な動きがあるようだが, 今度は不確実性が増している.

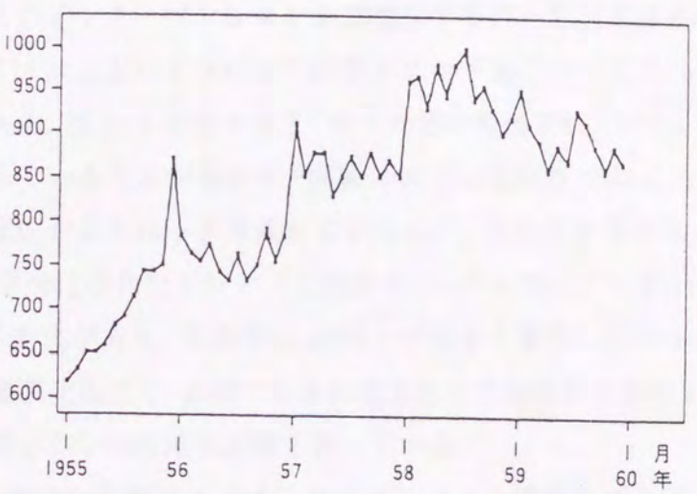


図15 販売データのプロット

モデルは, トレンドの部分を線形(レベルと成長)とする. 初期の事前分布は, レベル平均が600, で標準偏差が100; 成長平均が0で, 標準偏差が10; 分散推定値が1.5%で自由度は5である. 割引ファクターは, レベルが0.90で分散を0.98としている. ここで, 次のようにしてこの情報をモデルに含める.

(1) モデルを定式化する段階で, 3つのあらかじめ分かっているインターベンションの時点を, 1955/12, 1957/1, 1958/1 としてモデルに含める.

(2) モデルを当てはめていく段階で各々のインターベンションの時点の手前で分析を止め, インターベンションを確認する. このデータを分析に用いるか, はずれ値としてモデルから外すかを決める.

(3) 最初の1955/12の時点では, この観測値を欠測値として分析を進める. 1期先予測をプロットした図17では, この点は欠測値となっている.

(4) 2番目のインターベンション時点, 1957/1, ではこれを分析に含め, 次の事前分布を設定する. レベル平均は900で, 標準偏差が20；成長の平均が0で, 標準偏差が5である.

(5) 最後のインターベンション時点, 1958/1, では事前情報を次のように変更する. レベル平均が940で, 標準偏差が50；成長平均が−2で, 標準偏差が7である.

図16は, このインターベンションを無視してモデルを当てはめたときの予測値であり, 図17は上記のようにして計算された予測に対するインターベンションの結果である. 図16と比較すると, はずれ値は無視され, レベル変化ははっきりと処理されていることが分かる. 回顧分析では変化点でのインターベンションに対する反応がよりはっきり表れているので, その方を示すことにする. 図18と図19は平滑化されたトレンドと成長で, レベルではインターベンション時点で2回上昇変化があり, 成長率は1957/1で大きく変化しているのが分かる. 図20はその標準化残差で, 点線で示された上側と下側限界はその t 分布からの, ゼロの回りで対称な90%確率区間を表している.

次に, 自動的な予測監視とインターベンション機能を組み込んだ分析の例である. この例では, 同じ事前情報を用いている. 主な違いは割引ファクターに関するもので, モデル監視機能が指示したときに, 新しいデータを素早く取り入れるかどうかを決める. 自動監視の場合の割引ファクターは, レベルを0.10, 分散を0.90としてある. 図21はこの分析による1期先予測で, 図の横軸にモデル監視機能からの印が付けられている. 図22は自動モデル監視機能を動かした場合での回顧分析であり, 図23と図24は前の分析での図19と図20に対応するものである. このモデルでの1期先回顧予測誤差は図25に, そして1年先までの予測値は図26にある.

6.6 多変量モデルへの一般化

前節までに述べた単変量DLMを多変量時系列データを取り扱うモデルに拡張し, 一般化することができる. 多変量時系列モデルは, Tiao and Box(1981),

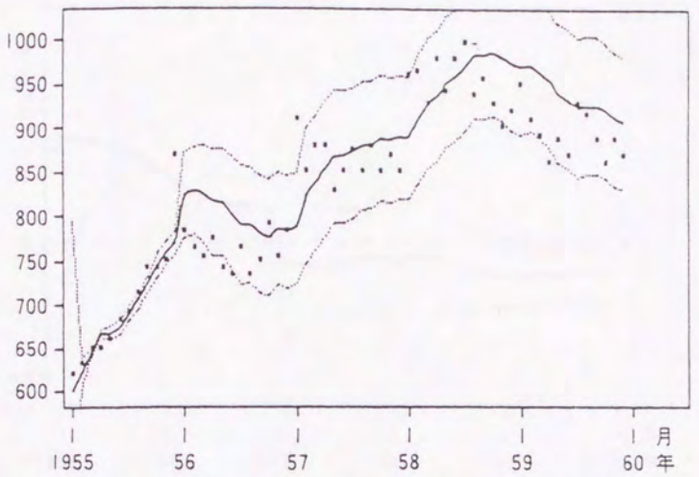


図 16　販売データの 1 期先予測と 90% 信頼区間

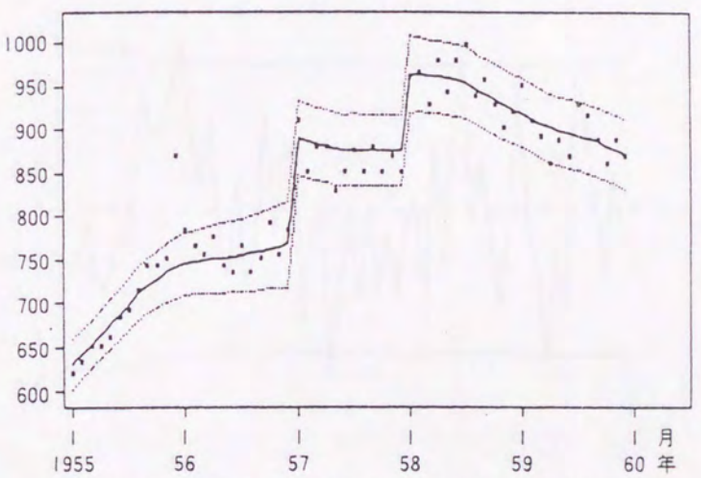


図 17　回顧的な 1 期先予測と 90% 信頼区間

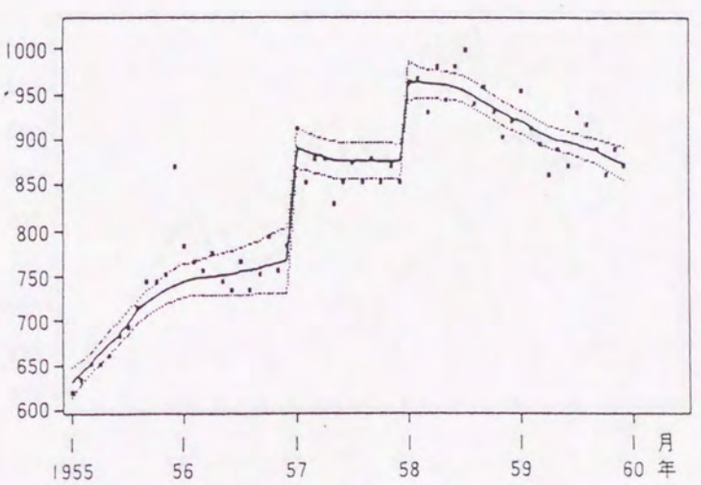


図 18　回顧的な平滑化レベル

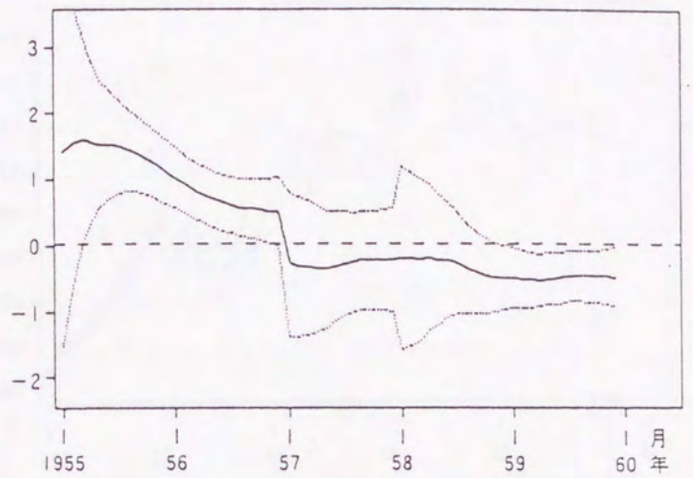


図 19 回顧的な平滑化成長率

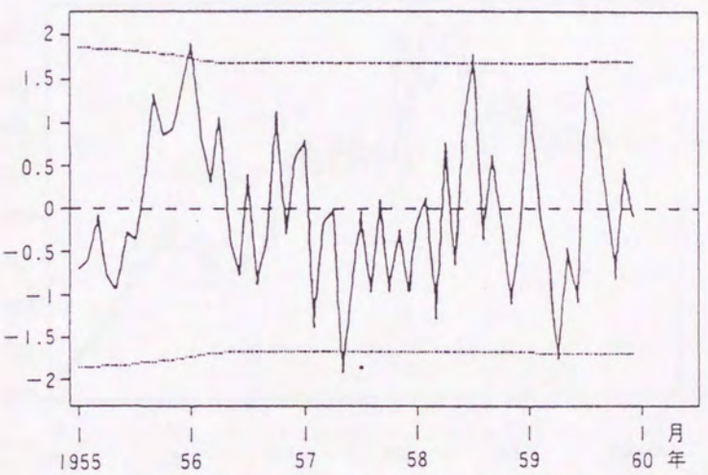


図 20 回顧的な平滑化標準化残差

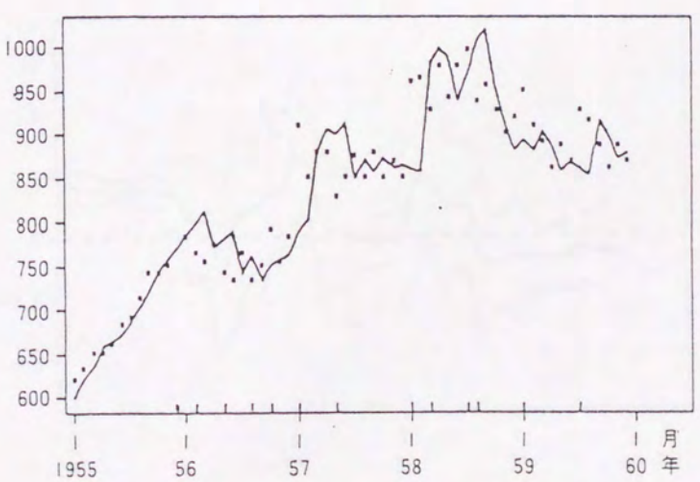


図 21 自動的予測監視による 1 期先予測

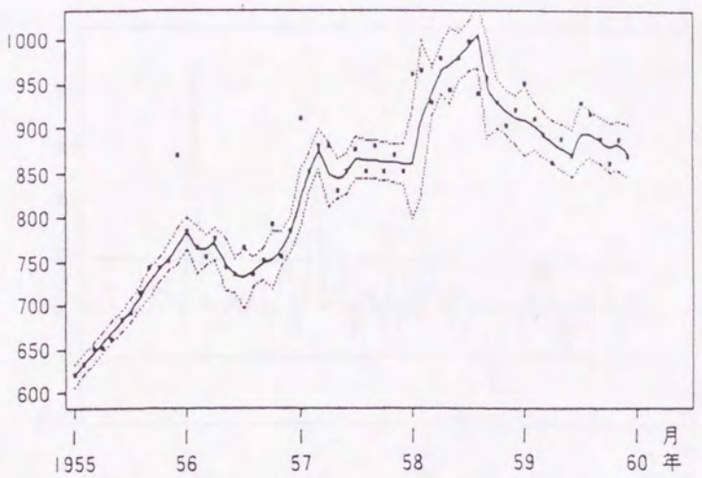


図 22　回顧的な販売データの 1 期先予測(自動的)

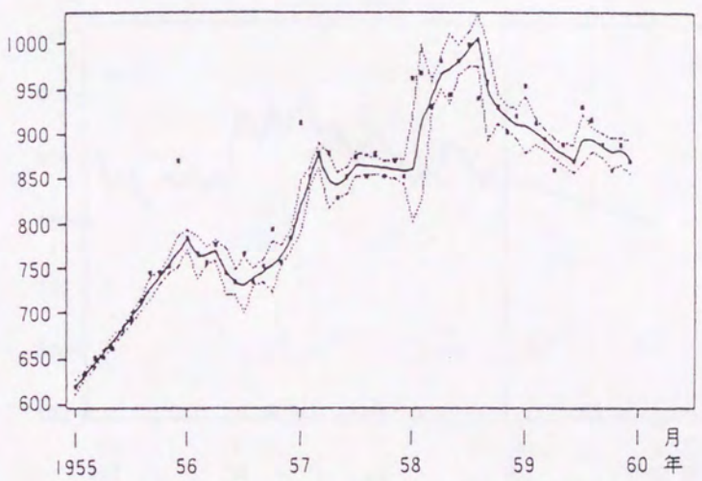


図 23　回顧的な平滑化レベル(自動的)

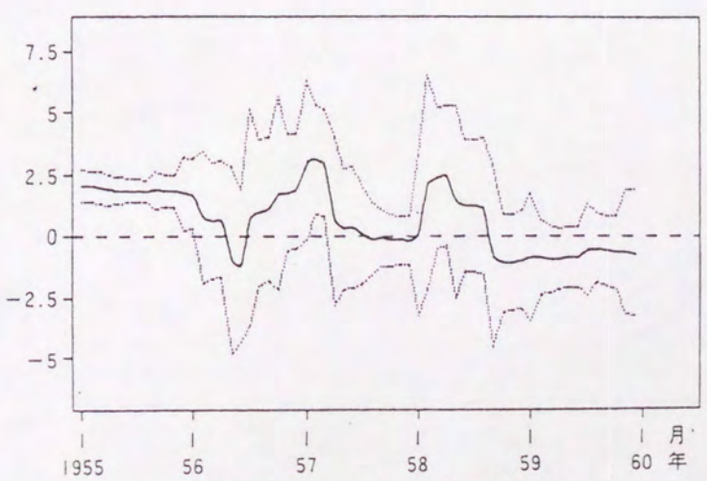


図 24　回顧的な平滑化成長率(自動的)

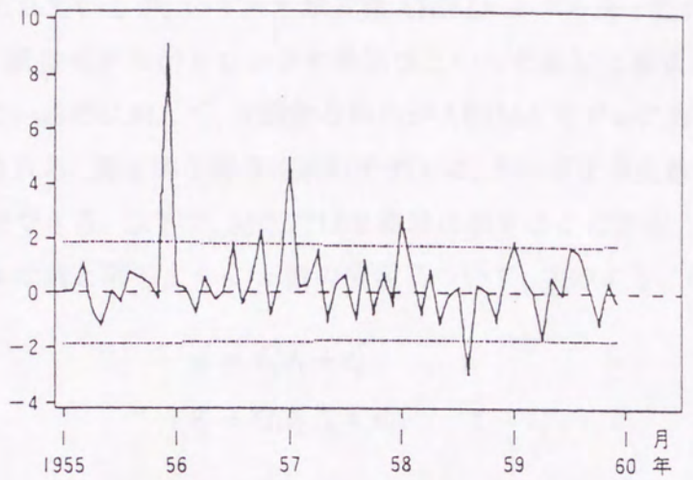


図 25 回顧的な平滑化標準化残差(自動的)

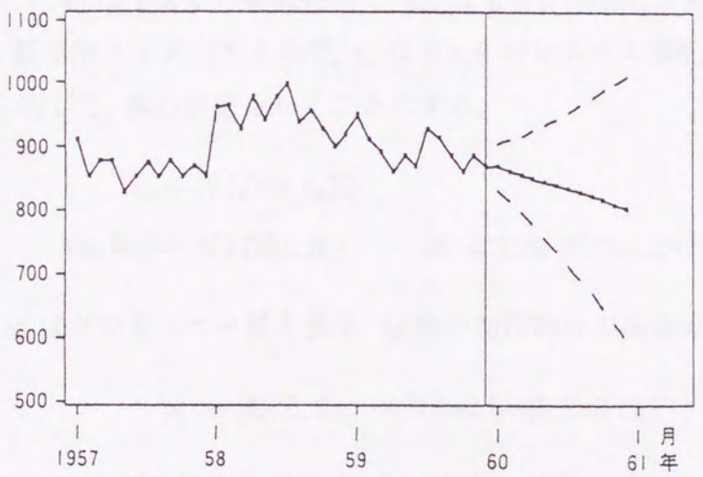


図 26 販売データの将来予測(自動的)

Chan and Wallis (1978), Jenkins and Alavi (1981) そして Abraham (1980) などで扱われているが, いずれも単変量ARIMAモデルを一般化したものである. これは, 構造モデルがトレンドや季節性といった直接に解釈可能な成分で定式化されているのに対して, 先験的な制約がARIMAモデルにおかれるので誘導型と見なされる. 基本的な構造時系列モデルは, そのまま多変量の場合に一般化することができる. 以下で, 結果だけを簡単に示すことにする.

モデルは前と同じように k 個の変数について, 次のように表す.

$$y_t' = F_t'\theta_t + \epsilon_t' \tag{6.49}$$

$$\theta_t = G_t\theta_{t-1} + u_t, \qquad t = 1, \cdots, T \tag{6.50}$$

ここで, y_t は k 個の従属変数についての $k \times 1$ ベクトルで, F_t は $p \times 1$ 説明変数ベクトル, θ_t は $p \times k$ の係数行列で, G_t は $p \times p$ の既知の定数行列である. ϵ_t は $k \times 1$ 観測値ノイズベクトルで, u_t は $p \times k$ のシステム攪乱項行列である. 誤差分布に関して, 次の仮定をおくことにする.

$$\epsilon_t \sim NID(0, v_t\Sigma) \tag{6.51}$$

$$\mathrm{vec}\,[u_t] \sim NID(0, H_t), \qquad H_t = \Sigma \otimes W_t \text{で} \epsilon_t \text{とは独立} \tag{6.52}$$

ここで, $\otimes$ はクロネッカー積を表す. 係数の初期時点の事前分布は

$$p(\mathrm{vec}\,[\theta_0|\Sigma, D_0) \sim N(\mathrm{vec}\,[m_0], \Sigma \otimes C_0) \tag{6.53}$$

Σ の $k(k+1)/2$ 個の異なる要素は, 次のような逆ウィッシャート分布に従うと考える.

$$p(\Sigma|D_0) \propto \det(\Sigma)^{-\eta/2} \exp(-\mathrm{tr}\,[\Sigma^{-1}S_0]/2) \tag{6.54}$$

ここで, $\eta = n_0 + k$ であり, S_t は $k \times k$ の対称行列である. これらの仮定の下で単変量の場合と同様にして多変量カルマンフィルター逐次計算は次のように表される.

(i) t 期で y ベクトルを観測する以前

$$a_t = Gm_{t-1} \tag{6.55a}$$

$$R_t = GC_{t-1}G' + W_t \tag{6.55b}$$

したがって, $p(\mathrm{vec}\,[\theta_0|\Sigma, D_{t-1}) \sim N(\mathrm{vec}\,[a_t], \Sigma \otimes R_t)$

$$f_t = F_t' a_t \tag{6.55c}$$

$$Q_t = v_t I + F_t' R F_t \tag{6.55d}$$

したがって, $(y_t|\Sigma, D_{t-1}) \sim N(f_t, \Sigma \otimes Q_t)$

$$A_t = R_t F_t Q_t^{-1} \tag{6.55e}$$

(ii) t 期で y ベクトルを観測した以後

$$e_t = a_t - f_t \tag{6.55f}$$

$$m_t = a_t + A_t e_t' \tag{6.55g}$$

$$C_t = (I - A_t F_t') R_t \tag{6.55h}$$

したがって, $(\mathrm{vec}\,[\theta]|\Omega, D_t) \sim N(\mathrm{vec}\,[m_t], \Sigma \otimes C_t)$.

$$n_t = n_{t-1} + 1 = n_0 + t \tag{6.55i}$$

$$S_t = S_{t-1} + e_t e_t' / Q_t \tag{6.55j}$$

したがって, $(\Sigma|D_t) \sim IW(S_t, n_t)$ にしたがって分布する.

v_0 と行列 θ_0, C_0, S_0 と $W_t, t = 1, \cdots, T$ のパラメータを決める必要がある. Σ の $k(k+1)/2$ 個のユニークなパラメータに対する多変量散漫事前分布は, v_0 をゼロ, S_0 をゼロ行列とすると得られる. 多変量構造モデルにおける個々の時系列データは, 同じ形の k 本の単変量構造モデルを考えればよい. k 本の単変量モデルは, これらの方程式の観測誤差 e_{it} が時間に関して独立ならば, 多変量システムとして書くことができる. これらの誤差間の同時点相関は, Σ の非対角要素に表れる.

単変量の場合と同様に周辺事前分布, 予測密度, 事後密度を導出することができる (Zellner[1971], Ch.8 参照). 周辺事前分布は

$$p(\theta_t|D_{t-1}) \propto (\det[S_{t-1} + (\theta_t - a_t)' R_t (\theta_t - a_t)])^{\gamma/2} \tag{6.56}$$

ここで, $\gamma = n_{t-1} + p$, すなわち一般化多変量 t 分布に従う. 1期先予測密度は

$$p(y_t|D_{t-1}) \propto [Q_t + (y_t - f_t)' S_t' (y_t - f_t)]^{-(\eta+k)/2} \tag{6.57}$$

ここで, $\eta = n_{t-1} - K + 1$, すなわち多変量 t 分布に従う. 事後密度は

$$p(\theta_t|D_t) \propto (\det[S_t + (\theta_t - m_t)'C_t(\theta_t - m_t)])^{\tau/2} \tag{6.58}$$

ここで, $\tau = n_t + 1$ である.

任意の t に対して $W_t \neq 0$ のとき, (6.49) と (6.50) のモデルは係数が時間で変化するケースになる. これはモデルを新しい状態に対してある意味で適応するようにする1つの方法である. これを示すために, 多変量での (6.55g) 式を見てみよう. これらの方程式ではパラメータの事後分布の平均は, 1期先予測誤差の関数を事前分布の平均に加えることによって得られる. 事前分布の平均がこれらの誤差によって修正される大きさは, しばしばカルマン・ゲインと呼ばれる A_t に依存している. (6.55c) と (6.55e) における f_t と A_t の定義を用いて, 事後平均 m_t を次のように表すことができる.

$$m_t = (R_t^{-1} + F_tF_t')^{-1}(F_tF_t' + R_t^{-1}a_t) \tag{6.59}$$

この式は, 事前精度が増加するにつれて, すなわち R_t^{-1} の対角要素が大きくなるにつれて, 事前平均 a_t は m_t を求める際にデータを支配 (dominate) し, 極限では $m_t = a_t$ となることを表している. 逆に言えば, 事前精度が低下するにつれて, t 期おいて観測される新しいデータはより重要になる. (6.55b) より W_t が大きくなれば, R_t^{-1} は小さくなることが分かる. したがって, ゼロでない W_t は, 古典的な固定パラメータの線形モデルの場合における重要性に比べて過去のデータの重要性を割り引く効果を持っているといえる. 前述したように, Ameen and Harrison (1985) ではこの割引法を (6.50) 式の u_t の共分散行列と結びつけた W_t 行列を割引要素行列と置き換えたモデルを用いている. このモデルは W_t を次のように定義すると, 単変量の場合と同様に, いわゆる正規割引 (*normal discount*) ベイズモデルの形になる.

$$W_t = \Delta_t^{-\frac{1}{2}}C_{t-1}\Delta^{-\frac{1}{2}} - C_{t-1} \tag{6.60}$$

ここで Δ_t は $0 \leq \Delta_{ii.t} \leq 1$ のような割引要素からなる対角行列である. (6.60) 式を (6.55b) に代入すると

$$R_t = \Delta_t^{-\frac{1}{2}}C_{t-1}\Delta^{-\frac{1}{2}} \tag{6.61}$$

となり, $\Delta_t = I$ のとき $W_t = 0$ である. 応用に際しては, 通常 W_t を直接与える代わりに, すべての t に対して $\Delta_t = \Delta$ とおかれる.

6.7 おわりに

ここで述べられたベイズ予測モデルは, 「基本的構造モデル」を基に定式化されている. このモデルを使ってある変数についての将来分析と回顧分析を行うことによって, 意志決定と予測が関連する「もしこうしたら, どうなるか?」という分析と, 「なにが起こったのか?」という分析をすることができる. このような回顧的な分析を行う意味は, 単にある出来事の影響を記述するだけでなしに, それから学びとることによって似たような将来の状況に対する対処の仕方を身につけることができることにある. この学習という側面は, この種の予測システムでは非常に重要であり, モデルを改良し適切なものへ変えることも目的の1つであるといえる. そしていろいろな人とコミュニケーションをするためには, 結果を数値情報だけでなくグラフィックに表示することも重要である. 多変量モデルへの一般化の方法についての概要は述べたが, 応用はまだあまりされていない. 多変量時系列モデルの応用では, 多くの研究で利用されているSims(1980)により提唱されたVARモデルが経済予測の分野でその実用性が認められているこのモデルに基づいたLitterman(1981)によるベイジアンVARモデルは, 制約を穏やかに導入することによって, 予測に有効なモデルを提供しようとする方法である. VARモデルは一種の誘導型方程式ともいうべきものであるので, 個々のパラメータについての経済学的解釈は明確でない. したがって, ある意味では, Litterman (1980, 1984) や Doan, Litterman and Sims (1984) で開発されたベイズVARの方法は, 前節までに扱った構造状態空間モデルと競争的なモデルとして見ることができるが, 潜在的にはマクロ経済予測という目的には利用価値があるように思える (Litterman (1986) を参照). このベイズVARモデルを状態空間モデルの形に書き直して用いると, Litterman他によって採用されたVARの方法をもっと明確にし, 拡張し, ある意味では単純化することができる. しかしながら, このベイズVARにはいくつかの問題点がある. 紙数の関係でその詳細を述べることはできないが, たとえば係数行列への事前分布の与え方や, パラメータの推定そしてモデル選択の問題があげられるが, それに関し

ては, Litterman (1980) や Doan, Litterman and Sims (1984) などでいくつかの方法が提案されている. また, モデルをもっと一般的で実用的にするためには, 高速で有効な数値計算の方法とグラフィック表示によるコミュニケーションの方法が必要になろう.

参考文献

1. Abraham, B., (1980), "Intervention Analysis and Multiple Time Series," *Biometrika,* 67,73–78.

2. Ameen, J.R.M. and Harrison, P.J., (1985), "Normal Discounted Bayesian Models," In J. Bernardo, M.H. Degroot, D.V. Lindley and A.F.M. Smith, eds. *Bayesian Statistics 2*, North-Holland, Amsterdom.

3. Anderson, B.O.D. and Moore, J.B. (1979), *Optimal Filtering,* New Jersey, Prentice-Hall.

4. Box, G.E.P. and Jenkins, G.M. (1976), *Time Series Analysis: Forecasting and Control,* Holden-Day, San Francisco.

5. Chan, W.Y.T. and Wallis, K.F. (1978), "Multiple Time Series Modelling: Another Look at the Mink-Muscrat Interaction",*Applied Statistics,* 27, 168–175.

6. DeGroot, M.H. (1970), *Optimal Statistical Decisions,* McGraw-Hill.

7. Doan, T., Litterman, R. and Sims, C. (1984), "Forecasting and Conditional Projection Using Realistic Prior Distributions (with discussion)," *Econometric Reviews*, 3-1, 1-144.

8. Engle, R.F. and Watson, M. (1981), "A One Factor Multivariate Time Series of Metropolitan Wage Rate," *Journal of American Statistical Association,* 76, 774–781.

9. Engle, R.F. and Watson, M. (1987), "The Kalman Filter: Application to Forecasting and Rational Expectations Models," Truman F. Bewley ed. *Advances in Econometrics : Fifth World Congress,* 245–284.

10. Gersch, W. and Kitagawa, G. (1983), "Prediction of Time Series with Trend and Seasonalities," *Journal of Business and Economic Statistics,* 1, 253–264.

11. Harrison, P.J. (1967), "Exponential Smoothing and Short-term Sales Forecasting," *Management Sciense,* 13, 102–139.

12. Harrison, P.J. and Stevens, G.F. (1971), "A Bayesian Approach to Short Term Forecasting," *Operation Research Quarterly,* 22, 341–362.

13. Harrison, P.J. and Stevens, G.F. (1976), "Bayesian Forecasting (with discussion)," *Journal of the Royal Statistical Society, Ser. B.,* 38, 205–247.

14. Harvey, A.C. (1984), "Unified View of Statistical Forecasting Procedure," *Journal of Forecasting,* 3, 245–275.

15. Harvey, A.C. and Todd, P.H.J. (1983), "Forecasting Economic Time Series with Structural and Box-Jenkins Models: A Case Study," *Journal of Business and Economic Statistics,* 1-4, 299–315.

16. Jenkins, G.M. and Alavi, A.S. (1981), "Some Aspects of Modelling and Forecasting Multivariate Time Series," *Journal of Time Series Analysis,* 2, 1–48.

17. Kalman, R.E. (1960), "A New Approach to Linear Filtering and Prediction Problems," Transaction of ASME, Series D, *Journal of Basic Engineering,* 82, 35–45.

18. Kitagawa, G. (1981), "A Nonstationary Time Series Model and its Fitting by Recursive Filter," *Journal of Time Series Analysis,* 2, 103–116.

19. Kitagawa, G. and Gersch, W. (1984), "A Smoothness Priors-State Space Modelling of Time Series with Trend and Seasonality," *Journal of American Statistical Association,* 79, 378–389.

20. Kitagawa, G. and Gersch, W. (1985), "A Smoothness Priors Time Varying AR Coefficient Modeling of Nonstationary Covariance Time Series," *IEEE Transactions on Automatic Control,* AC-30, 1, 48–56.

21. Lindley D.V. and Smith, A.F.M. (1972), "Bayes' Estimates for the Linear Model," *Journal of the Royal Statistical Society, Ser. B.,* 34, 1–41.

22. Litterman, R. (1980), "A Bayesian Procedure for Forecasting with Vector Autoregressions," Working Paper MIT.

23. Litterman, R. (1984), "Specifying Vector Autoregressions for Macroeconomic Forecasting," Research Department Staff Report 92, Federal Bank of Minneapolis.

24. Litterman, R. (1986), "Forecasting with Bayesian Vector Autoregression–Five Years of Experience," *Journal of Business and Economic Statistics,* 4, 25–38.

25. Plackett, R.L. (1950), "Some Theorems in Least Squares," *Biometrika,* 37, 149–157.

26. Sims, C. (1980), "Macroeconomics and Reality," *Econometrica,* 48, 1–48.

27. 鈴木 雪夫 (1987) 「統計学」朝倉書店

28. Tiao, G.C. and Box, G.E.P. (1983), "Modelling Multiple Time Series with Applications" *Journal of American Statistical Association,* 76, 802–816.

29. West, M. (1986), "Bayesian Model Monitoring," *Journal of the Royal Statistical Society, Ser. B.,* 48, 70–78.

30. West, M., Harrison P.J. and Migon, H.S. (1985), "Dynamic Generalized Linear Models and Bayesian Forecasting (with discussion)," *Journal of American Statistical Association,* 80, 73–97.

31. West, M. and Harrison P.J. (1986), "Monitoring and Adaptation in Bayesian Forecasting Models," *Journal of American Statistical Association,* 81, 741–750.

32. Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, John Wiley and Sons. (福場庸・大沢豊訳『ベイジアン計量経済学入門』培風館、1984年)