

# 金融工学の概要

田中 佑弥

Version 1.0 (2023/3/31)

## 序文

本稿は、金融工学に関する基本的な事項を抜粋し纏めたものである。

### 【更新履歴】

Version 1.0 2023/3/31 初稿作成.

時系列解析（基礎），債券投資理論の基礎，株式運用の基礎，  
市場リスクの基礎，信用リスクの基礎

## 第I部

# Excelと金融工学の基礎



このパートでは、金融工学の基本的な題材を使いながら Excel の基礎的な関数の習得を目指す。



# 第1章 リターン・リスク、ポートフォリオ理論の基礎

本章では、金融工学を始めるにあたって基本的な内容となるリスク・リターンおよびポートフォリオ理論の基礎を題材に Excel に関する基本操作を演習を通じて習得する。

## 1.1 ポートフォリオのリターン・リスク

株式に投資すれば、株価の値動きや配当によりリターン（投資収益）を得ることができる。しかし、そのリターンは確実に得られるというものではなく、不確実性を有する。そのため、個別の証券における不確実なリターンの取り扱いについて説明したのち、ポートフォリオのリターンやリスクについて解説する。

### 1.1.1 単一証券のリターン・リスク

株式のリターンは、株価の値動きによって得られる**キャピタルゲイン**と、企業から株主に支払われる配当など**インカムゲイン**に分けることが出来る。ある証券の時点  $t-1$  から  $t$  の間のリターン  $R_t$  は

$$R_t = \underbrace{\frac{P_t - P_{t-1}}{P_{t-1}}}_{\text{キャピタルゲイン}} + \underbrace{\frac{D_t}{P_{t-1}}}_{\text{インカムゲイン}}$$

で表される。ここで、 $P_t$  は時点  $t$  における証券の価格、 $D_t$  は時点  $t$  における証券からの配当を表す。

我々は時点  $t-1$  から  $t$  にかけてのリターンに興味があるわけだが、 $t-1$  時点において  $t$  に価格や配当がどのようになるかを予め知ることが出来ないため、確率変数だということとする。そのことを明示するべく  $\hat{P}_t, \hat{D}_t$  と書くと、先ほどの式は次のようになる：

$$\hat{R}_t = \underbrace{\frac{\hat{P}_t - P_{t-1}}{P_{t-1}}}_{\text{キャピタルゲイン}} + \underbrace{\frac{\hat{D}_t}{P_{t-1}}}_{\text{インカムゲイン}} .$$

この  $\hat{P}_t, \hat{D}_t$  が分からない以上、証券のリターン  $\hat{R}_t$  も確率的にしか分からない。そこで、証券のリターンにある種の確率分布を仮定し、リターンの期待値を考えよう。証券の期待リターンは、シナリオを複数考え、リターンがある値となるシナリオとその確率をもとに

$$\mu = \mathbf{E}[\hat{R}_t] := \sum_{i=1}^N \hat{R}_{t,i} p_i$$

で与えられる。ここで、 $N$  は考え得るすべてのシナリオの数で、起こり得るリターンを  $\hat{R}_{t,i}$ 、リターンが  $\hat{R}_{t,i}$  となる確率を  $p_i$  で表す<sup>1</sup>。

上式によって計算される期待リターンが等しければ、どのような証券に投資しようと変わらないかといえばそうではない。例えば、安定的な業種と不安定な業種のように、その証券のリターンのばらつき度合いは証券ごとに異なる。そこで、リターンのブレの大きさを証券のリスクと考え、リターンのブレの尺度を与えよう。一般的には標準偏差が用いられる：

$$\sigma = \sqrt{\text{Var}[\hat{R}_t]}.$$

これまでの、各証券のシナリオごとのリターンを予測して期待リターン・リスクを求めてきたが、これらのシナリオを立てることは難しく、仮にシナリオを作ったとしてもそこに恣意性が含まれること、そう捉えられる可能性があることも少なくない。そこで、ヒストリカルなリターンデータをもとに期待リターンやリスクを推計する手法をよく使う。本テキストにおいても、基本的にはこちらの手法で各種分析を行うことを想定していく。

過去の月次リターンが10か月分 ( $R_1, R_2, \dots, R_{10}$ ) あるときに、期待リターン ( $\mu$ )・リスク ( $\sigma$ ) は

$$\mu = \sum_{i=1}^{10} R_i, \sigma = \sqrt{\frac{10}{10-1}} \cdot \sqrt{\frac{1}{10} \sum_{i=1}^{10} (R_i - \mu)^2}$$

で計算される。リスク ( $\sigma$ ) においては、標本から母集団の標準偏差を推定するため、標本標準偏差を用いている<sup>2</sup>。

### 1.1.2 ポートフォリオのリターン・リスク

$N$  個の証券を組み合わせてポートフォリオを作成した場合、ポートフォリオ全体のリターンはココの証券のリターンを組入れ比率 (ウェイト) で加重平均したものとなる。一方、ポートフォリオのリスクは個々の証券についてウェイト加重したものとはならず、証券間の相関が影響を与える。

$N$  個の証券のリターンを  $\hat{R}_1, \hat{R}_2, \dots, \hat{R}_N$  で表し、組入れ比率を  $w_1, w_2, \dots, w_N$  で表す (なお、 $w_1 + w_2 + \dots + w_N = 1, w_i \geq 0$  に注意)。また、証券  $i$  と証券  $j$  のリターンの間の相関係数を  $\rho_{ij}$  とする。このとき、ポートフォリオの期待リターン  $\mu_P$  とリスク  $\sigma_P$  は夫々、

$$\mu_P = \mathbf{E}[\hat{R}_P] = \sum_{i=1}^N \mathbf{E}[\hat{R}_i] \cdot w_i,$$

$$\sigma_P = \sqrt{\text{Var}[\hat{R}_P]} = \sqrt{\mathbf{E}[\hat{R}_P - \mathbf{E}[\hat{R}_P]]^2} = \sqrt{\sum_{i=1}^N \sum_{j=1}^N \sigma_i \sigma_j \rho_{ij} w_i w_j}$$

となる。

<sup>1</sup>有限の場合で考えているが、連続的な分布があればその確率密度を用いた積分の形で書けるので本質的には変わらない。

<sup>2</sup>詳細は統計学の教科書などで「不偏性」についてを参照。



2 証券のポートフォリオを考える場合、計算式の入力もさほど手間ではないが、証券の数が 3, 4, ... と増加するとポートフォリオの分散

$$\sum_{i=1}^N \sum_{j=1}^N \sigma_i \sigma_j \rho_{ij} w_i w_j$$

を計算式として入力するのはなかなか大変になってくる。そこで、これを行列演算を用いて計算できるようにしよう。  $\sigma_{ij} := \rho_{ij} \sigma_i \sigma_j$  を共分散という。

$$\mathbf{w} := (w_1, w_2, \dots, w_N), \boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_N), \Sigma := (\sigma)_{ij} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_N^2 \end{pmatrix}$$

とおくと<sup>3</sup>

$$\mu_P = \langle \mathbf{w}, \boldsymbol{\mu} \rangle, \sigma_P^2 = \mathbf{w} \Sigma \mathbf{w}^\top$$

で計算できる。

ポートフォリオを構成する際、銘柄の組み合わせ・ウェイトは無数に考えられる。すると、その中には同じ期待リターンを持ち、最もリスクが低いポートフォリオや（或いは同じリスクで最も期待リターンの高いポートフォリオ）が存在する。このようなポートフォリオを**効率的ポートフォリオ**といい、効率的ポートフォリオの集合を**効率的フロンティア**と呼ぶ。

効率的ポートフォリオは以下のように最小化問題を解けば求めることが出来る。ポートフォリオのリスク（分散でもボラティリティでも良い）を目的関数  $\theta(\mathbf{w})$  とし、これを以下の制約条件

1.  $\sum_{i=1}^N w_i \mathbf{E}[\hat{R}_i] = C,$
2.  $w_i \geq 0, w_1 + w_2 + \cdots + w_N = 1.$

の下で最小化すれば、ポートフォリオのリターンが  $C$  である効率的ポートフォリオを求めることができる。あとは、この  $C$  を考え得るパターンで計算していけば、効率的フロンティアを得ることができる。Excel においては、ソルバーを用いることでこの最小化問題をコンピュータに解かせることができる。

### 1.1.3 ポートフォリオの選択

期待リターンが同じで、リスクの水準が異なる証券が 2 つある場合、投資家はどちらを選択するだろうか。投資家のリスクに対する考え方に応じて、投資家は 3 つのタイプに分類できる：

- (A) リスクの低い証券を選ぶ = リスク回避的な投資家
- (B) リスクの高い証券を選ぶ = リスク選好的な投資家
- (C) どちらでもよい = リスク中立的な投資家

期待リターンとリスクを 2 軸にプロットし、リスクの許容度の異なる投資家の効用が等しくなるような期待リターンとリスクの組み合わせを結んだ曲線を**無差別曲線**と呼ぶ。

<sup>3</sup>このとき、 $\Sigma$  を分散共分散行列と呼ぶ。

1. 同じ期待リターンであれば、よりリスク水準の低いポートフォリオを選ぶ。
2. 同じリスク水準であれば、より期待リターンの高いポートフォリオを選ぶ。
3. 効用が等しい場合、リスクが増加するに従い期待リターンも増加する。
4. 効用を等しくするために、投資家が要求する期待リターンの増加分は、リスク水準が高くなるに従い大きくなる。

投資家の効用が最大になるポートフォリオを考える。ここでは(1)リスク資産しか存在しない場合、(2)無リスク資産を含む場合の2つを順を追って考える。

### リスク資産しか存在しない場合のポートフォリオ選択

投資家がリスク回避的であると仮定すると、この投資家は自分の無差別曲線と効率的フロンティアが接するポートフォリオ  $P$  を選択する。

効率的フロンティアより内側のポートフォリオにも投資することは出来るが、リスク回避的な投資家はそのようなポートフォリオには投資しない。実際、そのようなポートフォリオよりも同一リスクで高い期待リターンのポートフォリオが効率的フロンティア上にある。また、コウリツテキフロンティアと無差別曲線の交点にあたるポートフォリオも選択されない。なぜなら、効率的フロンティア上でより効用の高い点が存在するからである。

投資家によって無差別曲線の形状は異なるため、選択されるポートフォリオも投資家によって異なる。

### 無リスク資産も存在する場合のポートフォリオ選択

無リスク証券のリターンを  $R_f$  とし、無リスク証券を含めてポートフォリオを組む場合を考える。このとき、無リスク証券とあらゆる証券やポートフォリオの組み合わせが新たに投資可能集合となる。

標準偏差を  $x$  軸、期待リターンを  $y$  軸として、点  $R_f : (0, R_f)$  から効率的フロンティアに向かって接線を引き、効率的フロンティアとの接点を  $T$  とする。このとき、同じ期待リターンで最もリスクの小さいポートフォリオは直接  $R_f T$  上に存在する。 $T$  以外のリスク資産と  $R_f$  を組み合わせても直線  $R_f T$  より下に位置するポートフォリオはしか作ることが出来ない。よって、効率的ポートフォリオは、すべて  $R_f$  と  $T$  の組み合わせで作られていることが分かる。無リスク証券を加えた場合に、次のことがわかる：

1. 投資家はリスク資産として必ず接点ポートフォリオ  $T$  を選ぶ。
2. 接点ポートフォリオと無リスク証券の保有ウェイトは、投資家のリスク回避度によって決まる。

全くリスクをとりたくない投資家は無リスク資産だけを購入するであろうし、もう少しリスクを取りたいならば、無リスク資産のウェイトを接点ポートフォリオに移すこともあるだろう。もし、さらに追加的にリスクをとりたい投資家があり、金利  $R_f$  で無制限に借入可能であれば、レバレッジを効かせてリスク資産へ投資することによって、直線  $R_f T$  の延長線上にあるポートフォリオを選択することが出来る。

このことを纏めたものが、**ポートフォリオ選択の分離定理**である。

**定理 1.1** (ポートフォリオ選択における分離定理). ポートフォリオ選択におけるすべての意思決定は：

**Step 1** リスク資産だけを対象に最適な投資比率の危険資産ポートフォリオ（接点ポートフォリオ）を作成

**Step 2** 接点ポートフォリオと安全資産の投資配分の比率を決める

という2つに分離できる。

無リスク資産のある場合の効率的フロンティアの導出は、

$$\begin{aligned} & \text{maximize } \theta := \frac{\mathbf{E}[\hat{R}_p] - R_f}{\sigma_P}, \\ & \text{subject to } \sum_{i=1}^N w_i = 1 \end{aligned} \quad (1.1)$$

によって定式化される。

#### 1.1.4 下方リスクの考え方と Value at Risk

これまでは、投資におけるリスクを標準偏差によって計測してきた。一方、リターンが上にぶれるのはリスクではなく、あくまで下方へのブレだけをリスクとしてとらえる方がより自然なリスクの考え方にも思える。このようなリスクの考え方を下方リスク（ダウンサイドリスクという。）

##### 下方リスクの例

個別証券やポートフォリオのリターンがある一定の利率を下回る確率を**ショートフォール確率**という。リターンが正規分布に従うとき、正規分布の累積密度関数によって容易に計算することが出来る。

保有ポートフォリオの現在価値を  $V_0$ 、将来のある時点  $t$  におけるポートフォリオの価値変動額（あるいは率）を  $\Delta V_t$  とする。この  $\Delta V$  がある水準  $\ell$  を下回る確率が  $\alpha\%$  であるとき、この  $\ell$  をこのポートフォリオのホライゾン  $T$  での確率  $(100 - \alpha)\%$  での **Value at Risk (VaR)** という。数学的に記述すれば

$$\mathbf{P}(\Delta V \leq \ell) = \alpha \quad (1.2)$$

を満たす  $\ell$  である。

## 1.2 CAPM

### 1.2.1 投資家の期待効用最大化

効用がリターンによって決まるので、効用はポートフォリオのリターン  $R_P$  の関数で  $U(R_P)$  の形で考えられる。この関数  $U(\cdot)$  を**効用関数** (utility function) と呼ぶ。

**仮定 1.** 効用はリターンによって決まり、効用関数は連続で少なくとも2階までは微分可能であると仮定する。リターンが大きくなれば、効用は高まるので  $U'(R_P) > 0$  である。また、リターンが0の状態からスタートした場合、わずかでもリターンが増えれば大きく効用が高まる。しかし、すでに大きなリターンを得ている状態から始めた場合、追加のリターンを得たとしてもたいして効用は高まらないということ表すべく、 $U''(R_P) < 0$  ということを仮定する。

**注意 1.** こういった効用関数の性質については、リターンではなく、最終的な収益や、収益に基づく消費などによって効用は定まるという主張があったり、効用関数が連続であるとか、微分可能であるといった仮定についてもしばしば反論される場合がある。そのため、こういった効用関数に関する仮定はあくまで議論展開上の簡単のため置くものとする。

上述の仮定  $U' > 0, U'' < 0$  というをおいても効用関数にはいくつかの種類が考えられる。そこで、代表的なものを説明する。

例えば、

$$U(R_P) := 1 - e^{-\gamma R_P} \quad (1.3)$$

の形で表されるものを指数効用という。ここで、 $\gamma$  は  $\gamma > 0$  なる定数で、**絶対的リスク回避度** (ARA; absolute risk aversion) という。また、

$$U(R_P) := \frac{1}{1-\gamma} (1 + R_P)^{1-\gamma} - 1 \quad (1.4)$$

の形で表されるものをベキ効用といい、定数  $\gamma$  を**相対的リスク回避度** (RRA; relative risk aversion) という。

ベキ効用における  $\gamma$  が1に近づくと、効用関数は

$$U(R_P) = \log(1 + R_P) \quad (1.5)$$

という対数型の形となること知られている。

投資家によって、効用は高いほど良いが、ポートフォリオから得られるリターン  $R_P$  は確率変数であるため、最終的にどの程度の効用が得られるかは投資を終えるまではわからない。そこで、リターンから期待リターンを計算した時のように、効用から期待効用を計算する。期待効用は、

$$u(\pi_0, \pi) = \mathbf{E}[U(R_P)] \quad (1.6)$$

の形で与えられる。期待効用  $u$  の引数には、安全資産のポートフォリオウェイト  $\pi_0$  と、危険資産のポートフォリオウェイト  $\pi$  で与えた。期待効用は個々の事象に対するポートフォリオのリターン  $R_P$  に依存するが、この  $R_P$  は安全資産と危険資産のウェイトとリターンから計算されるため、違和感はないだろう。

この期待効用の最大化が投資家の目標となるため、

$$\begin{aligned} & \max_{\pi_0, \pi} u(\pi_0, \pi), \\ & \text{subject to } \pi_0 + \pi^\top \mathbf{1} = 1. \end{aligned} \quad (1.7) \quad \boxed{\text{eq:MeanUtilMax}}$$

が目標となる。

## 効率的フロンティアとの関係

簡単のため、<sup>eq:MeanUtilMax</sup>(1.7) 式の期待効用の近似<sup>4</sup>

$$\tilde{u}(\pi_0, \boldsymbol{\pi}) := \mu_P - \frac{\gamma}{2} \sigma_P^2 \quad (1.8) \quad \text{eq:HeikinApprox}$$

の最大化を行う。ここで、 $\gamma > 0$  は定数である。この  $\tilde{u}$  を平均分散型の期待効用関数という。

## 1.2.2 CAPM

危険資産  $S_1, S_2, \dots, S_M$  の期待リターンをそれぞれ  $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_M)^\top$  とする。安全資産の期待リターンを  $\mu_0$  とする。これらの資産からなるポートフォリオにより、代表的投資家の期待効用は以下のような平均分散型で与えられることを仮定する：

$$u_R(\pi_0, \boldsymbol{\pi}) = \mu_R - \frac{\gamma_R}{2} \sigma_R^2. \quad (1.9) \quad \text{eq:HeikinBunsan}$$

ここで、 $\gamma_R > 0$  は定数。このとき、均衡で危険資産の期待リターンは以下のように決まることが知られている

$$\boldsymbol{\mu} - \mu_0 \mathbf{1} = \beta(\mu_M - \mu_0). \quad (1.10) \quad \text{eq:CAPM1}$$

ここで、 $\beta$  はベータとよび、

$$\beta = (\sigma_{1,M}/\sigma_M^2, \sigma_{2,M}/\sigma_M^2, \dots, \sigma_{M,M}/\sigma_M^2)^\top$$

で定義する。 $\mu_M$  は市場ポートフォリオの期待リターン、 $\sigma_{i,M}$  は危険資産  $S_i$  と市場ポートフォリオの共分散である。この <sup>eq:CAPM1</sup>(1.10) 式は、平均分散型の期待効用 <sup>eq:HeikinBunsan</sup>(1.9) の最大化問題を解くことで得られる。

**問題 1.1.** 投資家の期待効用が

$$u_R(\pi_0, \boldsymbol{\pi}) = \mu_R - \frac{\gamma_R}{2} \sigma_R^2$$

で与えられると仮定する。 $\mu_R, \sigma_R$  はそれぞれ代表的投資家のポートフォリオの期待リターンとボラティリティで、 $\gamma_R > 0$  は定数である。

1.  $\mu_0$  を安全資産のリターン、 $\boldsymbol{\mu}$  を危険資産のリターンからなるベクトル、 $\sigma_{i,M}$  は危険資産  $S_i$  のリターン  $R_i$  と、市場ポートフォリオのリターン  $R_M$  との共分散とする。均衡で、次が成り立つことを示せ：

$$\begin{pmatrix} \sigma_{1,M} \\ \sigma_{2,M} \\ \vdots \\ \sigma_{M,M} \end{pmatrix} = \frac{\boldsymbol{\mu} - \mu_0 \mathbf{1}}{\gamma_R} \quad (1.11)$$

<sup>4</sup> $\tilde{u}$  が近似となっていることは問題1.2で確認する。この近似と用いるメリットは、この近似式は  $\mu_P, \sigma_P$  によって期待効用が買っていることで、効率的フロンティア上のどの点を選択することが期待効用最大化になるかがはっきりと分かることにある。

2.  $\mu_M$  を市場ポートフォリオの期待リターンとする。均衡において、次が成り立つことを示せ：

$$\boldsymbol{\mu} - \mu_0 \mathbf{1} = \beta(\mu_M - \mu_0).$$

**定理 1.2.** 安全資産と複数の危険資産からなるポートフォリオを考え、夫々の期待リターンを  $\mu_0, \boldsymbol{\mu}$ 、ポートフォリオウェイトを  $\pi_0, \boldsymbol{\pi}$  とする。ポートフォリオの期待リターンを  $\mu_P$ 、ボラティリティ  $\sigma_P$  に対して、以下の期待効用最大化問題を考える：

$$\begin{aligned} \max_{\pi_0, \boldsymbol{\pi}} \quad & \mu_P - \frac{\gamma}{2} \sigma_P^2, \\ \text{subject to} \quad & \pi_0 + \boldsymbol{\pi}^\top \mathbf{1} = 1. \end{aligned} \quad (1.12)$$

ただし、 $\gamma > 0$  は定数である。このとき、最適なポートフォリオウェイトは

$$\pi_0^* = 1 - \frac{1}{\gamma} \mathbf{1}^\top \Sigma^{-1} (\boldsymbol{\mu} - \mu_0 \mathbf{1}), \quad (1.13)$$

$$\boldsymbol{\pi}^* = \frac{1}{\gamma} \Sigma^{-1} (\boldsymbol{\mu} - \mu_0 \mathbf{1}). \quad (1.14)$$

で与えられる。このとき、期待リターンとボラティリティは夫々

$$\sigma_P^* = \frac{\kappa}{\gamma}, \quad (1.15)$$

$$\mu_P^* = \mu_0 + \frac{\kappa^2}{\gamma} \quad (1.16)$$

となる。ここで、 $\kappa$  は最適ポートフォリオウェイトによるシャープレシオ  $\kappa := \frac{\mu_P^* - \mu_0}{\sigma_P^*}$ 。また、これらを所与として得られる無差別曲線は平均分散フロンティアと接する。

*Proof.* まずは最適ポートフォリオウェイトから求めよう。ラグランジュ関数

$$L(\pi_0, \boldsymbol{\pi}, \lambda) = \mu_P - \frac{\gamma}{2} \sigma_P^2 + \lambda(1 - \pi_0 - \boldsymbol{\pi}^\top \mathbf{1}). \quad (1.17)$$

ここで、 $\mu_P = \pi_0 \mu_0 + \boldsymbol{\pi}^\top \boldsymbol{\mu}$ 、 $\sigma_P^2 = \boldsymbol{\pi}^\top \Sigma \boldsymbol{\pi}$  であることに注意すれば

$$\frac{\partial L}{\partial \pi_0} = \mu_0 \lambda = 0, \quad (1.18) \quad \text{eq:CAPMOpt\_L1}$$

$$\frac{\partial L}{\partial \pi_0} = \boldsymbol{\mu} - \gamma \Sigma \boldsymbol{\pi} - \lambda \mathbf{1} = 0, \quad (1.19) \quad \text{eq:CAPMOpt\_L2}$$

$$\frac{\partial L}{\partial \pi_0} = 1 - \pi_0 - \boldsymbol{\pi}^\top \mathbf{1} = 0 \quad (1.20) \quad \text{eq:CAPMOpt\_L3}$$

が得られる。<sup>eq:CAPMOpt\\_L1</sup> (1.18) 式より  $\lambda = \mu_0$  がわかるので、この結果を <sup>eq:CAPMOpt\\_L2</sup> (1.19) 式に代入し、 $\boldsymbol{\pi}^* = \frac{1}{\gamma} \Sigma^{-1} (\boldsymbol{\mu} - \mu_0 \mathbf{1})$  を得る。また、これをさらに <sup>eq:CAPMOpt\\_L3</sup> (1.20) 式に代入すれば  $\pi_0^* = 1 - \frac{1}{\gamma} \mathbf{1}^\top \Sigma^{-1} (\boldsymbol{\mu} - \mu_0 \mathbf{1})$  となる。

この結果をもとに期待リターン・ボラティリティを計算する。

$$\begin{aligned} \mu_P^* &= \pi_0^* \mu_0 + (\boldsymbol{\pi}^*)^\top \boldsymbol{\mu} \\ &= \mu_0 - \frac{1}{\gamma} \mu_0 \mathbf{1}^\top \Sigma^{-1} (\boldsymbol{\mu} - \mu_0 \mathbf{1}) + \frac{1}{\gamma} (\boldsymbol{\mu} - \mu_0 \mathbf{1})^\top \Sigma^{-1} \boldsymbol{\mu} \\ &= \mu_0 + \frac{1}{\gamma} (\boldsymbol{\mu} - \mu_0 \mathbf{1})^\top \Sigma^{-1} (\boldsymbol{\mu} - \mu_0 \mathbf{1}) \end{aligned} \quad (1.21)$$

ポートフォリオのボラティリティは危険資産のウェイトだけで定まり、

$$\begin{aligned} (\sigma_P^*)^2 &= (\boldsymbol{\pi}^*)^\top \Sigma \boldsymbol{\pi}^* \\ &= \frac{1}{\gamma^2} (\boldsymbol{\mu} - \mu_0 \mathbf{1})^\top \Sigma^{-1} (\boldsymbol{\mu} - \mu_0 \mathbf{1}). \end{aligned} \quad (1.22)$$

これらの結果をもとにシャープレシオ  $\kappa$  は

$$\kappa := \frac{\mu_P^* - \mu_0}{\sigma_P^*} = \sqrt{(\boldsymbol{\mu} - \mu_0 \mathbf{1})^\top \Sigma^{-1} (\boldsymbol{\mu} - \mu_0 \mathbf{1})}. \quad (1.23)$$

これを使えば、

$$\begin{aligned} \sigma_P^* &= \frac{\kappa}{\gamma}, \\ \mu_P^* &= \mu_0 + \frac{\kappa^2}{\gamma} \end{aligned}$$

となる。

最後に、最適ポートフォリオウェイトが所与として得られる無差別曲線は平均分散フロンティアと接することを確認しよう。最適ポートフォリオウェイトに  $y$  による期待リターンとボラティリティを  $u(\pi_0, \boldsymbol{\pi}) = \mu_P - \frac{\gamma}{2} \sigma_P^2$  に代入し、期待効用を計算すると、最適な効用水準は

$$\begin{aligned} u^* &= \mu_0 + \frac{\kappa^2}{\gamma} - \frac{\gamma}{2} \frac{\kappa^2}{\gamma^2} \\ &= \mu_0 + \frac{\kappa^2}{2\gamma} \end{aligned} \quad (1.24)$$

この効用水準を所与とする無差別曲線は  $\mu_P = \frac{\gamma}{2} \sigma_P^2 + u^*$  である。無差別曲線の傾きは

$$\frac{d\mu_P}{d\sigma_P} = \gamma \sigma_P$$

であるから、ボラティリティが  $\sigma_P^*$  のとき、 $\kappa$  となることから、無差別曲線の傾きは

$$\mu_P = \mu_0 + \kappa \sigma_P$$

の傾きと一致する。 □

この結果をみると、分散共分散行列  $\Sigma$  の構造や期待リターン  $\boldsymbol{\mu}$  の大きさにもよるが、直観的に言えば、 $\gamma$  が大きければ危険資産のウェイトは小さくなり、そうでない場合は大きくなること分かり、リスク回避的な投資家の行動と直観的に一致する。

### 1.2.3 市場の均衡と CAPM

### 1.2.4 確率的ディスカウントファクター (SDF)

## 演習

**問題 1.2** (Markowitz 2014). 効用関数が、指数型・べき型・対数型のとき、以下の近似がそれぞれ成り立つことを示したい：

(a) 指数型のとき： $u(\pi_0, \pi) \approx \gamma_e(\mu_P - \frac{1}{2}\gamma_e(\sigma_P^2 + \mu_P^2))$

(b) ベキ型のとき： $u(\pi_0, \pi) \approx \mu_P - \frac{1}{2}\gamma_P(\sigma_P^2 + \mu_P^2)$

(c) 対数型のとき： $u(\pi_0, \pi) \approx \mu_P - \frac{1}{2}P(\sigma_P^2 + \mu_P^2)$

次の(1)-(4)の手順に沿って、このことを示そう。

1. 2回連続微分可能な効用関数  $U(R_P)$  について、2次の Taylor 展開を 0 および  $\mu_P = \mathbf{E}[R_P]$  の周辺で施し、次の2式を得よ：

$$U(R_P) = U(0) + U'(R_P) + \frac{1}{2}U''(0)R_P^2 + \cdots, \quad (1.25) \quad \text{eq:ProbMarko1}$$

$$U(R_P) = U(R_P) + U'(R_P)(R_P - \mu_P) + \frac{1}{2}U''(0)(R_P - \mu_P)^2 + \cdots. \quad (1.26)$$

2. 0 周辺のテイラー展開 (I.25) 式について、期待値  $\mathbf{E}[U(R_P)]$  を計算せよ。
3. 期待効用  $u(\pi_0, \pi) = \mathbf{E}[U(R_P)]$  が一定となるようにポートフォリオを調整したとする。このとき、横軸を  $\sigma_P$ 、縦軸を  $\mu_P$  とした平面において、期待リターンを  $-U'(0)/U''(0)$  以下の範囲のみを考えるとする。このとき、 $U'(0) \geq 0, U''(0) < 0$  の仮定のもとで

$$u(\pi_0, \pi) = U'(0)\mu_P + \frac{1}{2}U''(0)(\sigma_P^2 + \mu_P^2) + C, \quad C: \text{定数}$$

が右上がりの曲線となることを確認せよ。

4. 効用関数が、指数型・ベキ型・対数型のとき、それぞれについて (a),(b),(c) が成り立つことを示せ。

### 1.3 ダウンサイドリスク

### 1.4 本章の演習問題を解くうえで用いる Excel 関数

関 数	意 味
AVERAGE( データの範囲 )	データの範囲に含まれるデータの平均値を返す。
STDEV.S( データの範囲 )	データの範囲に含まれるデータの標本標準偏差を返す。 ※母集団に対する標準偏差は STDEV.P で求めることができる。 STDEV とだけ入力してもエラーが出ないため、注意が必要。 分散を返す VARIANCE.S, VARIANCE.P 関数もある。
CORREL( データ A, データ B )	データ A と データ B の相関係数を返す。
SUMPRODUCT( データ A, データ B )	データ A と データ B の無い席を返す。
MMULT( 配列 A, 配列 B )	配列 A と 配列 B の行列積の結果を返す。
TRANSPOSE( 配列 )	配列を転置した配列を返す。 ※配列を 2 つ以上のセルに跨る配列を返す演算にあたっては、 Ctrl + Shift + Enter の押下が必要。



## 1.5 最小分散フロンティア、効率的フロンティア

**定義 1.1** (最小分散ポートフォリオ). ある機体収益率  $\mu_p^*$  に対して、 $(\mu_p, \sigma_p)$  の組み合わせの中で文章が最小となるもの（投資機会集合における境界上の点）を最小分散ポートフォリオと呼ぶ。

最小分散ポートフォリオは次の2次計画問題を解くことで決定することができる：

$$\begin{aligned} \min \text{Var}[R_p] & \left( \sum_i \sum_j w_i w_j \sigma_{ij} \right), \\ \text{subject to } & \sum_i w_i = 1, \sum_i w_i \mathbf{E}[R_i] = \mu_p^* \end{aligned} \quad (1.27)$$

**定義 1.2** (最小分散フロンティア). 様々な機体収益率に対応する最小分散ポートフォリオからなる集合を最小分散フロンティアと呼ぶ。

**定義 1.3** (帯域的最小分散ポートフォリオ). 最小分散境界の中で最も分散の小さいポートフォリオを、大域的最小分散ポートフォリオと呼ぶ。

**定義 1.4** (効率的フロンティア). 最小分散教会の中で、帯域的最小分散ポートフォリオよりも上側の部分を効率的フロンティアと呼ぶ。

## 1.6 演習問題

**問題 1.3.** 3つの資産  $S_1, S_2, S_3$  による投資を考えよう。各資産のリターンを  $R_1, R_2, R_3$ 、分散共分散行列  $\Sigma$  を

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 & \sigma_1\sigma_3 \\ \sigma_2\sigma_1 & \sigma_2^2 & \sigma_2\sigma_3 \\ \sigma_3\sigma_1 & \sigma_3\sigma_2 & \sigma_3^2 \end{pmatrix}$$

とする。初期時点  $t = 0$  で  $S_3$  を1単位購入し、その資金を  $S_1, S_2$  をロング・ショートすることによって調達する戦略、すなわち

$$w_1 S_1(0) + w_2 S_2(0) = S_3(0)$$

を満たすような戦略を取るとする。この戦略の時点  $t = 0$  における価値  $V(0)$  は

$$V(0) = -w_1 S_1(0) - w_2 S_2(0) + S_3(0)$$

と表される。時点  $T$  におけるポートフォリオの価値  $h$  以下の式で



## 第2章 債券投資理論の基礎

### 2.1 債券価値の評価

1900年代はじめには多くの債券が市場で発行され、取引されていた。当時は、長期の安定資金確保の必要性から、電力会社や鉄道会社が満期100年にもおよぶ永久債に近い債券を発行していた。こういったことも相まって、満期が非常に長い債券の価値を知る必要があった。Fischerは[Fischer:ToI]において、複利での最終利回りの計算式

$$P = \frac{C}{1+r} + \frac{C}{(1+r)^2} + \cdots + \frac{C}{(1+r)^{n-1}} + \frac{C+F}{(1+r)^n} \quad (2.1) \quad \text{eq:BondPrice}$$

を示した。ここで、 $P$ は債券の現在価値、 $C$ は債券のクーポン、 $r$ は複利ベースでの最終利回り、 $n$ は満期、 $F$ は債券の償還額である。(2.1)式は、後の投資理論の発展に大きな影響を与えたとされる。この式は、将来のキャッシュフロー（ここでは、クーポンと償還額）と現在価値、そして、複利ベースでの最終利回りの関係を示している。この考え方は、将来価値を金利で「割り引く」ことによって、現在価値が定まるということを示しており、「時間に価値はあるか？」という疑問に肯定的な答えを導いたといえるだろう。

#### 2.1.1 連続複利金利

現在価値と将来価値の関係から、時間に価値があることが示されたが、この価値を具体的な数値で表したものを「金利」として捉えることができる。しかし、単に金利が「年率 $X\%$ 」といっても、利息が年に1度支払われるのか、2度支払われるのか、はたまた毎月支払われる（年12回）のかによって結果は異なってくる。この違いは、利息が支払われるたびに、その利息が元本に組み込まれて新しい元本になっていく所謂「複利効果」によるものである。初期投資額100に対して、年率表示された金利が同じであっても、利息の支払頻度が高くなるにつれて受取総額の差が大きくなることは次の表より明らかだろう。

年5%での複利における受取額				
期間	1回/年	2回/年	12回/年	$\infty$ (連続複利)
1年	105.00	105.06	105.12	105.13
2年	110.25	110.38	110.49	110.52
3年	115.76	115.97	116.15	116.18
4年	121.55	121.84	122.09	122.14
⋮				
10年	162.89	163.86	164.70	164.87

$r$  を年率で表した金利、 $n$  を1年間に利息が支払われる回数として、 $R := r/m$  とおくと、1年後の受取総額は

$$\left(1 + \frac{r}{n}\right)^n = (1 + R)^{r/R} = \{(1 + R)^{1/R}\}^r$$

となる。ここで、 $n \rightarrow \infty$  とすると

$$\lim_{n \rightarrow \infty} \left(1 + \frac{r}{n}\right)^n = \lim_{R \rightarrow 0} \{(1 + R)^{1/R}\}^r = e^r$$

となる。この式から、1円を年率 $r\%$ で連続的に再投資することで1年後に $e^r$ 円得られるということがわかる。このように、利息の支払い回数を無限大としたときに得られるか理論上の仮想的な金利を連続複利金利といい、以後の議論で重要な役割を果たすことになる。

この逆に、1年後の1円の現在価値を連続複利で表すと $e^{-r}$ 円となる。したがって、 $t$ 年先のキャッシュフロー $CF_t$ の現在価値を連続複利で表すと、割引率が年率 $r\%$ の無リスク金利であるとすれば、

$$\text{現在価値} = CF_n \cdot e^{-rt} \quad (2.2)$$

となる。なお、現実の世界では、利息が連続的に支払われるわけではないため、上式は現実の近似表現ということになる。

債券の将来キャッシュフローには、大きさと支払時期が予め確定しているものと確定していないものがある。以下では、予め将来キャッシュフローが確定していて、デフォルト（債務不履行）リスクのない債券を対象として議論を薦める。

このような前提のもとで、将来キャッシュフローのパターンに応じて、債券を3つに大別することが出来る。[1] 将来キャッシュフロー（償還）が1回のみ発生する割引債、[2] 将来キャッシュフロー（クーポン）が複数回発生し、満期に最も大きなキャッシュフロー（償還とクーポン）が発生する利付債、[3] 毎回同じ大きさの将来キャッシュフローが複数回発生するタイプに分類することが出来る。

### [1] 割引債

割引債の理論価格は、発生時期（満期）を $T$ 、時点 $T$ に発生するキャッシュフローを $F$ とし、割引率 $r$ が1年複利であれば、

$$P = \frac{1}{(1 + r)^T} \quad (2.3)$$

となる。 $r$ が連続複利であれば

$$P = e^{-rT} F \quad (2.4)$$

となる。逆に、割引債の市場価格 $P_M$ が与えられていれば、1年複利の割引率や連続複利の割引率はそれぞれ、

$$r = \sqrt[T]{F/P_M} - 1, r = -\frac{1}{T} \ln \frac{P_M}{F}$$

となる。これらを用いれば、残存年数の異なる割引債の市場価格を用いて残存年数ごとの割引率を計算することが出来る。

## [2] 利付債

利付債の理論価格は、キャッシュフローの発生時期を  $t$ 、 $t$  期で支払われる利息を  $C_t$ 、満期を  $T$  として、満期で支払われる償還額を  $F$ 、1 年複利の割引率を  $r_t$  とすると

$$\text{利回り } P = \frac{C_1}{1+r_1} + \frac{C_2}{(1+r_2)^2} + \cdots + \frac{C_{T-1}}{(1+r_{T-1})^{T-1}} + \frac{C_T + F}{(1+r_T)^T} \quad (2.5)$$

となる。市場価格  $P_M$  が与えられ、割引率が  $t$  に依存しない一定値  $r$  であるという仮定のもとでは、

$$P_M = \frac{C_1}{1+r} + \frac{C_2}{(1+r)^2} + \cdots + \frac{C_{T-1}}{(1+r)^{T-1}} + \frac{C_T + F}{(1+r)^T}$$

となり、この関係式から割引率を求めることが出来る。この  $r$  は一般に**最終利回り**と呼ばれる。

ファイナンスの世界では、資産の価値は、それを保有することによって将来得られるキャッシュフローの現在価値を合計したものである。これらの典型的なキャッシュフローをもった債券も、理論価格算出の基本は、「将来キャッシュフローの現在価値の総和が適正価値である」という考え方で、元本やクーポンといったキャッシュフローを纏めて  $CF_t$  とおいて一般化すると、これらの債券の理論価格はいずれも

$$P = \sum_{t=1}^T \frac{CF_t}{(1+r_t)^t} \quad (2.6)$$

と表せる（1 年複利の場合）。

## 2.1.2 Spot rate, Discount factor

債券の現在価値算出の基本は、上述の記述をより一般に、 $i$  番目のキャッシュフローが時点  $t_i$  に発生するとき、

$$P = \sum_{i=1}^N \frac{CF_{t_i}}{(1+r_{t_i})^{t_i}} \quad (2.7)$$

を用いて計算できる。債券毎にキャッシュフローは異なるため、分子の  $CF_{t_i}$  は各債券に固有のものである。デフォルト（債務不履行）を考えていないため、分母の  $r_{t_i}$  は時刻  $t_i$  での無リスク金利を表している。 $CF_{t_i}$  とは異なり、 $r_{t_i}$  は対象としているすべての再建に共通する割引率である。割引債について考えれば、 $r_{t_i}$  は割引債の最終利回りとなる。

$r_{t_i}$  のことをスポットレート (spot rate) と呼ぶ。残存期間の異なる割引債が存在すれば、各残存期間に応じたスポットレートを計算することが出来る。すなわち、残存  $t_i$  年で満期に 100 円が支払われる割引債の現在価値を  $P(t_i)$  とし、スポットレートを  $s_{t_i}$  とおくと

$$P(t_i) = \frac{100}{(1+s_{t_i})^{t_i}} \quad (2.8)$$

となり、残存期間  $t_i$  の種々の割引債から各残存期間に応じたスポットレートを求めることが出来る。ここで、 $DF_{t_i} := \frac{1}{(1+s_{t_i})^{t_i}}$  をディスカウントファクターと定義すると、 $P(t_i) = 100 \times DF_{t_i}$  となる。スポットレートとディスカウントファクターは 1 対 1 の関係にあり、実質的に同じことを意味するものであることが分かる。

将来のある時点  $t$  で残存期間が  $t$  年である割引債を購入するとき、現時点で取引価格を約束する場合の満期までの保有期間利回りをインプライド・スポットレート (implied spot rate) と呼ぶ。このとき、将来のある時点を現時点と一致させる ( $t = 0$  とする) と、インプライド・スポットレートとスポットレートは等しくなる。

### 2.1.3 Forward rate, Spot rate

現時点を 0 とし、 $t$  年後からみた 1 年先の金利  $f(t, t+1)$  は  $t+1$  年と  $t$  年のスポットレートを

$$f(t, t+1) = \frac{(1+r(t+1))^{t+1}}{(1+r(t))^t} - 1 \quad (2.9)$$

となる。このように、将来のある時点からみた、さらに先の将来金利に関する現在の見通し金利をフォワードレート (forward rate) という。一般に、 $t$  年後からみた  $s$  年先の金利  $f(t, t+s)$  は

$$f(t, t+s) = \frac{(1+r(t+s))^{t+s}}{(1+r(t))^t} - 1 \quad (2.10) \quad \text{eq:SpotForward}$$

となる。この式は、 $(1+r(t+s))^{t+s} = (1+r(t))^t \cdot (1+f(t, t+s))^s$  という関係から導くことが出来る。

フォワードレートもスポットレート同様、横軸に期間、縦軸にフォワードレートをとると、フォワードレート曲線を描くことが出来る。両者の間には (2.10) 式のような関係が成り立っているため、フォワードレート曲線とスポットレート曲線の間にも一定の関係があり、一方の曲線が定まればもう一方も自動的に特定される。

**命題 2.1.** スポットレート曲線が右上がりであるとき、任意の  $s, t$  について  $r(t+s) < f(t, t+s)$  が成り立つ。また、右下がりであるときはその逆が成り立つ。

フォワードレート曲線を描くにあたっては、債券の市場価格データからスポットレート曲線を推計し、そこからフォワードレート曲線を描くことが多い。このとき、スポットレート曲線が滑らかで安定していないと、フォワードレート曲線が凹凸の激しい不自然な形状となってしまう。

各時点でのフォワードレートについて、将来を見通す期間を無限小とした場合のフォワードレートを瞬間フォワードレート (Instantaneous forward rate) という。これを  $F(t, T)$  とし、時点  $t$  での満期  $T$  の割引債価格を  $P(t, T)$  とすると

$$F(t, T) = -\frac{\partial}{\partial T} \ln P(t, T) \quad (2.11)$$

となる。ここで、 $F(t, T)$  を  $t$  から  $T$  まで積分することで

$$P(t, T) = \exp \left( - \int_t^T F(t, s) ds \right) \quad (2.12)$$

となり、割引債価格と瞬間フォワードレートを区間  $[t, T]$  で積分した値の関係がわかる。

### 2.1.4 裁定取引・一物一価の法則

2つの市場で同一視さんに異なる価格がついているとする。このとき、その資産のうち、安い価格のついている市場で購入し、高い価格のついている市場で売却することで、確実に利益を得ることができる。あるいは、金融市場では、ある資産の将来キャッシュフローについて、それと同じ将来キャッシュフローをもつ資産（これを複製資産という）を比較的簡単に複製することが出来る。このとき、原資産と複製資産の価格が異なっていると安い価格のついている資産を購入し、高い価格のついている資産を売れば確実に利益を得ることが出来る。このような例で示したように、投資資金が0（以下）であるのに、確実に利益を獲得できる取引を**裁定取引**という。

このような市場では、安い資産の価格は上昇し、高い資産の価格は下降していずれ両者の価格は等しくなる。同じ価値の資産の価格が一致することを**一物一価の法則**という。合理的な投資家を中心となる市場では一物一価の法則が成立すると考えられ、裁定取引の機会はすぐに消滅してしまう。このとき、**市場は無裁定**であるという。

## 2.2 利回り

債券投資の収益は利息収入部分、キャピタルゲイン（キャピタルロス）、利息の再投資収益という3つの要素によって決定される。これらの要素をどこまで厳密に考慮するかによって、債券の投資収益率の測定方法がいくつか存在する。以下では、これらの考え方を整理する。

### 2.2.1 単利：再投資収益を考慮しない収益率

#### 直利（直接利回り）

もっとも単純な収益率の尺度の1つ。支払われる利息と債券の購入価格との比をとったもの。3つの要素のうち利息収入の部分しか考慮されていない。

$$\text{直利（直接利回り）} = \frac{\text{利息}}{\text{債券購入価格}} \times 100(\%)$$

#### 最終利回り（日本式）

3つの要素のうち利息収入と、キャピタルゲイン（キャピタルロス）を考慮している。

$$\text{最終利回り} = \frac{\text{利息} + \{(\text{償還額} - \text{購入金額}) / \text{残存年数}\}}{\text{購入価格}} \times 100(\%)$$

#### 所有期間利回り

債券は必ずしも償還まで保有されず、満期を迎えるまでに売却してしまうこともある。このような場合に、所有期間に応じた収益率を測定する方法として所有期間利回りがある。考えとしては日本式の見最終利回りと同様である。

$$\text{所有期間利回り} = \frac{\text{利息} + \{(\text{売却額} - \text{購入金額}) / \text{残存年数}\}}{\text{購入価格}} \times 100(\%)$$

### 2.2.2 複利：再投資収益を考慮した収益率

複利の場合には、再投資の影響を考慮することから、単利の場合ほど簡単にその尺度を計算することが出来ない。こういった場合には、利回りを変数パラメータとして Solver や最適化ツールを用いて、市場価格と等しくなるように最適化を行うことで値を得る。

#### 最終利回り

期中に支払われる利子の再投資収益を考慮した収益率の尺度。最終利回りはすべての利息を償還時まで同一の金利で運用できるという仮定の下で計算しており、内部収益率 (IRR) などとも呼ばれる。

$$\text{市場価格} = \frac{1 \text{ 年目の利息}}{1 + \text{最終利回り}} + \frac{2 \text{ 年目の利息}}{(1 + \text{最終利回り})^2} + \cdots + \frac{\text{償還額}}{(1 + \text{最終利回り})^T}$$

#### 実効利回り

最終利回りにおける「再投資レートがすべて等しい」という前提は必ずしも現実的とはいえ兄。そこで、再投資レートを外部から与えて複利収益率を計算したものを実効利回りという。 $r_g$  を再投資レートとすると：

$$\text{市場価格} = \frac{1 \text{ 年目の利息} \times (1 + r_g)^{T-1}}{(1 + \text{実効利回り})^1} + \frac{1 \text{ 年目の利息} \times (1 + r_g)^{T-2}}{(1 + \text{実効利回り})^2} + \cdots + \frac{\text{償還額}}{(1 + \text{実効利回り})^T}$$

**注意 2.** 最終利回りは実務で広く利用されている指標の1つである。一般に金利は期間の長さに応じて変わるもの（後述する金利の期間構造）であるが、最終利回りの計算ではこれを単純化し、フラットな金利の期間構造を想定している。ことに留意しよう。定義より明らかに、最終利回りの計算においては、ある債券について発生時期が異なるすべての将来キャッシュフローを単一の割引率によって割引いて得られる現在価値の総和がその債券の価格に等しいと仮定している。したがって、ある債券の将来キャッシュフローと市場価格が与えられれば、その最終利回りは一意に定まる。これによって最終利回りと市場価格の間に1対1の対応関係が得られる訳だが、最終利回りは実際に投資家が受け取ることのできる収益を示しているわけではない。機中で受け取ったクーポンの再投資期間は相対的に短くなるため、その再投資レートは低下するのが一般的である。

債券の市場価格を通じてスポットレートと最終利回りの表現を繋げると

$$\begin{aligned} P &= \frac{CF_1}{1+r(1)} + \frac{CF_2}{(1+r(2))^2} + \cdots + \frac{CF_n}{(1+r(n))^n} \\ &= \frac{CF_1}{1+r} + \frac{CF_2}{(1+r)^2} + \cdots + \frac{CF_n}{(1+r)^n} \end{aligned} \quad (2.13)$$

となる（ここで、 $r$  を最終利回りとする）。この式を見てわかる通り、最終利回りは異なる期間のスポットレートを1つのに集約した値となっている。また、最終利回りは各々のスポットレートの大きさ・将来キャッシュフローの大きさによって決定される。したがって、満期が同じ債券であっても、キャッシュフローのパターンが異なれば、最終利回りは異なることになる。また、これらの値は、相対的に大きい将来キャッシュフローに対応するスポットレートに大きく影響を上げる。このため、最終利回りの大きさが債券の魅力度を判断するには注意が必要である。このように、将来のキャッシュフローの大きさが最終利回りに与える影響を**クーポン効果**という。



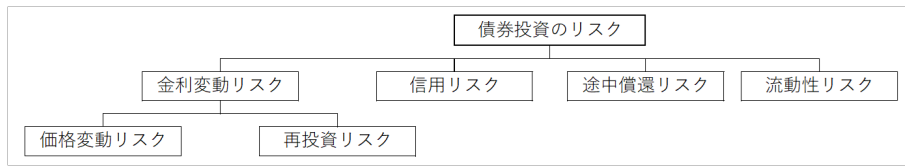


図 2.1: 債券投資のリスク

BondRisk

## 2.3 リスク評価

債券投資のリスクを考えるにあたって、一般には様々なリスクが考えられる。これらを1つの図に表すと以下ようになる。債券の理論価格は将来キャッシュフロー、割引率、投資期間の3つの要因によって決定される。割引率が変動すると、利息自体は変化しないものの、キャピタルゲイン（キャピタルロス）が発生するとともに、再投資収入が大きな影響を受けることになる。このため、金利変動リスクは債券投資に関する最大のリスク要因といえる。信用リスクは、債務不履行（デフォルト）リスクと呼ばれ、債券の発行体が利払いや償還を行えなくなるリスクである。途中償還リスクは、繰り上げ償還条項などがついている債券で、満期到来前に途中償還されてしまうリスクである。流動性リスクは、債券を計画通り売買しようとした際に、実行できないリスクである。

これらのリスクのうち、最大のリスク要因である金利変動リスクについて詳しく見ていこう。債券の理論価格をみてわかるように、債券価格に最も影響を与えるのはキャッシュフローの大きさと割引率（金利）である。一般に、キャッシュフローがあらかじめ決まっているため、理論上は割引率の変化、すなわち金利の変化が債券価格を変化させることにあつる。金利が上昇すれば債券価格は下落し、金利が低下すれば債券価格は上昇する。そこで、金利変動による債券価格の変化を表す指標として以下、デュレーションとコンベクシティを解説する。

### 2.3.1 デュレーション

金利変動に対する債券価格の変化を表す指標の1つとしてデュレーション (Duration) という考え方がある。これは、最終利回りの変化に対する債権価格の感応度によって定義される。

債券価格  $P$  は

$$P = \sum_{k=1}^n \frac{C_k}{(1+r)^k}$$

と表されるので、最終利回りの微小変化 ( $dr$ ) に対する債券価格  $P$  の変化率は

$$\begin{aligned} \frac{dP}{dr} &= -\frac{C_1}{(1+r)^2} - \frac{2C_2}{(1+r)^3} - \cdots - \frac{nC_n}{(1+r)^{n+1}} \\ &= -\frac{1}{1+r} \left( \frac{C_1}{1+r} + \frac{C_2}{(1+r)^2} + \cdots + \frac{nC_n}{(1+r)^n} \right) \end{aligned} \quad (2.14)$$

となる。両辺を  $P$  で和って、変形すれば

$$\frac{dP}{P} = -\frac{1}{1+r} \cdot \frac{1}{P} \left( \frac{C_1}{1+r} + \frac{C_2}{(1+r)^2} + \cdots + \frac{nC_n}{(1+r)^n} \right) dr$$

となる。ここで、 $dP = \Delta P$ ,  $dr = \Delta r$ , そして、

$$D := \frac{1}{P} \left( \frac{C_1}{1+r} + \frac{C_2}{(1+r)^2} + \cdots + \frac{nC_n}{(1+r)^n} \right)$$

とおくと,

$$\Delta P = -\frac{D}{1+r} \Delta r \cdot P$$

となる. このとき,  $D$  をデュレーションあるいは, マコーレー・デュレーション (Macaulay Duration) と呼ぶ. また,

$$D_{\text{mod}} := \frac{D}{1+r}$$

を修正デュレーション (modified duration) と呼ぶ.

デュレーションは, 「債券のキャッシュフローを受け取るまでの平均残存期間」と考えられる. このことを確認しよう. デュレーションの定義式を次のように分解する:

$$D = \underbrace{\frac{C_1}{P(1+r)}}_{=:w_1} \times 1 + \underbrace{\frac{C_2}{P(1+r)^2}}_{=:w_2} \times 2 + \cdots + \underbrace{\frac{C_n}{P(1+r)^n}}_{=:w_n} \times n.$$

すなわち, 各期のキャッシュフローの現在価値が債券価値全体に占める比率に年数をかけたものの総和をデュレーションと呼んでいることがわかる.  $w_1, w_2, \dots, w_n$  を合計すれば1となり,

$$D = w_1 \cdot 1 + w_2 \cdot 2 + \cdots + w_n \cdot n$$

となる. これが, デュレーションが債券のキャッシュフローを受け取るまでの平均残存期間と呼ばれる所以である.

**注意 3** (デュレーションの2つの解釈). デュレーションを経済的に意味づけしようとする2つの解釈が可能となる.

**金利感応度としてのデュレーション** 修正デュレーションは最終利回りが1単位変化したときの債券価格の経ぬお度合いを表したつ問であった. その値が大きければ金利感応度が大きいことになり, 最終利回りの変化に対してその債券の価格変動度合が大きいことを意味する.

**平均回収期間としてのデュレーション** マコーレー・デュレーションは, 各キャッシュフローの現在価値が債券の現在価値に占める比率に各々の残存年数をかけて合計をとったものである. つまり, キャッシュフローの大きさに残存年数を重みづけしてあることから, 加重平均残存年数を表しているとみることができる. デュレーションが大きければ, 加重平均残存年数が長いことになり, キャッシュフロー回収に時間がかかることを意味するため, 平均回収期間とも呼ばれる.

### 2.3.2 コンベクシティ

債券価格の金利感応度としてのデュレーションをみて分かる通り,  $\Delta r$  が十分に小さければ, デュレーションによる1次近似は価格変化のよい近似となる. 一方, 債券価格と最終利回りとの間には下に凸の関係が成り立っているため,  $\Delta r$  がある程度大きい, 或いは近似精度として高い精度を求める場合には, 1次だけではなく, 2次以上の影響も考慮する必要がある.

最終利回り  $r$  が  $\Delta r$  だけ変化した場合の債券価格を  $P(r + \Delta r)$  とすると, Taylor 展開により

$$P(r + \Delta r) \approx P(r) + P'(r)\Delta r + \frac{1}{2}P''(r)(\Delta r)^2 + \cdots + \frac{1}{n!}P^{(n)}(r)(\Delta r)^n$$

と表せる．デュレーションの議論から

$$\frac{\Delta P}{\Delta r} = P'(r) = -\frac{1}{1+r} \sum_{t=1}^n \frac{t \cdot C_t}{(1+r)^t}$$

なので,

$$\begin{aligned} P'(r)\Delta r &= -\frac{\Delta r}{1+r} \sum_{t=1}^n \frac{t \cdot C_t}{(1+r)^t} \\ &= -\frac{\Delta r}{1+r} D \cdot P \end{aligned}$$

となる．同様に

$$\frac{\Delta^2 P}{(\Delta r)^2} = P''(r) = \frac{1}{(1+r)^2} \sum_{t=1}^n \frac{t \cdot (t+1) \cdot C_t}{(1+r)^t}$$

とすれば,

$$P''(r)(\Delta r)^2 = \frac{(\Delta r)^2}{(1+r)^2} \sum_{t=1}^n \frac{t \cdot (t+1) \cdot C_t}{(1+r)^t}$$

となる．ここで,

$$C := \sum_{t=1}^n \frac{t \cdot (t+1) \cdot C_t}{P \cdot (1+r)^t}$$

を**コンベクシティ**と呼び, Taylor 展開の第 2 項までを用いると

$$P(r + \Delta r) \approx P(r) - \frac{D}{1+r} P(r) \Delta r + \frac{1}{2} \frac{C}{(1+r)^2} P(r) (\Delta r)^2$$

となる．修正デュレーション同様,  $C/(1+r)^2$  を**修正コンベクシティ**と呼び,  $C_{\text{mod}}$  と書く．これらを用いれば

$$P(r + \Delta r) \approx P(r) - D_{\text{mod}} P(r) \Delta r + \frac{1}{2} C_{\text{mod}} P(r) (\Delta r)^2$$

となる．

なお, 債券価格と最終利回りは下に凸であることから, 2 次の項は必ず正值となる．

**注意 4.** デュレーションやコンベクシティは, 債券あるいはポートフォリオの金利変動リスクの評価やほかの再建との金利感応度の比較をする上で有益な指標である．しかし, これらを利用する際には,

- 金利の期間構造がフラットである
- 金利の変化パラレルシフトである
- 金利の変化が微小である

といった前提をおいていることを忘れてはならない。

### 金利変動と再投資収入の変化

利息の再投資収入は再投資レートによって決まる。そのため、金利と再投資レートが連動している限り、金利変動は再投資収入に影響を与える。金利が上昇すれば再投資レートも上昇するので再投資収入は増加し、金利が下降すれば再投資レートも s 加工するので再投資収入は減少する。

これまでに見たように、金利変動によって債券価格の変化と再投資収入の変化という2つの変化が生まれるが、両者の効果は大きさに違いがあれども相殺する関係にあることがわかる。

#### 2.3.3 フィッシャー・ワイル・デュレーション

これまでのリスク指標は金利の基幹構造がフラットであることを仮定していたが、この仮定は現実から大きく乖離している。現実の金利の期間構造を考慮して、利付債の価格を算出すると、 $i$  期のスポットレートを  $r_i$  として

$$P_t(r) = \sum_{i=1}^t CF_i \exp(-r_i \cdot i) \quad (2.15)$$

と表せる。ここで、 $r_i$  が微小変化するならば、

$$\frac{\partial^n P_i}{\partial r_i^n} = (-i)^n CF_i \exp(-r_i \cdot i)$$

となるため、

$$\begin{aligned} \Delta P_t &= P_t(r + \Delta r) - P_t(r) \\ &= \sum_{i=1}^t \left\{ \frac{\partial P}{\partial r_i} \Delta r_i + \frac{1}{2} \frac{\partial^2 P}{\partial r_i^2} (\Delta r_i)^2 + \frac{1}{3!} \frac{\partial^3 P}{\partial r_i^3} (\Delta r_i)^3 + \dots \right\} \\ &\approx \sum_{i=1}^t \left\{ \frac{\partial P}{\partial r_i} \Delta r_i + \frac{1}{2} \frac{\partial^2 P}{\partial r_i^2} (\Delta r_i)^2 \right\} \end{aligned} \quad (2.16)$$

となる。すると、

$$\frac{\Delta P_t}{P} \approx \sum_{i=1}^t \frac{CF_i \exp(-r_i \cdot i)}{P} \left\{ -i \Delta r_i + \frac{1}{2} i^2 (\Delta r_i)^2 \right\} \quad (2.17)$$

となる。ここで、金利の期間構造が平行シフトしたとすると ( $\Delta r_1 = \Delta r_2 = \dots = \Delta r_t = \Delta r$  となるので)

$$\frac{\Delta P_t}{P} \approx -\hat{D}(r_i) \Delta r + \frac{1}{2} \hat{C}(r_i) (\Delta r)^2 \quad (2.18)$$

となる。ただし、

$$\begin{aligned} \hat{D}(r_i) &= \sum_{i=1}^t \frac{i CF_i \exp(-r_i \cdot i)}{P} \\ \hat{C}(r_i) &= \sum_{i=1}^t \frac{i^2 CF_i \exp(-r_i \cdot i)}{P} \end{aligned}$$

であり、この  $\hat{D}(r_i)$  をフィッシャー・ワイル・デュレーションと呼ぶ。

なお、最終利回りを連続複利で表した場合の修正デュレーションと修正コンベクシティは、 $r_1 = r_2 = \dots = r_t = r$  (一定) とした場合の  $\hat{D}(r_i), \hat{C}(r_i)$  に相当する。

### 2.3.4 実効デュレーション・キーレートデュレーション

これまでに説明してきた債券は、将来のキャッシュフローがあらかじめ確定していて、金利の変化は、全期間一定の幅で変化（パラレルシフト）するとして、議論を進めてきた。しかし現存する債券には、金利の変化に応じて将来のキャッシュフローが変化するようなオプション的性質を有するものも存在する。また、金利の変化は全期間一定ということは稀であり、ある特定期間の金利が上昇したり、下降したりする方が一般的である。

このようなオプション性を有する代表的な債券として、変動利付債やMBS（金利の水準によって期限前償還が大量発生する）がある。これらは、将来の金利の発生経路によりキャッシュフローが異なってくることになり、これまでに説明してきた修正デュレーションを使うには課題が残る。

**実効デュレーション**は、金利の変化などにより、将来のキャッシュフローが変化することを考慮した、より一般性をもったデュレーションであり、金利の変化に対する債券価格の変化の一般化された感応度と考えることが出来る。この考え方を使得、さらに複雑な債券の金利変化に対する債券価格の変化を捉えることができる（実際には、実効デュレーションの計算に、モンテカルロシミュレーションと呼ばれる手法によって行われることが多い）。

また、金利の変化は全期間一定の幅で変化するわけではなく、ある特定の期間のみ上昇または下降する場合が多い。**キーレートデュレーション**は、イールドカーブ上のある年限のスポットレートが微小変化し、他の年限のスポットレートは変化しないとした場合の、債券価格の変化の大きさを表している。この考え方を使得、複雑な金利期間構造の変化に対する債券価格への影響を考慮することが出来る。キーレートデュレーションは、金利の期間構造の一部の変化に対するデュレーションという意味で、**パーシャルデュレーション**と呼ばれることもある。

## 2.4 期間構造

### 2.5 短期金利モデルの概要

生命保険会社の保有する資産や負債は金利の影響を大きく受ける。そのため、詳細については別に章を立てて説明することとする。

債券価格は金利の期間構造に大きく左右される。したがって、適当な債券評価モデルを得るためには金利の期間構造をモデル化することが望ましい。金利の期間構造モデルには様々なものがあるが、短期金利をモデル化した例の概要を紹介する。

短期金利をモデル化し、そこにおなじみの「市場は無裁定である」という過程を加味すれば、リスク中立確率という確率の下で短期債投資を繰り返した時の資産価値の将来分布を知ることが出来る。このとき、将来の各事典について資産価値の期待値を計算すると、金利の期間構造を算出できる。

短期金利をモデル化する方法としては、短期金利そのものだけを取り扱う1ファクターモデルと、短期金利に加えて関連する属性などもあわせてモデル化するマルチファクターモデルが考えられる。

ファクターの数とは異なる切り口でも、短期金利モデルを大きく2つに分類することが出来、一つは均衡モデルと呼ばれるもので期間構造の時系列データよりファクターの変動過程を記述する考え方でVasicekモデルやCIRモデルと呼ばれるものが代表例となる。これらは金利の期間構造

にゆがみが存在しうることを前提とする長期の投資家に適したモデルである。一方、この考え方では現在の金利の期間構造を再現できるとは限らないため、金利商品に価格付けを行うマーケット・メーカーには都合が良くないモデルであるため、均衡モデルと対をなす無裁定モデルと呼ばれるものがある。これは、現在の金利の期間構造に完全に一致させるようにファクターの変動過程を記述する考え方である。つまり、現在の金利の期間構造（債券の市場価格）は正しい、という立場で作られており、マーケットメーカーが利用するモデルである。こちらの代表例は Ho-Lee モデルや Black-Derman-Toy モデルなどがある。

以下、1 ファクターの短期金利モデルについて概説する。

### 2.5.1 Merton モデル

Merton は

$$dr = \mu dt + \sigma dW(t) \quad (2.19)$$

というモデルを考案した。ここで、 $dr$  は短期金利  $r$  の微小変化、 $\mu$  は短期金利  $r$  のトレンド（ドリフト）項、 $\sigma$  は  $r$  のボラティリティ、 $dW(t)$  は Wiener 過程に従うランダム項である。このモデルは、短期金利の確率的な変動過程を一般化 Wiener 過程で表したものである。ただ、金利水準が大きくプラスまたはマイナスになってしまうことがあるため、実際の金利変動とは異なった動きとなってしまう。

### 2.5.2 Vasicek モデル

（マイナス金利が導入されるまでの間）金利の市場での動きを実際にみると、マイナスになることもなければ 10% を大きく超えるようなことも殆どなく、長期的には一定の水準を中心に上下ゆっくりと変動しているように見える。これは、**平均回帰**鳥屋れる現象の一例で、金利が高いときは景気が減速して金利低下につながり、逆に金利が低ければやがて景気が回復して金利が上昇するといった解釈がなされている。このような平均回帰の考え方をモデル化したものは Ornstein-Uhlenbeck 過程と呼ばれており

$$dX(t) = \alpha(X_0 - X(t))dt + \beta dW(t) \quad (2.20)$$

といった数式で表され、粘度のある溶液中の粒子のランダムな変動が粘性の抵抗で抑制されることを Ornstein と Uhlenbeck がモデル化したものである。このモデルを使って金利変動のメカニズムを記述したのが Vasicek で、短期金利の変動が平均回帰型の確率プロセスに従うと考え、

$$dr(t) = \kappa(\theta - r(t))dt + \sigma dW(t) \quad (2.21)$$

という式を得た。 $\theta$  は  $r$  が回帰する水準、 $\kappa(>0)$  はその回帰の強さを表している。回帰水準  $\theta$  に対して、金利が高いか低いかによって金利の動く方向が決定されるようになっていることがみてわかるだろう。

短期金利が Vasicek モデルに従う場合、時点  $t$ 、残存  $T$  の割引債（無裁定）価格は

$$P(t, T) = e^{A(T) - B(T)r(t)} \quad (2.22)$$

で与えられる。ここで、

$$A(T) := \frac{1}{\kappa^2} \left\{ (B(T) - T) \left( \kappa\theta - \frac{1}{2}\theta^2 \right) \right\} - \frac{1}{4\kappa} \sigma^2 B(T)^2,$$

$$B(T) := \frac{1}{\kappa} (1 - e^{-\kappa T}).$$

このとき、スポットレート  $r(t, T)$  が

$$r(t, T) = -\frac{1}{T} A(T) + \frac{1}{T} B(T) r(t) \quad (2.23)$$

となることは簡単な計算よりわかる。

### 2.5.3 CIR モデル

Vasicek モデルでは、金利が負値をとる可能性が残されている。これは、今となっては自然なことではあるものの、金利が負値をとることが考えられない時代において、この点を解決したいという動きがあった。Cox, Ingersoll, Ross が提案したモデルは、金利が負値をとる可能性を排除している。このモデルもまた、Vasicek モデル同様に平均回帰性をもっているが、変動の大きさが金利水準に合わせて上下する構造となっており、

$$dr(t) = \kappa(\theta - r(t))dt + \sigma\sqrt{r(t)}dW(t) \quad (2.24)$$

という式でモデル化した。ここで、

$$A(T) = \left( \frac{2re^{(\kappa+r)T/2}}{(\kappa+r)(e^{rT-1} + 2r)} \right)^{2\kappa\theta/\sigma^2},$$

$$B(T) = \frac{2(e^{rT-1})}{(\kappa+r)(e^{rT-1} + 2r)}$$

であり、割引債価格は

$$P(t, T) = e^{A(T) - B(T)r(t)} \quad (2.25)$$

で与えられる。

### 2.5.4 Ho-Lee モデル

Ho と Lee が提案したこの方法は、短期金利を直接モデル化する代わりに、無裁定条件のもとでの、割引債価格の確率的な挙動を分析するものである。

状態  $i$ 、時点  $t$  における残存期間  $T$  年の割引債価格を  $P(i, t, T)$  とする。このとき、投資期間を 1 念とする無リスクリターンは  $1/P(i, t, 1)$  である。したがって、リターンに不確実性のない世界を想定すれば、そこではすべての割引債から無リスクでリターンを得られるため、残存期間  $T+1$  年の割引債の時点  $t+1$  における将来価値を  $P(i, t, T+1)/P(i, t, 1)$  と表現できる。

次に、リターンに不確実性のある世界を創造する。ここで、時点  $t+1$  における不確実性を「状態  $i$  を維持」または「状態  $i+1$  への移行」のいずれかになるという形で表現することにする。また、状態  $i$  が維持される場合には金利が上昇し、一方で、状態  $i+1$  へ移行する場合には金利が

下落するものとする。このとき、時点  $t$  での残存期間が  $T + 1$  年であった割引債の時点  $t + 1$  での価格を以下の2式で表すこととする。

$$P(i + 1, t + 1, T) = \frac{P(i, t, T + 1)}{P(i, t, 1)} h^{\text{UP}}(T), \quad (2.26)$$

$$P(i, t + 1, T) = \frac{P(i, t, T + 1)}{P(i, t, 1)} h^{\text{DOWN}}(T). \quad (2.27)$$

ここで、 $h^{\text{UP}}(T), h^{\text{DOWN}}(T)$  は節導関数と呼ばれる関数で、それぞれ  $0 \leq h^{\text{DOWN}}(T) \leq 1 \leq h^{\text{UP}}(T)$  という関係が成り立っていないといけない<sup>1</sup>。

### 2.5.5 Hull-White モデル

### 2.5.6 Black-Derman-Toy モデル

## 2.6 マルチファクターモデル

## 2.7 期間構造 (revisited)

## 2.8 3-Factor Affine Model

## 2.9 債券ポートフォリオ運用

債券は、株式などのほかの運用資産と比べると

1. 価格変動リスクが小さい
2. 株式との相関が低い
3. 一般企業だけでなく、国や公的機関などの多様な発行体が存在するうえ、満期や優先劣後の違いがあることから銘柄の種類が豊富

といった特徴がある。これらの特徴から、債券はその多寡に関わらず、ポートフォリオに組み込まれることが多い。債券の具体的な運用例をみていく。

### 2.9.1 ラダー運用

債券の特徴の1つは、キャッシュフローの形状にある。満期償還があるため、購入後に満期償還まで保有すれば、あらかじめ決まった期間収益を確保することが出来る<sup>2</sup>。債券のこのような特徴から、債券運用では、当初購入した債券を満期まで保有し続けるという、バイ&ホールドという運用戦略が古くからおこなわれていた。その代表的な手法がラダー運用である。

ラダー運用とは、残存年数の異なる債券を、短期から長期にかけて均等に保有し、満期償還を迎える都度、その償還金を長期債へと再投資することを、機械的に繰り返す運用のことをいう。

<sup>1</sup>状態

<sup>2</sup>再投資をすれば、その分の収益のブレは存在するし、外貨建ての債券であれば円建てでの収益は為替によるブレも生じる



このラダー運用は換金性の高い債券を一定割合保有することで資金の流動性を確保しながら、利回りの高い長期債への投資を行うことで、全体の収益率を高めようとする運用である。

ラダー運用では、金利期間構造の予測や、それによる銘柄売買が生じないことから、運営管理が容易で、売買頻度も少なく取引コストを負担せずにするという特徴がある。また、ラダー運用は基本的にバイ&ホールドであることから、保有債券の時価ぶれが生じても会計上、損益を認識せずに済むケースが多い。

### 2.9.2 バーベル運用・ブレット運用

ラダー運用と比較して、債券運用におけるリスクを積極的にとることによって、期間収益を高める運用手法に、バーベル運用・ブレット運用がある。

今後の金利水準全体が高まる（下がる）と予測するなら、ポートフォリオのデュレーションを短く（長く）すれば良い。また、期間構造において相対的に割安だと判断する年限があれば、その残存期間の債券に比重を置いたポートフォリオを組めば良い。

同じくして、期間構造の形状について予測を持つ場合、デュレーションは同じだが、年限構成に違いのあるポートフォリオをいくつか考えることが出来る。大きく分ければ2つあり、1つは、ポートフォリオ全体のデュレーションに近い年限の銘柄だけで構成されるポートフォリオで、これを**ブレット運用**と呼ぶ。ブレット運用は、中期債に重きを置くことから、曲率成分が下に凸化する曲面や、傾斜成分が急になるような局面に強い。一方で、短期債と長期債の組み合わせによって構成されるポートフォリオは**バーベル運用**と呼ばれ、曲率成分が上に凸化する場合や傾斜成分が緩やかになる局面に強く、これら将来の期間構造の形状に対する投資家のビューが必要となる。

### 2.9.3 キャッシュフローマッチング運用

保険をはじめ、年金などの将来定期的な現金支払いを伴う負債について、その支払の時期と金額とが一致したキャッシュフローをもたらす債券ポートフォリオを構築する運用を**キャッシュフローマッチング運用**という。確定したキャッシュフローを提供する、という債券の特徴を活用した運用スタイルであるが、対象となるキャッシュフローと完全一致するポートフォリオを構築するのは至難であり、デュレーションを合わせつつ、なるべく誤差を減らしてキャッシュフローを合わせるように構築する場合が殆どである。



## 第3章 株式運用の基礎

3.1 本質的価値に基づく株式価値評価

3.2 価格変動の要因

3.3 ファクターモデル

3.4 3-Factor モデル



## 第4章 デリバティブ

ここでは、金融派生商品（デリバティブ）を取り扱う。この章における目的はオプションについて基本的な事項を理解することと、VBA(Visual Basic for Application) を用いてユーザー関数を作る手順を理解することである。

### 4.1 先物

「将来のある時点に、一定の条件で受け渡しを約束する取引」を**先物取引**といい、受け渡しを約束した期日以前であれば反対売買を行うことで、値動きによって生じる差額だけの受け渡しにより取引を決済（差金決済）することができる。先物取引の例としては、国債先物や日経平均先物、金利先物などがある。

これらは取引所取引が中心であり、取引対象や方法は標準化されている。先物市場の参加者は、売買にあたって少額の証拠金を差し入れる必要がある。取引が開始されると、毎営業日の売買終了後の評価損益を反映し、証拠金は調整されることになる（これを**値洗い**という）。

#### 4.1.1 先物理論価格

時点  $t$  で無リスク金利  $r$  で資金を借入、現物債券を 1 単位購入し、その価格を  $B_t$ 、クーポンを  $C$  とする。同時に時点  $T$  で決済を行う先物を価格  $F_t$  で 1 単位売却することを考える。

この場合の収益は

$$\{F_t + C(T - t)\} - [B_t + B_t\{r(T - t)\}] = (F_t - B_t) + (C - B_t \cdot r)(T - t)$$

となる。キャッシュを仮入れてきて、現物を買って保有（キャリー）していることから、この運用手法はキャッシュ&キャリーと呼ばれる。この取引でプラスの収益が得られるならば、効率的な市場であれば、次々と市場参加者が増え、先物価格  $F_t$  は低下あるいは現物価格  $B_t$  は上昇し、得られる収益は次第に少なくなっていく。逆に収益がマイナスとなる場合には、逆の取引、つまり現物を空売りして先物を買う取引で収益を狙う市場参加者が増えて、先物価格は上昇（現物価格は下落）する。このように、効率的な市場では、無リスクで収益を得ようとすれば必ず収益機会がなくなるまで市場参入が相次ぐ所謂「裁定」が働くため、最終的には無リスクの取引による収益機会が 0 となるはずである。このことから、理論先物価格は次のように求められる：

$$F_t = B_t \left\{ 1 + \left( r - \frac{C}{B_t} \right) (T - t) \right\}. \quad (4.1) \quad \text{eq:Future1}$$

とくに、

$$B_t \left( r - \frac{C}{B_t} \right) (T - t)$$

はキャッシュ&キャリー取引に伴う損益を表しており、**キャリー損益**と呼ばれる。

**例 4.1.** クーポン 3% の債券価格が 100 円で、無リスク金利が 1% のとき、3 か月先の先物価格は

$$100(1 + 2\% - 3\%) \times \frac{3}{12} = 99.50 \text{ (円)}.$$

**問題 4.1.** 無配当の株が 5000 円で、無リスク金利 2% のとき、その株式を対象とする期間 6 か月の先物価格はいくらか。

ファイナンスの世界では、便宜上無リスク金利を連続複利で表すことが多い。期中収入の現在価値を  $I$ 、連続複利ベースでの無リスク利子率を  $r$  とすると (4.1) 式は <sup>eq:Future1</sup>

$$F_t = (B_t - I)e^{r(T-t)} \quad (4.2)$$

となる。配当がある場合にも、その配当が年率  $q$  で連続的に支払われるとして

$$F_t = (B_t - I)e^{(r-q)(T-t)} \quad (4.3)$$

とすればよい。

**例 4.2.** ある株が 5000 円で、無リスク利子率 2% (連続・年率) で、配当利回りが 3% (連続・年率) であるとき、その株を対象とする 6 か月先物の価格は

$$5000 * e^{(2\%-3\%)\cdot\frac{6}{12}} = 4975.062$$

となる。

#### 4.1.2 先物を用いたヘッジ

先物の利用方法として、現在保有している資産の価格変動リスクを回避する、「売りヘッジ」といった機能がある。先物で現物の反対のポジションをとることによって、現物市場の損失を先物市場での利益で相殺することによって保有資産の価値を保全することができる。

**例 4.3.** 長期国債のヘッジを考える。単価 97 円の長期国債を額面で 100 億円保有している投資家が、ヘッジのために債券先物を単価 108 円で 100 億円分売却した。その後、長期国債の単価が 94 円に、債券先物価格が 105 円にそれぞれ値下がりしたとき、投資家にはどのような損失が発生しているかを考えて見よう。

長期国債の資産価値は  $100 \text{ 億円} \times 97/100 = 97 \text{ 億円}$  なので、長期国債の値下がり、 $100 \text{ 億円} \times (94 - 97)/100 = -3 \text{ 億円}$  の損失。一方、債券先物では、 $-100 \text{ 億円} \times (105 - 108)/100 = 3 \text{ 億円}$  の利益が出てくる。そのため、投資家はこのような戦略をとることで、もともとの資産価値が完全に保全されている。

## 4.2 オプション

### 4.2.1 オプションとは

**定義 4.1** (オプション). オプションとは、ある特定の日 (または期間内) に、あらかじめ決められた価格で、ある特定の資産を買うあるいは売る権利を指す。

あらかじめ決められた価格のことを**権利行使価格**といい、特定の資産のことを**原資産**と呼ぶ。買う権利のことを**コールオプション**といい、売る権利のことを**プットオプション**という。

**定義 4.2** (ヨーロピアン・オプション). ある特定の日 (**満期日**) にしか権利を行使できないタイプのオプションを**ヨーロピアン・オプション**という。

**定義 4.3** (アメリカン・オプション). ある特定の期間内にいつでも権利を行使できるタイプのオプションを**アメリカン・オプション**という。

オプションのペイオフを考えて見よう。

**例 4.4.** A 社の株を原資産とするヨーロピアン・コールオプションを持っている場合を考える。権利行使価格が 100 円のオプションについて、満期時点で原資産価格が 100 円以上であるとき、オプションを行使し、100 円で株式を購入する。その後、原資産価格をその時の値段で売却すればその差額分の利益を得ることが出来る。逆に、原資産価格が 100 円未満であるときには、権利を行使しない (つまり、株式を 100 円で購入しない) ことにすれば、売買が行われず損益は発生しない。

一方、これがヨーロピアン・プットオプションである場合には、満期時点で権利行使価格 > 原資産価格なら原資産を購入し、オプションを行使することで 100 円で株式を売却することで差額を利益として得る。逆に、権利行使価格 < 原資産価格であるときには、権利を行使しないことにすれば、売買が行われず損益は発生しない。

そのため、オプションのペイオフは、コールオプションの場合は  $(S_T - K)^+$ 、プットオプションの場合は  $(K - S_T)^+$  で表される<sup>1</sup>。

オプションのペイオフだけをみれば、確実に得をするように見えるが、この権利を得るためには対価が必要となる。オプションを買った投資家は、契約時に**プレミアム** (オプション料) を支払う代わりに満期時に権利行使をする権利を得るのである。プレミアムをどのように計算するかについては、後程扱う。

**問題 4.2.** A 社の株式を原資産とする行使価格 2,000 円のプットオプションをプレミアム 300 円で 1 枚購入したとする。このとき、以下の夫々について答えよ。

1. 満期日の株価が 2,500 円の場合の損益はどのようなになるか。
2. 満期日の株価が 1,500 円の場合の損益はどのようなになるか。

**問題 4.3.** A 社の株式を原資産とする行使価格 2,000 円のコールオプションをプレミアム 300 円で 1 枚売却したとする。このとき、以下の夫々について答えよ。

1. 満期日の株価が 2,500 円の場合の損益はどのようなになるか。
2. 満期日の株価が 1,500 円の場合の損益はどのようなになるか。

<sup>1</sup> $(\bullet)^+$  は  $\max\{\bullet, 0\}$  の意味で、プラス部分を表す。

### 4.2.2 オプション価格の理論

コールオプションは、原資産価格が行使価格を上回った時に価値を持ち、プットオプションは原資産価格が行使価格を下回った時に価値を持つことは先述の通りであり、コールオプションの価値  $C$ 、プットオプションの価値  $P$  はそれぞれ、満期時点 ( $T$ ) における原資産価格  $S_T$  と権利行使価格  $K$  によって

$$C = (S_T - K)^+, P = (K - S_T)^+$$

となる。それでは、満期日以前のオプションの価値はどうだろうか？それをみていくのがここでの目的である。

オプションは満期日到来以前から、何らかの価値をもつものの、その価値は満期日と同じとは限らない。オプションの価値は、満期日の損益にあたる**本源的価値** (Intrinsic Value) と、原資産価格の将来変動により利益を得られる可能性を表す時間的価値 (Time Value) の2つに分解される。

本源的価値の有無によって、オプションの状態は In the Money (ITM), At the Money (ATM), Out of the Money (OTM) の3つの状態がある：

コール	プット	オプションの状態
$S_T > K$	$S_T < K$	In the Money (ITM)
$S_T = K$	$S_T = K$	At the Money (ATM)
$S_T < K$	$S_T > K$	Out of the Money (OTM)

オプションが ATM や OTM の状態にある場合であっても、通常何らかの価値を有している。これは、原資産の価格が将来変動することにより ITM になる可能性があることを評価したものが時間的価値である。

原資産の動きが激しいほど、現時点で OTM にあるオプションが ITM となる可能性は大きくなるため、そのオプションの時間的価値も大きくなる。原資産の価格変動を統計的に測定した指標であるボラティリティにオプションの時間的価値は関係する。

満期日が近づくにつれて現時点で OTM にあるオプションが ITM になる可能性が減少するため、オプションの時間的価値も次第に小さくなる。これをタイム・ディケイ (Time Decay) という。

このように、時間的価値は原資産がどのような値動きをするかによって大きく違ってくるため、オプションの価値を推計するのは容易ではない。ここからは、オプションの価値を求めるアイデアの1つとして2項モデルを紹介する。

## 2項モデル

現在の原資産価格を  $S$ 、現時点から満期日にかけて値上がりした場合には  $Su$ 、値下がりした場合には  $Sd$  になるものとする ( $u$  は  $1+$  上昇率、 $d$  は  $1+$  下落率)。

1 期間モデルに対応するコールオプションを例に考える。行使価格  $K$  ( $Sd < K < Su$ ) について、満期日におけるコールオプションの価値は、原資産価格が上昇した時  $Su - K$ 、原資産価格が下落した場合  $0$  となる。

続いて、原資産とオプションを組み合わせる無リスクポートフォリオを作成する。仮に現時点で原資産を  $\Delta$  単位、コールオプションを  $-1$  単位もっていたとする。このポートフォリオの満期日における価値は、原資産価格が上昇した時  $\Delta \cdot Su - c_u$ 、原資産価格が下落した場合  $\Delta \cdot Sd - c_d$



になる。このポートフォリオが無リスクであるためには、原資産が上がろうと下がろうと満期時の価値が同じでなければならないから

$$\Delta \cdot Su - c_u = \Delta Sd - c_d$$

とならなければならない、これを整理すると

$$\Delta = \frac{c_u - c_d}{S(u - d)}. \quad (4.4) \quad \text{eq:2model\_1}$$

また、無リスクポートフォリオ  $(\Delta S - c)$  は満期日までに安全資産と同等の収益を生むはずである<sup>2</sup>から、 $R := 1 + r$  ( $r$  は無リスク金利) として、

$$\Delta \cdot S_u - c_u = (\Delta S - c)R. \quad (4.5) \quad \text{eq:2model\_1}$$

が成り立つ。<sup>eq:2model\\_1</sup> <sup>eq:2model\\_2</sup> (4.5) 式を (4.4) 式に代入し、 $c$  について解くとオプションの現時点での価値が求められる：

$$c = \frac{1}{R} \left( \frac{r - d}{u - d} c_u + \frac{u - r}{u - d} c_d \right). \quad (4.6)$$

上述の2項モデルは現時点と満期の2時点間（1期間）で考えたが、この1期間2項モデルの考え方を繰り返し適用することによって、任意のオプションの価値を算出することが出来る。

2項モデルは閉じた方程式ではなく、オプション価格を決定するためのアルゴリズムを示したものであるが、応用範囲が広い。途中で権利行使の可能性があるアメリカンタイプのオプションを解くのに用いることも出来る。

## プット・コールパリティ

行使価格と満期が同一なヨーロピアン・コールオプション、ヨーロピアン・プットオプションの価格の間には一定の関係があり、一方が決まれば他方決まるという仕組みになっている。この関係を**プット・コールパリティ**という。

原資産とそのコールオプション、プットオプションおよびキャッシュからなる無リスクポートフォリオを作成し、プット・コールパリティを導く。

満期までの期間  $T$  かつ行使価格  $K$  のコールオプションを1単位売却し、同じくプットオプションを1単位購入する。原資産を1単位購入し、キャッシュを  $K \cdot e^{-rT}$  借入れるポートフォリオを構築する ( $r$  は連続複利ベースの無リスク金利)。このポートフォリオは満期時における原資産価格がどうなろうともその価値は0となる。

現在のポジション	$S_T \geq K$ のとき	$S_T < K$ のとき
$-c$	$K - S_T$	0
$p$	0	$K - S_T$
$S_0$	$S_T$	$S_T$
$K \cdot e^{-rT}$	$-K$	$-K$
合計	0	0

<sup>2</sup> そうでない場合、無リスク金利で借入を行い、無リスクポートフォリオを保有するか、空売りして無リスク金利で運用することによって、無リスクで収益が得られる。

満期時点のポートフォリオの価値が0であるならば、現在のポートフォリオの価値も0でなければならないため

$$p + S_0 - K \cdot e^{-rT} - c = 0$$

となる。これを整理すれば

$$p = c - S + K \cdot e^{-rT}. \quad (4.7) \quad \text{eq:PutCallParity}$$

**注意 5.** このプット・コールパリティが成り立つのは (1) ヨーロピアンタイプオプションであること、(2) 配当・クーポンがないこと、この2条件をともに満たすオプションに限られる。

## Black-Scholes モデル

2項モデルの基本的を示したが、オプション評価には **Black-Scholes model** (BS モデルとも表記される) が用いられることも多い。このモデルは、オプション評価モデルとしてポピュラーなものであるが、もともとは配当のない、ヨーロピアンタイプの、株式オプションの評価モデルとして作成されたものである。そのため、それ以外のオプション評価に対してはモデルを修正する必要がある。このモデルによる評価式は一見複雑であるが、導出の基本は2項モデルとほぼ同様で、無リスクポートフォリオと裁定の考え方がベースとなる。

ヨーロピアン・コールオプション、ヨーロピアン・プットオプションの価格  $C(S, K, \tau, r, \sigma)$ ,  $P(S, K, \tau, r, \sigma)$  は以下で与えられる：

$$C = C(S, K, \tau, r, \sigma) = S \cdot \Phi(d_1) - K \cdot e^{-r\tau} \cdot \Phi(d_2), \quad (4.8) \quad \text{eq:BS-Call}$$

$$P = P(S, K, \tau, r, \sigma) = K \cdot e^{-r\tau} \cdot \Phi(-d_2) - S \cdot \Phi(-d_1). \quad (4.9) \quad \text{eq:BS-Put}$$

ここで、 $S$  は原資産価格、 $K$  は権利行使価格、 $r$  は（連続複利ベースの）無リスク金利、 $\sigma$  は原資産のボラティリティ、 $\tau$  は計算時点から満期までの期間で、 $d_1$  および  $d_2$  は

$$d_1 := \frac{\ln(S/K) + (r + \sigma^2/2)\tau}{\sigma\sqrt{\tau}}, \quad (4.10)$$

$$d_2 := \frac{\ln(S/K) - (r + \sigma^2/2)\tau}{\sigma\sqrt{\tau}} = d_1 - \sigma\sqrt{\tau}. \quad (4.11)$$

さて、ここで一つ問題となるのが、ボラティリティである。

Black-Scholes モデルを利用する場合、最も入力に気を遣うのが原資産のボラティリティである。ボラティリティの推計方法には何種類もあり、その一つは、ポートフォリオのリスクとして使ったように、過去の一定期間の対象原資産の価格変化率の標準偏差を用いた**ヒストリカル・ボラティリティ**である。

## ヒストリカル・ボラティリティ

$t$  日の原資産価格を  $S_t$ 、日次収益率を  $R_t := \ln(S_t/S_{t-1})$  と表すと標本機関を過去  $n$  日間とするヒストリカル・ボラティリティ (HV) は、以下のように定義づけされる：

$$HV = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (R_i - \bar{R})^2}, \quad \bar{R} := \frac{1}{n} \sum_{i=1}^n R_i \quad (4.12)$$

ヒストリカル・ボラティリティは年率で表示されることが一般的である。日次ベースのボラティリティや月次ベースのボラティリティを年率換算するには、営業日数・月数が用いられる：

$$HV(\text{年率}) = HV(\text{日次ベース}) \times \sqrt{250},$$

$$HV(\text{年率}) = HV(\text{週次ベース}) \times \sqrt{52},$$

$$HV(\text{年率}) = HV(\text{月次ベース}) \times \sqrt{12}.$$

**注意 6.** オプションの評価に際して必要とされるボラティリティはオプションの満期日までの将来のボラティリティであることに注意しておく。過去のボラティリティであるヒストリカルボラティリティを、オプション評価に用いて良いのか？という問題はあるが、計算が容易なこともあり将来のボラティリティの推計値として用いられている。

**注意 7.** ヒストリカル・ボラティリティの計算は簡単であるが、標本期間によってその値が異なる。標本機関として、どの期間がふさわしいのか？という問いに対して決定的な解があるわけでもない。将来のボラティリティの推計値としてふさわしいように標本期間を工夫するしかない。例えば、長期のデータを使えば統計的にあり得る正確性という観点では説明がつくのかかもしれないが、直近の値動きをうまく反映できなくなってしまう。この課題を解決するために  $(R_t - \bar{R})^2$  を単純平均するのではなく、加重平均を取ることにして、直近ほどウェイトを高めるなどの方法をとることがある。

### インプライド・ボラティリティ

一方で、Black-Scholes モデルから逆算されるボラティリティーを**インプライド・ボラティリティ**と呼ぶ。インプライド・ボラティリティ(*IV*)は、Black-Scholes モデルを用いて、Newton-Raphson 法などの手法を用いて収束計算によって算出されたボラティリティである。*IV* は現時点でのオプション市場参加者が予想している将来に向けてのボラティリティを表したもので、オプション評価に用いるボラティリティとしては *HV* に比べると適しているという見方ができる。一方で、債券先物オプションのような上場オプション市場や指標銘柄のオプション市場以外では、オプション価格を観察することが難しく、*IV* を常に計算できるわけではないことに注意が必要である。

**注意 8.** Black-Scholes モデルでは、同一限月のオプションならば、ボラティリティも一定であると考えている。一方で、実際のオプション市場では、行使価格 (*K*) が異なれば *IV* も異なり、ATM よりも OTM の *IV* が高くなる**ボラティリティ・スマイル**と呼ばれる現象がよく観察される。これは、オプション市場の需給を反映した結果で、市場があれど需給関係<sup>3</sup>が不安定になっている場合は必ずスマイル状にはならないものの、権利行使価格ごとの割高・割安などを判断する材料にはなり得る。

#### 4.2.3 Black-Scholes モデルの限界

Black-Scholes モデルには色々な仮定が設けられているが、モデルの限界を考えるにあたっては次の2つの仮定は重要になってくる。

<sup>3</sup>例えば、コールオプションの売り手は、タイムディケイの効果が大きい ATM オプションに集中し、コールオプションの買い手は、レバレッジの大きい OTM に集中することが多いため、OTM の *IV* が割高になるといったことが考えられる。

1. 微小な時間における原資産の収益率が正規分布に従い、将来の任意時点の原資産価格が対数正規分布に従う：

$$\frac{dS_t}{S_t} = \mu \cdot dt + \sigma \cdot dW_t.$$

2. ボラティリティ( $\sigma$ )や無リスク金利( $r$ )が一定である。

このような仮定は、原資産が株式や通貨の場合はそこまで不自然ではない<sup>4</sup>が、債券を原資産とする場合には様々な点で不都合が生じる。

#### 4.2.4 オプションを用いたヘッジとスペキュレーション

オプションの活用手段として、ヘッジとスペキュレーションの例を挙げておく。

**例 4.5** (オプションを用いたヘッジ). とある企業の株式を保有している時、プットオプションを購入することによって、一定以上の損失を抑えることが出来る。

**例 4.6** (オプションを用いたスペキュレーション). 将来原資産の値上がりを予想するときはコール・オプションを、値下がり予想する場合はプット・オプションを購入すればスペキュレーションが出来る。また、権利行使価格の異なる複数のオプションを用いることで、ボラティリティの変動にも賭けることが出来、現物や先物にはないオプションの特徴である。例えば、原資産と満期日が同じで、権利行使価格が異なるコールとプットを同数買う（または売る）戦略を**ストラングル**という。また、原資産と満期日と行使価格が同じコールとプットを同数買う（または売る）戦略を**ストラドル**という。

#### 4.2.5 オプションのリスク

Black-Scholes モデルによるオプションの価格は、原資産価格  $S$  とそのボラティリティ  $\sigma$ 、残存期間  $\tau$ 、無リスク金利  $r$  の変化など様々な要因で変動する。そこで、どの要因でオプション価値がどれだけ変動するかを捉えるため、**グリークス (Greeks)** と呼ばれる各種リスク指標を概説する。

##### デルタ ( $\Delta$ )

原資産価格の変化に対するそのオプションのプレミアムの感応度（変化率）を**デルタ**と呼ぶ。例えば、あるオプションのデルタ ( $\Delta$ ) が 0.5 というのは、原資産価格が微小に変化したとき、オプションのプレミアムはその原資産価格の変化の 70% 変動することを意味している。

配当のないヨーロピアン・オプションのデルタは、<sup>eq:BS-Call</sup>(4.8) 式・<sup>eq:BS-Put</sup>(4.9) 式をそれぞれ  $S$  で偏微分すればよく、

$$\Delta_{\text{Call}} = \Phi(d_1), \quad (4.13) \quad \text{eq:BS-Call-Delta}$$

$$\Delta_{\text{Put}} = \Phi(d_1) - 1 \quad (4.14) \quad \text{eq:BS-Put-Delta}$$

となる。

<sup>4</sup>実際には、市場で観測されるジャンプなども踏まえると、このモデルの仮定はやや厳しいのであるが。

ガンマ ( $\Gamma$ )

原資産価格の変化に対するそのオプションのデルタの感応度（変化率）を**ガンマ**と呼ぶ。

配当のないヨーロピアン・オプションのガンマは、<sup>eq:BS-Call-Delta</sup>(4.13)式・<sup>eq:BS-Put-Delta</sup>(4.14)式をそれぞれ  $S$  で偏微分すればよく、

$$\Gamma = \Gamma_{\text{Call}} = \Gamma_{\text{Put}} = \frac{\varphi(d_1)}{S\sigma\sqrt{\tau}} \quad (4.15) \quad \text{eq:BS-Gamma}$$

となる。ここで、 $\varphi$  は標準正規分布の確率密度関数である。

セータ ( $\Theta$ )

残存の変化に対するそのオプションのプレミアムの感応度（変化率）を**セータ**と呼ぶ。プレミアムのタイム・ディケイとも呼ばれている。

配当のないヨーロピアン・オプションのガンマは、<sup>eq:BS-Call</sup>(4.8)式・<sup>eq:BS-Put</sup>(4.9)式をそれぞれ  $\tau$  で偏微分すればよく、

$$\Theta_{\text{Call}} = -\frac{S\varphi(d_1)\sigma}{2\sqrt{\tau}} - rKe^{-r\tau}\Phi(d_2), \quad (4.16) \quad \text{eq:BS-Call-Theta}$$

$$\Theta_{\text{Put}} = -\frac{S\varphi(d_1)\sigma}{2\sqrt{\tau}} + rKe^{-r\tau}\Phi(-d_2) \quad (4.17) \quad \text{eq:BS-Put-Theta}$$

となる。

ロー ( $\rho$ )

無リスク金利の変化に対するそのオプションのプレミアムの感応度（変化率）を**ロー**と呼ぶ。

配当のないヨーロピアン・オプションのローは、<sup>eq:BS-Call</sup>(4.8)式・<sup>eq:BS-Put</sup>(4.9)式をそれぞれ  $r$  で偏微分すればよく、

$$\rho_{\text{Call}} = K\tau e^{-r\tau}\Phi(d_2), \quad (4.18) \quad \text{eq:BS-Call-Rho}$$

$$\rho_{\text{Put}} = -K\tau e^{-r\tau}\Phi(-d_2) \quad (4.19) \quad \text{eq:BS-Put-Rho}$$

となる。

ベガ ( $v$ )

ボラティリティの変化に対するそのオプションのプレミアムの感応度（変化率）を**ベガ**と呼ぶ。

配当のないヨーロピアン・オプションのベガは、<sup>eq:BS-Call</sup>(4.8)式・<sup>eq:BS-Put</sup>(4.9)式をそれぞれ  $\sigma$  で偏微分すればよく、

$$v = v_{\text{Call}} = v_{\text{Put}} = S\sqrt{\tau}\varphi(d_1) \quad (4.20) \quad \text{eq:BS-Vega}$$

となる。

## グリークスの特徴

グリークスについて、特徴を挙げておく。具体的にはグラフを描画し、確認されたい。

- オプションが ITM になるにつれて、オプション価格は原資産と殆ど 1 対 1 に動くようになり、デルタの絶対値は 1 に近づく。一方、オプションが OTM になるにつれ、オプション価格はゼロに近づくと同時に原資産価格の変化に対する感応度は鈍くなり、デルタは 0 に近づく。
- ガンマは ATM 近辺で最大になり、OTM・ITM になるに従って次第にゼロに近づいていく。
- セータは ITM や OTM より ATM で、また ATM の時には満期が近づくにつれてその絶対値はより大きくなり、プレミアムの減少スピードが早くなる。

## デルタヘッジ

デルタはオプションを 1 単位持つことと、原資産を  $\Delta$  単位持つことは微小な時間内での価格変化によるエクスポージャーが同じであることを意味している。よって、オプション 1 単位のロングポジションに対して、原資産を  $\Delta$  単位ショートすることで、微小な時間内では価格変動リスクを打ち消すことが出来る。この考え方を**デルタ・ヘッジ**という。

**例 4.7.** デルタ 0.5 の A 社株のコールオプションを 20 単位 (1,000 株単位) 保有していたとする。このとき、オプション 20 単位で  $20 \times 1,000 = 20,000$  株が買えるものとする、 $0.5 \times 20,000 = 10,000$  株の株式を売却することで、株式の価格変化に対するオプションの価格変動リスクをヘッジすることが出来る。

**注意 9.** 原資産をオプションでデルタ・ヘッジしても、それはあくまで微小な時間内で価格変動リスクを排除したに過ぎず、(1) 時間経過によりデルタ自体も変化すること、(2) 原資産価格の変化によりデルタの感応度も変化することもあり、原資産のポジションの大きさを適宜調整しなければならない。

時間の経過に従ってヘッジ量を定期的に調整するヘッジ手法を**動的ヘッジ・ダイナミックヘッジ**と呼ぶ<sup>5</sup>。

デルタ  $\Delta$  に加えて、ガンマ  $\Gamma$  を用いることで、原資産価格の変化  $\partial S$  に対するプレミアムの変化は以下のように近似できる：

$$\Delta \cdot \partial S + \frac{1}{2} \Gamma \cdot (\partial S)^2. \quad (4.21)$$

これを利用することで、動的ヘッジをより精緻に実施することが可能となる。

<sup>5</sup>ダイナミックヘッジについて詳細には Nassim Nicholas Taleb, *Dynamic Hedging: Managing Vanilla and Exotic Options*, 1997, Wiley などがある。

## 4.2.6 VBA によるユーザー関数の作成

```
Option Explicit
```

```
Public Function BS_Calc_d1(S As Double, K As Double, _  
                        Tau As Double, r As Double, Sigma As Double)  
    BS_Calc_d1 = (Log(S / K) + (r + Sigma * Sigma * 0.5) * Tau) / (Sigma * Sqr(Tau))  
End Function
```

```
Public Function BS_Calc_d2(S As Double, K As Double, _  
                        Tau As Double, r As Double, Sigma As Double)  
    BS_Calc_d2 = (Log(S / K) + (r - Sigma * Sigma * 0.5) * Tau) / (Sigma * Sqr(Tau))  
End Function
```

```
Public Function BS_Call(S As Double, K As Double, _  
                    Tau As Double, r As Double, Sigma As Double)  
    Dim d1 As Double  
    Dim d2 As Double  
    d1 = BS_Calc_d1(S, K, Tau, r, Sigma)  
    d2 = BS_Calc_d2(S, K, Tau, r, Sigma)  
    BS_Call = S * Application.WorksheetFunction.Norm_S_Dist(d1, True) _  
        - K * Exp(-r * Tau) * Application.WorksheetFunction.Norm_S_Dist(d2, True)  
End Function
```





## 第5章 市場リスクの基礎

### 5.1 Value at Risk

市場リスク計測の基本的なものとして、Value at Risk (VaR) を概説する。VaR とは、「 $N$  営業日のうちに、 $\alpha\%$  の確率で被る損害はいくらか」を表すもので、この VaR を用いることで、

#### 5.1.1 ポートフォリオの VaR

まずは、簡単な例を紹介しよう。1 日の価格変化 (標準偏差:  $\sigma$ ) が 2% の資産を 10 億円保有しているとする。この資産の 10 日間の信頼区間 99% の VaR を計算してみよう。

まず、10 日間の資産の価格変動の標準偏差を計算すると

$$1,000,000,000 \times 2\% \times \sqrt{10} = 63,245,553$$

となる。今回は信頼区間 99% の VaR を求めたい。累積密度を思い出せば、 $2.33\sigma$  の点がちょうど 1% 点となることから、99% の 10 日間 VaR は

$$2.33 \times 63,245,553 = 147,362,139$$

つまり、約 1 億 4,700 万円となる。

2 資産の場合は、どうなるだろうか。資産  $A$ ,  $B$  の 2 資産を保有する場合について考察しよう。 $A$ ,  $B$  の保有割合を  $\theta_A, \theta_B$  ( $\theta_A + \theta_B = 1, \theta_A \geq 0, \theta_B \geq 0$ ) と標準偏差を  $\sigma_A, \sigma_B$  で表し、相関係数を  $\rho$  とする。この場合、 $A$  を  $\theta_A$ ,  $B$  を  $\theta_B$  保有するポートフォリオの標準偏差は

$$\sigma_{\text{Port}} = \sqrt{(\theta_A \sigma_A)^2 + (\theta_B \sigma_B)^2 + 2\rho \theta_A \theta_B \sigma_A \sigma_B}$$

となるのであった。よって、先ほどの手順に沿って計算すれば 10 日間の 99% VaR は、ポートフォリオを 10 億円保有していたとすると、

$$1,000,000,000 \times \sigma_{\text{Port}} \times \sqrt{10} \times 2.33$$

によって計算できる。

#### 5.1.2 分散共分散法

**例 5.1.**  $t_1, t_2$  の 2 時点へマッピングしたポジションを  $\pi_1 = 10$  億円,  $\pi_2 = -5$  億円とし、ボラティリティと相関係数をそれぞれ  $\sigma_1 = 10\%$ ,  $\sigma_2 = 20\%$ ,  $\rho_{12} = 0.6$  とすれば

$$\pi = \begin{pmatrix} 10 \\ -5 \end{pmatrix}, \Sigma = \begin{pmatrix} 10\% & 0 \\ 0 & 20\% \end{pmatrix}, R = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix}$$

と表せわせ，さらに計測期間  $T$  を1日とすると  $T = 1/252$ . 従って，求める VaR は

$$\text{VaR}(\pi) = 2.33 \times \sqrt{\frac{1}{252}} \times \sqrt{\pi^\top \Sigma^\top R \Sigma \pi} = 0.13$$

となり，計測期間1日で信頼区間99%の VaR は約1,300万円となる．

## 第6章 信用リスクの基礎



## 第II部

# 業務に向けた基礎知識・基礎技術



## 第7章 時系列解析（基礎）

TSA1

本章では、時系列解析に関する基礎的事項を解説する。時系列解析における厳密な理論については Hamilton<sup>hamilton:tsa</sup>[3] を参照すれば良い。時系列解析の基礎的な事項の学習においては、日本アクチュアリー会<sup>actuary:PSMText</sup>[5] が安価で内容も優しい。[3] は発展的な内容を含めて幅広く、深い洞察を与えており（但し和訳書は絶版となっている）、その訳者である沖本氏による<sup>okimoto:tsa</sup>[4] は経済・ファイナンスデータに焦点を当てた解説を与えている。本章では、ヒストリカルなファイナンスデータが与えられたときに、それを適当な時系列モデルでモデル化し、シミュレーション出来るようになることを目標とする。

### 7.1 準備

本稿では、株価や為替レートのように、時間の推移とともに観測されるデータを**時系列データ**と呼ぶことにする。このような時系列データは、各時点で観測された値の集まりで、時点  $t$  に観測されたデータを  $y_t$  で表すこととする。時系列データと言に纏めても、それがもつ特徴は大きく異なり、グラフへプロットした場合にその挙動・性質は様々である。時系列解析における1つの目的は、このようなデータが持っている多様な特徴を記述できるモデルを構築することにある。そして、それらのモデルをもとに、目的に応じた分析を行っていくことである。

#### 7.1.1 言葉の定義

時系列解析を行う上で、必要となる用語について説明する。

時系列データそのものは、**原系列**と呼ばれることが多い。上に述べてきた時系列解析の目的は、殆どこの原系列の性質を明らかにすることだが、実際に扱う上では、原系列に何らかの変換を施したデータに対して分析を行うことが多い。そこで、変換の手法についていくつか述べておく。

経済データには、値が大きくなるにつれてばらつきが大きくなるデータが多く、後述する「定常性」という仮定が満たされないことがある。そのようなデータについて、対数変換を行うことで、その問題を解決できることが多く、原系列に対して、対数系列を施した系列を**対数系列**とよび、 $\log y_t$  と表記される。

また、観測時点間のデータの差も多様される。例えば、1日・1週・1年での株価の変化量・変化率がその例である。変化量を表す場合には  $y_t - y_{t-1}$  という系列を用いられ、これを**差分系列**あるいは**階差系列**などと呼び  $\Delta y_t$  で表す。変化率に興味がある場合には、通常の変化率  $(y_t - y_{t-1})/y_{t-1}$  を使ったり、**対数差分系列**  $\Delta \log y_t := \log y_t - \log y_{t-1}$  を使ったりする。対数差分系列を用いるのは、変化分が小さいとするときの1次の Taylor 展開から

$$\Delta \log y_t \approx \frac{y_t - y_{t-1}}{y_{t-1}}$$

が成り立つからである。

データにある種の「調整」を行うこともある。代表的なものとして**季節調整**があり、これは例えば、季節の変動によるトレンド変化などの影響を受ける時系列データについて、季節変動そのものの影響を取り除いた動向を知りたい場合に、季節の影響を排除する必要がある。そこで、原系列から季節変動を取り除いた**季節調整済み系列**を使ったりする。ただし、季節変動の影響を適切に抽出するということは別の大きな問題として存在するため、本稿でこれについて扱うことはない。

### 7.1.2 基本統計量

系列  $\{y_t\}$  について、その期待値を  $\mathbf{E}[y_t]$ 、分散を  $\text{Var}[y_t]$ 、 $x_t$  と  $y_t$  の共分散を  $\text{Cov}[x_t, y_t]$  で表す。時系列解析に特有の統計量について定義しておく。

**定義 7.1** (自己共分散). 同一の時系列データにおける異時点間の共分散を**自己共分散 (autocovariance)** といい、 $k$  時点間の自己共分散を  $k$  次の自己共分散といい

$$\gamma_{k,t} = \text{Cov}[y_t, y_{t-k}] = \mathbf{E}[(y_t - \mathbf{E}[y_t])(y_{t-k} - \mathbf{E}[y_{t-k}])]$$

で定義する。

自己共分散の解釈は、共分散の場合と同様にすることができて、例えば1次の自己共分散が正であれば、1時点離れたデータは期待値を基準として同じ方向に動く傾向があり、1次の自己共分散が負であれば、1時点離れたデータは期待値を基準として逆の方向に動く傾向がある。0次の自己共分散はまぎれもなく分散そのものである。

自己共分散は、異時点間のデータのふるまいについて特徴が掴めるという利点がある一方で、値が単位に依存してしまう点で、異なるデータ間での比較が出来ない。そこで、変数間の関係の強弱を図ることができるよう、自己共分散を基準化した自己相関を定義する。

**定義 7.2** (自己相関).  $k$  次の自己相関は

$$\rho_{k,t} := \frac{\text{Cov}[y_t, y_{t-k}]}{\sqrt{\text{Var}[y_t]\text{Var}[y_{t-k}]}} = \frac{\gamma_{k,t}}{\sqrt{\gamma_{0,t}\gamma_{k,t}}}$$

で定義される。

定義より明らかに  $\rho_{0,t} = 1$  で、 $|\rho_{k,t}| \leq 1$  となる。自己相関を  $k$  の関数として見たもの（つまり、 $k$  を与えると  $k$  次の自己相関がかえってくる関数として解釈したもの）を自己相関関数といい、これをグラフにプロットしたものを**コレログラム (correlogram)** という。

### 7.1.3 定常性

時系列モデルについて論じるうえで、根幹にあるのは「定常性」とよばれる概念である。定常性の仮定の下で、基礎的な時系列モデルが構築された上で、それらのモデルをもとに非定常なモデルが構築される。

定常性は、同時分布や基本統計量が時間に依らないということを主張するものであり、何を不変とするかにより弱い概念・強い概念が存在する。



**定義 7.3** (弱定常). 任意の  $t, k$  に対して

$$\mathbf{E}[y_t] = \mu, \quad (7.1) \quad \text{stable:1}$$

$$\text{Cov}[y_t, y_{t-k}] = \mathbf{E}[(y_t - \mu)(y_{t-k} - \mu)] = \gamma_k \quad (7.2) \quad \text{stable:2}$$

が成り立つとき、仮定は**弱定常 (weak stationary)** であるという.

定義より弱定常は、過程の期待値と自己共分散が時間を通じて一定であることを要求している. 式 (7.2) からわかるように、定常過程においては、自己共分散は時点に依存せずに時間差  $k$  のみに依存することになる. よって、任意の  $k$  に対して、 $\gamma_k = \gamma_{-k}$  が成り立っている. 過程が弱定常のとき、自己相関は

$$\frac{\gamma_{k,t}}{\sqrt{\gamma_{0,t}\gamma_{k,t}}} = \frac{\gamma_k}{\gamma_0} = \rho_k$$

と、時点に依存しないこともわかり、 $\rho_k = \rho_{-k}$  が得られる.

**定義 7.4** (強定常). 任意の  $t, k$  に対して  $(y_t, y_{t+1}, \dots, y_{t+k})$  の同時分布が同一となるとき、過程は**強定常 (strict stationary)** といわれる.

以下の性質が知られている<sup>1</sup>.

**定理 7.1.** 過程の分散が有限であるなら、強定常過程は弱定常過程となる.

以下、特に断らない限りは弱定常過程を定常過程と呼んで議論をすすめる.

最も基礎的な (弱) 定常過程は、

$$y_t = \mu + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2) \quad (7.3) \quad \text{eq:Basic}$$

である. このモデルは  $\mu, \sigma$  という 2 つのパラメータを持ち、それぞれ過程の期待値と標準偏差を表す. このモデルによって、パラメータを操作すれば、任意の平均・分散を持つ時系列を作成することはできるのだが、自己相関を持たず、ファイナンスのデータを表現するには不十分である. そこで、以下に述べる AR や MA あるいは、ARMA モデルなどを紹介し、ファイナンスのデータを表現する手法を説明していこう.

## 7.2 AR 過程・MA 過程・ARMA 過程

経済データの中には、自己相関をもつデータが数多く存在する. そこで、モデル化する上で、自己相関を表現する必要があるわけだが、これを実現する方法として簡単に考えられるものが 2 つある.

1 つ目は、 $y_t$  のモデルに 1 時点前の  $y_{t-1}$  を含める方法である. 例えば、

$$y_t = a \cdot y_{t-1} + b$$

の形で表現する場合である. このとき、 $y_t$  と  $y_{t-1}$  が相関を持つことは明らかであろう. この手法は自己回帰 (AR:Auto Regression) の考え方である.

<sup>1</sup>証明は[3]を参照.

2 つ目は、 $y_t$  と  $y_{t-1}$  に共通の成分を持たせる方法がある．例えば、

$$\begin{cases} y_t = a + b, \\ y_{t-1} = c + b \end{cases}$$

と言った具合に共通の要素  $b$  をそれぞれに持たせる手法である．このようにモデル化をすれば、共通の成分である  $b$  を通じて、 $y_t$  と  $y_{t-1}$  が相関を持つことが期待される．この手法は移動平均 (MA: Moving Average) の考え方に従う．

以下、本節では、AR 過程・MA 過程について性質を紹介する．

### 7.2.1 自己回帰過程：AR( $p$ )

自己回帰 (AR) 過程は、過程が自身の過去に回帰された形で表現されるものである．1 次の AR 過程 (AR(1)) は

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2) \quad (7.4) \quad \boxed{\text{eq: AR1}}$$

で定義される．AR(1) モデルは、トレンド・ノイズによるモデル式 (7.3) に  $\phi_1 y_{t-1}$  という項が追加された形をしており、上述の通り、 $y_t, y_{t-1}$  が相関を持つことは明らかであろう．AR(1) モデルは (7.3) によるモデルよりも  $\phi_1$  というパラメータが多くなっており、このパラメータが過程の定常性や自己相関の強さなどを決定する要因となる（後述）．

以下、定常な AR(1) 過程の性質を見てみよう．

**命題 7.2.** 定常な AR(1) 過程について、 $\mathbf{E}[y_t] = \phi_0 / (1 - \phi_1)$ ．

*Proof.*  $\mathbf{E}[y_t] = \mathbf{E}[\phi_0 + \phi_1 y_{t-1} + \varepsilon_t] = \phi_0 + \phi_1 \mathbf{E}[y_{t-1}] + \mathbf{E}[\varepsilon_t] = \phi_0 + \phi_1 \mathbf{E}[y_{t-1}] + 0$ ． $y_t$  は定常であるから、 $\mathbf{E}[y_t] = \mathbf{E}[y_t] =: \mu$  とし、先ほどの式を書き直せば  $\mu = \phi_0 + \phi_1 \mu$  となり、これより  $\mathbf{E}[y_t] = \phi_0 / (1 - \phi_1)$  を得る．  $\square$

**命題 7.3.** 定常な AR(1) 過程について、 $\text{Var}[y_t] = \gamma_0 = \sigma^2 / (1 - \phi_1^2)$ ．

*Proof.*  $\text{Var}[y_t] = \text{Var}[\phi_0 + \phi_1 y_{t-1} + \varepsilon_t] = \phi_1^2 \text{Var}[y_{t-1}] + \text{Var}[\varepsilon_t] + 2\phi_1 \text{Cov}[y_{t-1}, \varepsilon_t] = \phi_1^2 \text{Var}[y_{t-1}] + \sigma^2$ ． $y_t$  は定常であるから、 $\gamma_0 = \text{Var}[y_t] = \text{Var}[y_{t-1}]$  であるので、期待値の場合の議論と同様に  $\text{Var}[y_t] = \sigma^2 / (1 - \phi_1^2)$  を得る．  $\square$

**命題 7.4.** 定常な AR(1) 過程について、 $\gamma_k = \phi_1 \gamma_{k-1}$ ．

*Proof.*  $\gamma_k = \text{Cov}[y_t, y_{t-k}] = \text{Cov}[\phi_1 y_{t-1} + \varepsilon_t, y_{t-k}] = \text{Cov}[\phi_1 y_{t-1}, y_{t-k}] + \text{Cov}[\varepsilon_t, y_{t-k}]$  ここで、 $\varepsilon_t$  と  $y_{t-k}$  は独立だから  $\gamma_k = \phi_1 \gamma_{k-1}$ ．  $\square$

ここで、自己共分散  $\gamma_k$  について両辺を  $\gamma_0$  で割ると、

$$\rho_k = \phi_1 \rho_{k-1}$$

という関係が得られる．この方程式を Yule-Walker の方程式といい、AR 過程の自己相関が  $y_t$  が従う AR 過程と同一の係数をもつ差分方程式に従うことを示している．Yule-Walker の方程式と、 $\rho_0 = 1$  という事実を用いれば、自己相関は順次求めることができ、

$$\rho_k = \phi_1 \rho_{k-1} = \phi_1^2 \rho_{k-2} = \cdots = \phi_1^k \rho_0 = \phi_1^k$$

という関係が得られる．この結果と，定常性（AR(1) のいま， $|\phi_1| < 1$ ）から AR(1) 過程の自己相関の絶対値は指数的に減衰していくことが分かる．もちろん， $\phi_1$  の符号に応じて，それが単調あるいは振動するという違いは表れるが，減衰していくことについては変わらない．纏めると，AR(1) 過程は 2 次以降の自己相関を持つ．しかし，2 次以降の自己相関の挙動については，単調あるいは振動するという違いはあれども，制約的な挙動を示しており，複雑な自己相関構造を記述するには，一般的なモデルが必要となる．

続いて，AR(1) 過程を一般化し， $p$  次の自己相関過程を考える． $p$  次の AR 過程 ( $AR(p)$ ) は

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2) \quad (7.5) \quad \text{eq:ARp}$$

で定義される．これまでに確認した定常な AR(1) 過程の性質は，定常な  $AR(p)$  過程の場合にも一般化することができ，それを纏めておく．

#### 定理 7.5. 定常 $AR(p)$ 過程

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2)$$

は以下の性質を持つ：

1.  $\mu = \mathbf{E}[y_t] = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}$ ,
2.  $\gamma_0 = \text{Var}[y_t] = \frac{\sigma^2}{1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \cdots - \phi_p \rho_p}$ ,
3.  $k(\geq 1)$  次の自己共分散  $\gamma_k$  と自己相関  $\rho_k$  は  $y_t$  が従う AR 過程と同じ係数を持つ，以下の  $p$  次差分方程式に従う．特に， $\rho_k$  の式は Yule-Walker 方程式とよばれる：

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p}, \quad (7.6)$$

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p}. \quad (7.7)$$

4. AR 過程の自己相関は指数減衰．

ここまで，AR 過程について定常性を過程してきたが，定常性を持つ条件としては以下が知られている．

#### 命題 7.6. $AR(p)$ 過程

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2)$$

に対し，以下の式を特性方程式という：

$$\phi(x) = 1 - (\phi_1 x + \phi_2 x^2 + \cdots + \phi_p x^p). \quad (7.8) \quad \text{eq:CFAR}$$

$\phi(x) = 0$  の解の絶対値がすべて 1 より大きいとき， $AR(p)$  過程は定常性を持つ．

### 7.2.2 移動平均過程：MA( $q$ )

移動平均 (MA) 過程は，式 (7.3) の拡張で，ノイズの線形和の形で表される．1 次の MA 過程 (MA(1)) は

$$y_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1}, \varepsilon_t \sim N(0, \sigma^2) \quad (7.9) \quad \text{eq:MA1}$$

で定義される。MA(1) モデルは、式 (7.3) に  $\theta_1 \varepsilon_{t-1}$  という項が追加された形をしており、1 期前に遡ると

$$y_{t-1} = \theta_0 + \varepsilon_{t-1} + \theta_1 \varepsilon_{t-2}$$

となることから、 $y_t$  と  $y_{t-1}$  が共通の  $\varepsilon_{t-1}$  を持つ。これによって、 $y_t, y_{t-1}$  の間に相関が生じるのである。これは即ち、MA(1) モデルは、1 次の自己相関をもつモデルとなっていることに他ならない。

また、MA(1) モデルは式 (7.3) から  $\theta_1$  というパラメータが増えており、このパラメータが 1 次の自己相関の度合を決定する（後述）。以下、MA(1) 過程の性質を見てみよう。

**命題 7.7.**  $E[y_t] = \theta_0$ .

*Proof.*  $E[y_t] = E[\theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1}] = E[\theta_0] + E[\varepsilon_t] + E[\theta_1 \varepsilon_{t-1}] = \theta_0 + 0 + 0 = \theta_0$ . □

**命題 7.8.**  $\text{Var}[y_t] = \gamma_0 = (1 + \theta_1^2)\sigma^2$ .

*Proof.*  $\text{Var}[y_t] = \text{Var}[\theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1}] = \text{Var}[\varepsilon_t] + \theta_1^2 \text{Var}[\varepsilon_{t-1}] + 2\theta_1 \text{Cov}[\varepsilon_t, \varepsilon_{t-1}] = \sigma^2 + \theta_1^2 \sigma^2 + 0 = (1 + \theta_1^2)\sigma^2$ . □

gamma1

**命題 7.9.**  $\gamma_1 = \theta_1 \sigma^2$ .

*Proof.*  $\gamma_1 = \text{Cov}[y_t, y_{t-1}] = \text{Cov}[\theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1}, \theta_0 + \varepsilon_{t-1} + \theta_1 \varepsilon_{t-2}] = \text{Cov}[\varepsilon_t, \varepsilon_{t-1}] + \text{Cov}[\varepsilon_t, \theta_1 \varepsilon_{t-2}] + \text{Cov}[\theta_1 \varepsilon_{t-1}, \varepsilon_{t-1}] + \text{Cov}[\theta_1 \varepsilon_{t-1}, \theta_1 \varepsilon_{t-2}] = \theta_1 \text{Cov}[\varepsilon_{t-1}, \varepsilon_{t-1}] = \theta_1 \sigma^2$ . □

これによって、MA(1) 過程の 1 次の自己相関  $\rho_1$  が

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\theta_1}{1 + \theta_1^2}$$

となることも分かる。

**注意 10.**  $\theta_1 = \pm 1$  のとき MA(1) 過程の自己相関の絶対値が最大値 0.5 となる。言い換えると、1 次の自己相関の絶対値が 0.5 より大きな過程は MA(1) 過程でモデル化することは出来ない。

**注意 11.** 命題 7.9 の証明をたどれば、MA(1) 過程の 2 次以降の自己共分散については、 $\gamma_k = 0 (k \geq 2)$  となることが分かる。従って、2 次以降の自己相関についても 0 となることがわかる。よって、MA(1) 過程では、1 次の自己相関をモデル化することは出来るが、2 次以降の自己相関を表現することは出来ない。

これまでの主張を見ると、MA(1) 過程の期待値や自己共分散は時点  $t$  に依らないことがわかる。これは、MA(1) 過程が定常であるということであり、**実は、MA(1) 過程はパラメータの値のよらず、常に定常となる**ことがわかる。

続いて、MA(1) 過程を一般化し、 $q$  次の移動平均過程を考える。 $q$  次の MA 過程 (MA( $q$ )) は

$$y_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} = \theta_0 + \varepsilon_t + \sum_{k=1}^q \theta_k \varepsilon_{t-k} \quad (7.10) \quad \text{eq:MAq}$$

で定義される。これまでに確認した MA(1) 過程の性質は、MA( $q$ ) 過程の場合にも一般化することができ、それを纏めておく。

**定理 7.10.** MA( $q$ ) 過程 (7.10) は以下の性質を満たす：

1.  $\mathbf{E}[y_t] = \theta_0$ ,
2.  $\text{Var}[y_t] = (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma^2$ ,
3.  $k = 1, 2, \dots, q$  なる  $k$  について,  $\gamma_k = (\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q)\sigma^2$ ,  $\rho_k = (\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q)/(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)$ ,
4.  $k \geq q + 1$  なる  $k$  について,  $\gamma_k = 0, \rho_k = 0$ ,
5. MA( $q$ ) 過程は常に定常

### MA 過程の反転可能性

MA 過程は常に定常となることから, AR 過程に比べると扱い易い側面があるが, MA 過程に対してはまた別の問題が存在する.

任意の MA 過程に関して, 同一の期待値と自己相関をもつ異なる過程が複数存在することである. 例えば,

$$y_t = \varepsilon_t + \theta\varepsilon_{t-1}, \quad \varepsilon_t \sim N(0, \sigma^2), \quad (7.11) \quad \text{eq:MA1-1}$$

$$\tilde{y}_t = \tilde{\varepsilon}_t + \frac{1}{\theta}\tilde{\varepsilon}_{t-1}, \quad \varepsilon_t \sim N(0, \theta^2\sigma^2), \quad (7.12) \quad \text{eq:MA1-2}$$

の 2 つの MA(1) 過程を考える. このとき, 各種統計量について見てみると, 期待値, 分散, 1 次の自己共分散がすべて一致している. 時系列モデルを用いる 1 つの目的に, データの平均的な挙動・自己相関構造をモデル化することがあることは, これまでに述べてきたことであるが, その観点からすると, 同一の期待値・自己相関構造を持つ MA 過程が複数<sup>2</sup>存在するとき, どの MA 過程を用いるべきかが定かではないという問題が出てくる. そこで, 1 つの基準となるのが「反転可能性」とよばれる概念である.

**定義 7.5** (反転可能性). MA 過程が AR( $\infty$ ) 過程に書き直せるとき, MA 過程は**反転可能** (invertible) という.

MA 過程が反転可能であれば,  $\varepsilon_t$  は過去の  $y_t$  の関数として表現でき, さらに過去の  $y$  を用いて  $y_t$  を予測した時の予測誤差として解釈することができる. 同一の期待値・自己相関構造をもつ MA 過程は複数存在するが, 反転可能なものはただ一つとなることが知られている.

MA 過程が反転可能かを判定するにあたっては, 次の主張が知られている.

**定理 7.11.** MA( $q$ ) 過程 (7.10) の反転可能条件は, 以下の**特性方程式** (AR の場合と区別するべく MA 特性方程式とも呼ぶ) のすべての解の絶対値が 1 より大きいことである：

$$\theta(z) := 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q = 0. \quad (7.13) \quad \text{eq:CFMA}$$

**例 7.1** (MA(1) の反転可能条件と AR( $\infty$ ) 表現). MA(1) 過程 (7.9) の特性方程式は  $1 + \theta_1 z = 0$  となるので, 反転可能条件は  $|z| = |\theta_1^{-1}| > 1$  すなわち  $|\theta_1| < 1$  で与えられる. MA(1) が反転可能であるときに, AR( $\infty$ ) で書き直せることを確認しておこう.  $\varepsilon_t$  を AR 過程で書き直すと,

$$\varepsilon_t = -\theta_1 \varepsilon_{t-1} + y_t = (-\theta)m\varepsilon_{t-m} + \sum_{k=0}^{m-1} (-\theta)^k y_{t-k} \rightarrow \sum_{k=0}^{\infty} (-\theta)^k y_{t-k}$$

<sup>2</sup>一般に  $2^q$  個存在することが知られている

となる． $(-\theta)^m \varepsilon_{t-m} \rightarrow 0 (m \rightarrow \infty)$  であるから，MA(1) 過程 (7.9) は，反転可能であるとき，

$$y_t = - \sum_{k=1}^{\infty} (-\theta)^k y_{t-k} + \varepsilon_t$$

AR( $\infty$ ) の表現を持つ．

### 定常 AR 過程の MA( $\infty$ ) 表現

最後に，定常 AR 過程は，MA( $\infty$ ) 過程で表現できることを示しておく．主張として書くなら以下の通り．

**定理 7.12.** AR 過程が定常であることと，AR 過程を MA 過程で表現できることは同値．

簡単のため，AR(1) の特に  $\phi_0 = 0$  の場合についてこれを考察することで証明に替える．

AR(1) 過程

$$y_t = \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2)$$

について，繰り返し代入することで

$$\begin{aligned} y_t &= \phi_1 y_{t-1} + \varepsilon_t \\ &= \phi_1 (\phi_1 y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \dots \\ &= \phi_1^m y_{t-m} + \sum_{k=1}^{m-1} \phi_1^k \varepsilon_{t-k} \end{aligned}$$

となる．ここで，AR(1) が定常，即ち  $|\phi_1| < 1$  なら， $\phi_1^m y_{t-m} \rightarrow 0 (m \rightarrow \infty)$  となることから，

$$y_t = \sum_{k=0}^{\infty} \phi_1^k \varepsilon_{t-k}$$

と表現することが出来る．一方  $|\phi_1| \geq 1$  であるときには，MA 過程で表現することが出来ない．したがって，AR 過程の定常条件は MA 過程で表現できることと同値であることがわかった．

一般に，AR( $p$ ) 過程

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

について

$$y_t = \xi_0 + \varepsilon_t + \sum_{k=1}^{\infty} \xi_k \varepsilon_{t-k}$$

と表現される．

### 7.2.3 ARMA 過程：ARMA( $p, q$ )

ARMA(Auto-regressive moving average) 過程は，自己回帰項と移動平均項の両方を含んだ過程である，( $p, q$ ) 次 ARMA 過程 ( $ARMA(p, q)$ ) は

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \varepsilon_t \sim N(0, \sigma^2) \quad (7.14)$$

eq:ARMApq

で定義される. ARMA( $p, q$ ) 過程は, AR 過程と MA 過程の性質をあわせもっており, 両過程の性質のうち強いほうが ARMA 過程の性質となる. これはどういうことかということ, MA 過程は常に定常であるが, AR 過程は定常になるとは限らない. この場合には, AR 過程の性質が残って, ARMA 過程が定常になるとは限らないということになる. 定常性を持つための条件は, やや砕けた形になるが

**定理 7.13.** ARMA 過程が定常であるための条件は, AR 過程部分が定常であることである.

という主張で得られる. 今後は, 定常な ARMA 過程に絞って説明していく. ARMA 過程の性質については, AR 過程・MA 過程の性質を鑑みれば殆ど明らかなものであり, 以下のように纏められる.

**定理 7.14.** 定常 ARMA( $p, q$ ) 過程 (7.14) <sup>eq:ARMApq</sup> は以下の性質をもつ.

1.  $\mu = \mathbf{E}[y_t] = \frac{c}{1-\phi_1-\phi_2-\cdots-\phi_p},$
2.  $q+1$  次以降の自己共分散と自己相関は  $y_t$  の AR 部分と同一の係数を持つ  $p$  次差分方程式 (Yule-Walker 方程式) に従う. 即ち

$$\begin{aligned}\gamma_k &= \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} + \cdots + \phi_p\gamma_{k-p}, & k \geq q+1, \\ \rho_k &= \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \cdots + \phi_p\rho_{k-p}, & k \geq q+1,\end{aligned}$$

3. ARMA 過程の自己相関は指数的に減衰する.

**注意 12.**  $q$  次までの自己共分散と自己相関は AR・MA の影響があるため, 一般的に表現することが難しい.

MA 過程を含むことから, ARMA 過程においても反転可能性の問題が生じる. ただし, 定常性の場合の議論とほぼ同様に, 反転可能性の条件を理解することができる. 実際, ARMA 過程が AR 過程で表現できるためには, AR 過程の部分はすでに AR 過程で表現できているため, MA 過程の部分が反転可能であれば良いことになる. したがって MA 過程部分について, MA 特性方程式を考え, そのすべての解の絶対値が 1 より大きければ, ARMA 過程は反転可能となる.

## 7.3 モデルの推定と選択

Choice

ARMA モデルの推定法として, 最尤法を説明する.

### 7.3.1 最小 2 乗法

まず, ARMA モデルで最尤法を使う理由を知るため, AR モデルに対する最小 2 乗法 (OLS; ordinary least squares) を説明しよう.

OLS の考えかたは非常にシンプルであり, モデルが説明できない部分を最小化しようという考え方である. 以下, AR(1) モデル <sup>eq:AR1</sup> (7.4)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2)$$

を用いて説明する．OLSでは、推定するモデルのことを回帰モデルといい、左辺の変数  $y_t$  を被説明変数、 $\varepsilon_t$  のことを誤差項とよび、誤差項以外の右辺の変数（ここでは、 $\phi_0, y_{t-1}$ ）のことを説明変数という．今回のように定数以外の説明変数が1つしかないような回帰モデルを「単回帰モデル」、 $p \geq 2$  の  $AR(p)$  もでるのように、定数以外の説明変数が複数ある回帰モデルは「重回帰モデル」という．

回帰モデルにおける  $\phi_0, \phi_1$  の推定量を  $\tilde{\phi}_0, \tilde{\phi}_1$  とすると、OLSでは、モデルが説明できない部分（残差という）

$$e_t := y_t - \tilde{\phi}_0 - \tilde{\phi}_1 y_{t-1}$$

が平均的に0に近くなるようにすることを考える．しかし、残差は正負の値をとるため、単純な足し合わせが出来ないため、残差平方和 (SSR; sum of squared residuals)

$$SSR = \sum_{t=1}^T e_t^2 = \sum_{t=1}^T (y_t - \tilde{\phi}_0 - \tilde{\phi}_1 y_{t-1})^2 \quad (7.15) \quad \text{eq:SSR}$$

が最小になるよう  $\tilde{\phi}_0, \tilde{\phi}_1$  を選択する．具体的には、式 (7.15) の両辺を  $\tilde{\phi}_0, \tilde{\phi}_1$  でそれぞれ偏微分すると、

$$\begin{cases} \frac{\partial SSR}{\partial \tilde{\phi}_0} = -2 \sum_{t=1}^T (y_t - \tilde{\phi}_0 - \tilde{\phi}_1 y_{t-1}), \\ \frac{\partial SSR}{\partial \tilde{\phi}_1} = -2 \sum_{t=1}^T y_{t-1} (y_t - \tilde{\phi}_0 - \tilde{\phi}_1 y_{t-1}) \end{cases} \quad (7.16) \quad \text{eq:pSSR}$$

が得られる．OLS 推定量を  $\hat{\phi}_0, \hat{\phi}_1$  とすると、これらは (7.17) の2式が0になるような値であるから、

$$\begin{cases} \sum_{t=1}^T (y_t - \hat{\phi}_0 - \hat{\phi}_1 y_{t-1}) = 0, \\ \sum_{t=1}^T y_{t-1} (y_t - \hat{\phi}_0 - \hat{\phi}_1 y_{t-1}) = 0 \end{cases}$$

を満たす．この方程式を正規方程式といい、この正規方程式を解くことによって、OLS 推定量が

$$\hat{\phi}_0 = \bar{y}_t - \hat{\phi}_1 \bar{y}_{t-1}, \quad \hat{\phi}_1 = \frac{\sum_{t=1}^T (y_t - \bar{y}_t)(y_{t-1} - \bar{y}_{t-1})}{\sum_{t=1}^T (y_{t-1} - \bar{y}_{t-1})^2}$$

で与えられることがわかる．ここで、 $\bar{y}_t = 1/T \sum_{t=1}^T y_t$  である．

$AR(p)$  モデルにおいてもこれを拡張することは容易であるが、ARMA モデルでは、OLS を単純に適用できない．そこで、OLS の代わりに次に説明する最尤法が用いられることが多い．

### 7.3.2 最尤法

最尤法は最小2乗法では推定が難しい複雑な構造をもったモデルの推定に用いられることが多い．コンセプトの違いは、OLS の場合はモデルの説明できない部分が最小になるようにパラメータを選択した一方で、最尤法の場合は、得られた観測地をモデルが最も実現しやすくなるようにパラメータを選択するという点にある．



引き続き、AR(1) モデルを使った場合で考えよう。推定したいパラメータ  $\Theta := (\phi_0, \phi_1, \sigma^2)$  とする。  $y_1, y_2$  の同時密度について、Bayes の公式を使えば

$$f_{Y_2, Y_1 | Y_0}(y_2, y_1 | y_0; \Theta) = f_{Y_2 | Y_1, Y_0}(y_2 | y_1, y_0; \Theta) \cdot f_{Y_1 | Y_0}(y_1 | y_0; \Theta)$$

$$A = f_{Y_2 | Y_1}(y_2 | y_1; \Theta) \cdot f_{Y_1 | Y_0}(y_1 | y_0; \Theta)$$

が成立<sup>3</sup>する。続いて、  $y_1, y_2, y_3$  の同時密度を考えると

$$f_{Y_3, Y_2, Y_1 | Y_0}(y_3, y_2, y_1 | y_0; \Theta) = f_{Y_3 | Y_2}(y_3 | y_2; \Theta) \cdot f_{Y_2 | Y_1}(y_2 | y_1; \Theta) \cdot f_{Y_1 | Y_0}(y_1 | y_0; \Theta)$$

が得られる。これを繰り返すと、一般に

$$f_{Y_n, Y_{n-1}, \dots, Y_1 | Y_0}(y_n, y_{n-1}, \dots, y_1 | y_0; \Theta) = \prod_{k=1}^n f_{Y_k | Y_{k-1}}(y_k | y_{k-1}; \Theta)$$

が成り立つ。最尤法は、このようにして得られた同時密度をもとに計算された尤度あるいは以下のような対数尤度を最大化することによって、パラメータを推定する方法である：

$$\mathcal{L}(\Theta) = \sum_{k=1}^n \ln f_{Y_k | Y_{k-1}}(y_k | y_{k-1}; \Theta). \quad (7.17) \quad \boxed{\text{eq:pSSR}}$$

$\varepsilon_t \sim N(0, \sigma^2)$  のとき、  $y_t | y_{t-1} \sim N(\phi_0 + \phi_1 \cdot y_{t-1}, \sigma^2)$  となるため、

$$f_{Y_t | Y_{t-1}}(y_t | y_{t-1}; \Theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - \phi_0 - \phi_1 \cdot y_{t-1})^2}{2\sigma^2}\right)$$

となるため、  $\mathcal{L}(\Theta)$  に代入すると、AR(1) 過程の対数尤度が

$$\mathcal{L}(\Theta) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \sum_{k=1}^n \left(-\frac{(y_t - \phi_0 - \phi_1 \cdot y_{t-1})^2}{2\sigma^2}\right)$$

となることがわかる。これを最大にするような  $(\phi_0, \phi_1)$  は  $(y_t - c - \phi_0 \cdot y_{t-1})^2$  を最小化するペアで良いわけであるが、これは OLS 推定量そのものである。従って AR(1) モデルにおける係数の最尤推定量は OLS 推定量によって与えられる。また、  $\sigma^2$  の最尤推定量は

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{k=1}^n (y_t - \hat{\phi}_0 - \hat{\phi}_1 \cdot y_{t-1})^2$$

で与えられる。

以上の議論は、ARMA モデルに対しても一般化でき、基本的には  $\mathcal{L}(\Theta)$  の右辺にある  $f_{Y_t | Y_{t-1}}(y_t | y_{t-1}; \Theta)$  を適切な条件付き分布に置き換えればよい。

実際の分析においては、R や Python のパッケージを使ったりすることで尤度関数の計算などは容易に行うことができる。

### 7.3.3 モデルの候補

真のモデルがどのような形をしているのか検討するにあたって、しらみつぶしにすべてのモデルにおいて検討することは現実的ではない。実際、AR・MA ないし ARMA のパターンに限ったとしても、その次数として適切な  $p, q$  が何かという問題に遭遇する。モデルの選択にあたって、標本自己相関等からモデルの候補を選択することからはじめ、最終的には「情報量規準」とよばれるものを用いてモデルを選択するが、モデルを絞ったほうが効率的であることは変わらない。

<sup>3</sup>最後の等号は、AR(1) モデルのときに、直近 1 期の値にしか条件付き分布が依存しないことによる。

### 自己相関・偏自己相関

モデル候補の選択にあたって重要な指標となるのが、標本自己相関や、標本偏自己相関である。

自己相関の検定を行うためには、自己相関の推定値を計算する必要があるが、定常性の仮定のもとでは、期待値や自己相関などの基本統計量は時点に依存しない為さほど難しくない。データから対応する標本統計量を計算することで、自然な推定量を得ることが出来るので、期待値、自己共分散の推定量として

$$\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t,$$

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y}), k = 0, 1, 2, \dots$$

を使って、標本自己相関は

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}, k = 1, 2, 3, \dots \quad (7.18)$$

で与えられる<sup>4</sup>。

標本自己相関は、データから自己相関を計算したものであり、自己相関と同様の性質を持つと期待される。ここで、AR、MA、ARMA の夫々において自己相関がどのような性質を持つかおさらいすると、AR 過程と ARMA 過程については、自己相関の絶対値は指数的に減衰するのであった。一方、MA( $q$ ) 過程の自己相関は  $q+1$  次以降 0 になるという性質があった。これらの性質から分かることとしては、**標本自己相関が  $q+1$  次で切断され、それ以降 0 に近い値をとって** いれば MA( $q$ ) 過程の可能性が高いということである。

では、AR 過程と ARMA 過程についてはどのように区別するか？そこで新たに偏自己相関という概念を導入する。 $k$  次自己相関は単に  $y_t$  と  $y_{t-k}$  の相関を計算したものであったが、 $k$  次偏自己相関は  $y_t$  と  $y_{t-k}$  から、 $y_{t-1}, \dots, y_{t-k+1}$  の影響を取り除いたものの間の相関を考えたものである。 $y_t$  と  $y_{t-k}$  は、 $y_{t-1}$  や  $y_{t-2}$  を通じて相関している可能性があるので、間にあるデータの影響を取り除いたうえで、 $y_t$  と  $y_{t-k}$  の相関を評価したものを  $k$  次偏自己相関という。正確な定義は本稿のレベルを越えるため省略するが、 $k$  次偏自己相関  $\alpha_k$  は  $y_t$  を定数と  $y_{t-1}, y_{t-2}, \dots, y_{t-k}$  に線形射影した時の  $y_{t-k}$  の射影係数で定義される。

AR( $p$ ) 過程の偏自己相関について考えると、AR( $p$ ) 過程は、 $y_t$  は  $y_{t-1}, \dots, y_{t-p}$  の線形和で書ける。そのため、 $y_t$  から  $y_{t-1}, \dots, y_{t-p}$  の影響を取り除くと、残るのは自己相関をもたないホワイトノイズ  $\varepsilon_t$  の部分のみとなる。**したがって、AR( $p$ ) モデルにおける  $p+1$  次以降の偏自己相関は 0 となる。**MA( $q$ ) 過程については、AR( $\infty$ ) 過程で書き換えられるのであったから、 $y_t$  から  $y_{t-1}, y_{t-2}, \dots, y_{t-k+1}$  までの影響を取り除いたとしても、 $y_t$  と  $y_{t-k}$  が相関を持つような  $k$  は限りなく存在することになる。ただし、MA 過程を AR( $\infty$ ) 過程で書き換えたときの AR 係数の絶対値は指数的に減衰することが知られているため、MA 過程は無限の偏自己相関を持つが、その絶対値は指数的に減衰していく。これは、ARMA 過程においても同様である。よって、**標本偏自己相関は偏自己相関と同様の性質をもつことが期待されるので、標本偏自己相関の値が  $p+1$**

<sup>4</sup>標本自己共分散の定義では、 $t = k+1, k+2, \dots, T$  の範囲で足し合わせているため、要素の個数  $T-k$  で割っても構わないが、 $T$  で割ることによって標本自己共分散が整定値になるという性質を持つことから  $T$  で割るスタイルのテキストが多い（例えば、[4]）ため、ここではそれを踏襲する。

次で切断され、それ以降で 0 に近い値をとっていることが確認されれば、 $AR(p)$  過程の可能性が高いということになる。

以上を纏めると、標本自己相関と標本偏自己相関を用いて、ある程度最適なモデルが選択できることになる。

モデル	自己相関	偏自己相関
$AR(p)$	減衰	$p + 1$ 次以降 0
ARMA	減衰	減衰
$MA(q)$	$q + 1$ 次以降 0	減衰

**注意 13.** 標本自己相関と標本偏自己相関には、推定誤差を含むということに注意が必要である。従って、真値が 0 であるとしても、標本自己相関や標本偏自己相関が完全に 0 に一致することはまず考えられない。1つの目安は有意に 0 と異なるかを確認するかどうかであるが、真値が 0 であっても、推定値が 0 と有意に異ならない可能性がある。ゆえに自己相関や偏自己相関が 0 であるかの判断は慎重にならなければならない。減衰していることと切断されていることの差についても同様に注意が必要である。

**注意 14.** 標本自己相関と標本偏自己相関だけでは、ARMA 過程の字数  $p, q$  を定めることは出来ない。

これらの注意を受けて、通常は、標本自己相関や標本偏自己相関で保守的に複数のモデル候補を選択した後、客観的な基準を利用して裁量のモデルを選択することが多く、その客観的な基準を与えてくれるのが情報量規準とよばれる指標である。

### 情報量規準

**情報量規準 (information criterion)** は最尤法の推定結果をもとに、最適なモデルを選択する客観的な指標であり、

$$IC = -2\mathcal{L}(\hat{\Theta}) + p(n)k$$

のような形で定義される。ここで、 $\mathcal{L}(\hat{\Theta})$  は対数尤度を最尤推定値  $\Theta$  で評価した最大対数尤度、 $n$  はモデルの推定に用いた標本数、 $p(n)$  は  $n$  の何らかの関数、 $k$  は推定したパラメータの数である。

情報量規準は 2つの部分から構成されており、第 1 項 ( $-2\mathcal{L}(\hat{\Theta})$ ) はモデルの当てはまりを表し、第 2 項 ( $p(n)k$ ) はモデルが複雑となることに対する罰則項の 2つによって構成されている。モデルを複雑にすると、モデルの説明力は上昇する一方でパラメータの推定精度を犠牲とするため、第 1 項・第 2 項が別々の動きをする。これは、モデルを簡素化した場合に逆の動きをし、第 1 項・第 2 項の間でトレードオフが生じるので、モデルが過剰に複雑化することを防ぐことができる。

この情報量規準はどのように見ればよいのか？ 上述のようなトレードオフを捉えているため、パラメータを増やすペナルティ以上にモデルが改善しなくなれば、そのモデルは最適であると考えられる。それは、譲歩言う量規準を最小にするようなモデルのことであり、そこに分析者の主観が入ることを防ぐことができる。

情報量規準は客観的な基準ではあるが、罰則項の関数  $p(n)$  の選択には主観的な部分があり、この関数の違いによって、複数の情報量規準が存在する。本稿では有名な 2つの情報量規準を紹介する。

1つは、**赤池情報量規準 (AIC; Akaike's Information Criterion)** と呼ばれるもので、 $p(n) = 2$  において

$$\text{AIC} = -2\mathcal{L}(\hat{\Theta}) + 2k$$

で定義される。2つ目は **Bayes 情報量規準 (BIC; Bayes' Information Criterion)** と呼ばれるもので、 $p(n) = \ln(n)$  において

$$\text{BIC} = -2\mathcal{L}(\hat{\Theta}) + \ln(n)k$$

で定義される。罰則項に  $n$  が含まれるか否かが本質的な違いであり、 $n \geq 8$  のとき、 $\ln(n) > 2$  となるため、AIC よりも BIC のほうが罰則が強く働く。よって、 $n \geq 8$  のとき、BIC は AIC よりも小さいモデルを選択する傾向がある。これからも分かる通り、AIC, BIC の夫々が同一のモデルを選択するとは限らない。

AIC と BIC はどちらが優れているのか? この問題は一言で整理できるほど簡単な問題ではない。真のモデルを  $\text{AR}(p)$  とする。  $p$  が有限であるとすると、AIC が選択する  $\hat{p}$  について、

$$\lim_{n \rightarrow \infty} P(\hat{p} < p) = 0, \quad \lim_{n \rightarrow \infty} P(\hat{p} > p) > 0$$

となることが知られている。つまり、**標本数  $n$  を大きくすれば、AIC はモデル次数を過少に評価することはないが、過大に評価する確率が 0 とはならない**。AIC はその意味で一致性を持たない。一方で、BIC が選択する  $\hat{p}$  に関しては

$$\lim_{n \rightarrow \infty} P(\hat{p} = p) = 1$$

となることが知られており、一致性を持つことになる。したがって、BIC は標本数が十分であれば、ほぼ必ず正しいモデルを選択することができる。この点からすれば、BIC のほうが AIC よりも優れているように感じられる。

しかし、真のモデルが  $\text{AR}(\infty)$  の場合、AIC はある種の漸近最適性を持つことが知られており、また過大なモデルを選択したとしても、標本数が大きくなるにつれて過大な部分のパラメータの西邑推定量は真の値である 0 に収束するため、過大なモデルの選択はそれほど大きな問題とならない。

よって、AIC, BIC によって提示される最適なモデルが異なる場合、どちらのモデルを採用するかという問には画一的な解はなく、分析者に判断がゆだねられる。

## 7.4 モデルの予測

AR モデル、MA モデルの予測について簡単に述べる。

**定義 7.6 (最適予測)**. 2 乗平均誤差 (MSE; Mean Squared Error) が最小になるような予測を**最適予測**という。

### 7.4.1 AR モデルの予測

MSE を基準に最適予測を求めるならば、最適予測は条件付き期待値で与えられるので、「最適予測を求める」という問題は「条件付き期待値を求める」ことに帰着する。 $\varepsilon_t$  がホワイトノイズ

という仮定の下では、AR 仮定の条件付き期待値を求めるのは比較的容易で

$$\mathbf{E}[y_s|\Omega_t] = y_s, s \leq t \quad (7.19) \quad \text{eq:ARF1}$$

$$\mathbf{E}[\varepsilon_{t+k}|\Omega_t] = 0, k > 0 \quad (7.20) \quad \text{eq:ARF2}$$

という2つの結果を用いる。<sup>eq:ARF1</sup>(7.19)式の意味するところは、過去の $y$ に関しては、条件付き期待値は実際の値そのものであることを示すものであり、これは常に成り立つ。<sup>eq:ARF2</sup>(7.20)式は将来の $\varepsilon_t$ について、条件付き期待値が0であることを示すものであり、これは $\varepsilon_t$ が期待値0の独立系列であれば存在するため、「 $\varepsilon_t$ がホワイトノイズである」という仮定だけでは、<sup>eq:ARF2</sup>(7.20)式は常に一般に成り立つわけではないということに注意が必要である。

これらの式を用いて最適予測を計算するためには、 $y_{t+k}$ を過去の $y$ と将来の $\varepsilon$ に分解したうえで、過去の $y$ に関しては実績値を用い、将来の $\varepsilon_t$ については、0で置き換えればよい。例を挙げてみてみよう。

**例 7.2** (AR(1)における最適予測とその性質). AR(1) 過程

$$y_t = c + \phi_1 y_{t-1} + \varepsilon_t, \varepsilon \text{ i.i.d. } N(0, \sigma^2) \quad (7.21)$$

の最適予測とその性質を考えてみる。1期先の予測については、上式より

$$y_{t+1} = c + \phi_1 y_t + \varepsilon_{t+1}$$

と書くことが出来る。<sup>eq:ARF1</sup>(7.19)式と<sup>eq:ARF2</sup>(7.20)式を用いると、最適1期先予測は

$$\hat{y}_{t+1|t} = c + \phi_1 y_t$$

で与えられる。このときの予測誤差は

$$\hat{e}_{t+1|t} = y_{t+1} - \hat{y}_{t+1|t} = \varepsilon_{t+1}$$

であり、平均2乗誤差MSEは、

$$\text{MSE}(\hat{y}_{t+1|t}) = \mathbf{E}[\varepsilon_{t+1}^2] = \sigma^2$$

となる。2期先の予測も同様に

$$\begin{aligned} y_{t+2} &= c + \phi_1 y_{t+1} + \varepsilon_{t+2} \\ &= c + \phi_1 (c + \phi_1 y_t + \varepsilon_{t+1}) + \varepsilon_{t+2} \\ &= (1 + \phi_1)c + \phi_1^2 y_t + \phi_1 \varepsilon_{t+1} + \varepsilon_{t+2} \end{aligned}$$

と書き直すことが出来るため、2期先最適予測とMSEは

$$\begin{aligned} \hat{y}_{t+2|t} &= (1 + \phi_1)c + \phi_1^2 y_t, \\ \text{MSE}(\hat{y}_{t+2|t}) &= \mathbf{E}[(\phi_1 \varepsilon_{t+1} + \varepsilon_{t+2})^2] \\ &= (1 + \phi_1^2)\sigma^2 \end{aligned}$$

となる。これを一般の  $k$  期先予測まで展開すると

$$y_{t+k} = c \sum_{i=1}^k \phi_1^{i-1} + \phi_1^k y_t + \sum_{i=0}^{k-1} \phi_1^i \varepsilon_{t+k-i}, \quad (7.22) \quad \text{eq:0FAR1}$$

$$\hat{y}_{t+k|t} = \frac{(1 - \phi_1^k)c}{1 - \phi_1} + \phi_1^k y_t, \quad (7.23) \quad \text{eq:0FAR2}$$

$$\text{MSE}(\hat{y}_{t+k|t}) = \frac{(1 - \phi_1^{2k}\sigma^2)}{1 - \phi_1^2} \quad (7.24) \quad \text{eq:0FAR3}$$

となる。得られた最適予測をもとに性質をみてみよう。まず、第一に、最適予測が予測期間  $k$  によらず、 $y_t$  のみに依存するということである。つまり、 $t-1$  以前の  $y_{t-1}, y_{t-2}, \dots$  の情報は  $y_t$  さえあれば事足りるということである。これは、AR(1) 過程が1期前の値しか含まないことから直観的に明らかであろう。 $p > 1$  における、一般の AR( $p$ ) の場合には、後に確認するように使われる情報は増えるが、 $p$  期以上前の情報は使われない。続いて、予測期間が伸びる場合の最適予測のふるまいを考えて見よう。<sup>eq:0FAR2</sup>(7.23) 式と定常条件  $|\phi_1| < 1$  から、 $y_t$  の影響が指数的に減衰していき、

$$\lim_{k \rightarrow +\infty} \frac{(1 - \phi_1^k)c}{(1 - \phi_1)} + \phi_1^k y_t$$

を計算すれば、最適予測が過程の期待値である  $c/(1 - \phi_1)$  に収束していくことが分かる。この結果は、定常 AR(1) 過程は長期的には過程の平均に回帰することを意味しており、この性質を**平均回帰性 (mean reverting)** という。また、MSE については、予測期間が長くなるほどに単調増加していく。<sup>eq:0FAR3</sup>(7.24) 式をみれば明らかに分かるように予測期間が長くなるほどに予測が無図笠生なることを示している。 $\phi_1^{2k} \rightarrow 0 (k \rightarrow \infty)$  であるから、MSE は予測機関が長くなるにつれて、過程の分散  $\sigma^2/(1 - \phi_1^2)$  に近づいていくことがわかる。つまり、定常過程の場合には、平均回帰性があるために、MSE が発散することなく、過程の分散を上限として増大していくことが分かる。

一般に定常 AR( $p$ ) 過程の場合についてはどうであろうか？ AR( $p$ ) 過程：

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.} N(0, \sigma^2) \quad (7.25)$$

について、1期先、2期先の値  $y_{t+1}, y_{t+2}$  は

$$\begin{aligned} y_{t+1} &= c + \sum_{i=1}^p \phi_i y_{t-i+1} + \varepsilon_{t+1}, \\ y_{t+2} &= c + \phi_1 y_{t+1} + \sum_{i=1}^p \phi_i y_{t-i+2} + \varepsilon_{t+2} \\ &= (1 + \phi_1)c + \sum_{i=1}^p (\phi_1 \phi_i + \phi_{i+1}) y_{t-i+1} + \phi_1 \varepsilon_{t+1} + \varepsilon_{t+2} \end{aligned}$$

となる。このとき、 $\hat{y}$  および MSE は

$$\begin{aligned}\hat{y}_{t+1|t} &= c + \sum_{i=1}^p \phi_i y_{t-i+1}, \\ \text{MSE}(\hat{y}_{t+1|t}) &= \mathbf{E}[\varepsilon_{t+1}^2] \\ &= \sigma^2, \\ \hat{y}_{t+2|t} &= (1 + \phi_1)c + \sum_{i=1}^p (\phi_1 \phi_i + \phi_{i+1}) y_{t-i+1}, \\ \text{MSE}(\hat{y}_{t+2|t}) &= \mathbf{E}[(\phi_1 \varepsilon_{t+1} + \varepsilon_{t+2})^2] \\ &= (1 + \phi_1^2) \sigma^2\end{aligned}$$

となる。同様にして  $k$  期先の予測を求めることが出来るが、計算がより複雑さを増していく。そこで使われる手法が逐次予測と呼ばれる方法である。逐次予測は、1 期先予測の結果から初めて、予測に必要な将来の  $y$  の値を逐次予測値で置き換えていく手法で、例えば 2 期先の予測については、

$$\hat{y}_{t+2|t} = c + \phi_1 \hat{y}_{t+1|t} + \phi_2 y_t + \cdots + \phi_p y_{t-p+2}$$

と求めることが出来る。この式の中にある  $\hat{y}_{t+1|t}$  を先ほどの計算結果を代入すれば、2 期先予測が一致することが確認でき、この手法によれば AR( $p$ ) 過程の  $k$  期先予測は

$$\hat{y}_{t+k|t} = c + \sum_{i=1}^p \phi_i \hat{y}_{t+i-p|t}$$

と書くことが出来る。ここで、 $s \leq t$  であれば、 $\hat{y}_{s|t} = y_s$  である。

先ほどの例で AR(1) の場合に最適予測がどのような性質を持つかを確認したが、AR( $p$ ) についても結果だけ纏めておく。

**定理 7.15** (AR( $p$ ) の最適予測の性質). AR( $p$ ) 過程の最適予測は以下の性質をもつ：

1. 最適予測は直近の  $p$  期の  $y$  の値  $y_t, y_{t-1}, \dots, y_{t-p+1}$  のみに依存する。
2. 予測期間が長くなるにつれ、最適予測における過去の観測値の影響は指数減衰していき、最適予測は過程の期待値に収束する。
3. 予測期間が長くなるにつれて、平均 2 乗誤差 MSE は単調に増大し、過程の分散に収束する。

#### 7.4.2 MA モデルの予測

続いて、反転可能な MA 過程の予測について考える。予測する過程が MA 過程であっても、最適予測が条件付き期待値になることには変わらないが、MA 過程は観測することが出来ない  $\varepsilon$  の線形和で  $y$  が表現されるため、予測はやや複雑となる。

無限個の  $y$  の観測値が利用可能であれば、MA 過程の AR( $\infty$ ) 表現より、 $\varepsilon$  の値を求めることが出来る。したがって、この場合は、過去の  $\varepsilon$  の値が観測できるとして、予測を考えることが出来るため、この場合から始めていく。その後、有限個の  $y$  の観測値しか利用できない場合にどのように予測を構築するかを考えていく。

反転可能な MA 過程は一般的に、

$$y_t = \sum_{i=1}^{\infty} \eta_i y_{t-i} + \varepsilon_t$$

と  $\text{AR}(\infty)$  表現を持つ。もし、無限個の  $y$  の過去の観測値があれば、 $\Omega_t := \{y_t, y_{t-1}, \dots\}$  とすると、これによって

$$\varepsilon_t = y_t - \sum_{i=1}^{\infty} \eta_i y_{t-i} \quad (7.26) \quad \text{eq:MAeps}$$

として、過去の  $\varepsilon_t$  の値を求めることが出来る。条件付き期待値は

$$\mathbf{E}[y_s | \Omega_t] = y_s, s \leq t, \quad (7.27) \quad \text{eq:MAF1}$$

$$\mathbf{E}[\varepsilon_{t+k} | \Omega_t] = 0, k > 0, \quad (7.28) \quad \text{eq:MAF2}$$

$$\mathbf{E}[\varepsilon_s | \Omega_t] = \varepsilon_s, s \leq t \quad (7.29) \quad \text{eq:MAF3}$$

となり、AR 過程の時に加えて、<sup>eq:MAF3</sup>(7.29) 式を用いることが出来る。

AR 過程の予測の場合には、 $y_{t+k}$  が将来の  $y$  を含んでいるため、 $y_{t+k}$  を過去の  $y$  と将来の  $\varepsilon$  に書き直していた。MA 過程の場合には、 $y_{t+k}$  は過去から将来の  $\varepsilon$  しか含んでおらず、それらの条件付き期待値は、<sup>eq:MAF2</sup>(7.28) 式と <sup>eq:MAF3</sup>(7.29) 式により完全に求めることが出来る。MA 過程の最適予測を <sup>eq:MAF2</sup>(7.28) 式と <sup>eq:MAF3</sup>(7.29) 式を用いて求めていこう。

**例 7.3** (MA(2) 過程の最適予測とその性質). MA(2) 過程

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \varepsilon \sim \text{i.i.d. } N(0, \sigma^2) \quad (7.30)$$

の最適予測とその性質を考えて見る。1 期先の予測については、上式より

$$y_{t+1} = \mu + \varepsilon_{t+1} + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1}$$

と書くことが出来る。そこで、<sup>eq:MAF2</sup>(7.28) 式と <sup>eq:MAF3</sup>(7.29) 式を用いると、最適 1 期先予測と MSE は

$$\hat{y}_{t+1|t} = \mu + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1}, \quad (7.31) \quad \text{eq:MA2_1f}$$

$$\text{MSE}(\hat{y}_{t+1|t}) = \mathbf{E}[\hat{\varepsilon}_{t+1|t}^2] = \mathbf{E}[\varepsilon_{t+1}^2] = \sigma^2 \quad (7.32) \quad \text{eq:MA2_1fMSE}$$

となる。 $\varepsilon_t, \varepsilon_{t-1}$  は <sup>eq:MAeps</sup>(7.26) 式によって求めているため、<sup>eq:MA2\_1f</sup>(7.31) 式は過去すべての観測値に依存していることに注意しておく。この点が、AR 過程の予測とは本質的に異なる箇所、MA 過程の予測はすべての観測値に依存する。2 期先の予測も同様にすれば

$$y_{t+2} = \mu + \varepsilon_{t+2} + \theta_1 \varepsilon_{t+1} + \theta_2 \varepsilon_t, \quad (7.33)$$

$$\hat{y}_{t+2|t} = \mu + \theta_2 \varepsilon_t, \quad (7.34)$$

$$\text{MSE}(\hat{y}_{t+2|t}) = (1 + \theta_1^2) \sigma^2 \quad (7.35)$$

を得る。3 期先の予測では、

$$y_{t+3} = \mu + \varepsilon_{t+3} + \theta_1 \varepsilon_{t+2} + \theta_2 \varepsilon_{t+1}$$



となり、ついに将来の  $\varepsilon$  の値しか使わなくなる。最適予測においては、 $\varepsilon_{t+i}$  ( $i = 1, 2, 3, \dots$ ) はすべて 0 で置き換えることが出来るので、 $\hat{y}_{t+3|t} = \mu$  となり、単に過程の期待値に等しくなる。また、MSE についても

$$\text{MSE}(\hat{y}_{t+3|t}) = \mathbf{E}[(\varepsilon_{t+3} + \theta_1\varepsilon_{t+2} + \theta_2\varepsilon_{t+1})^2] = (1 + \theta_1^2 + \theta_2^2)\sigma^2 = \gamma_0$$

となるため、MSE は過程の分散に等しくなる。4 期先以降は、3 期先と同じ結果を得ることは最早自明であろう。

例で確かめたことを、一般の  $\text{MA}(q)$  に拡張して、性質を纏めておく。

**定理 7.16** ( $\text{MA}(q)$  過程の最適予測の性質).  $\text{MA}(q)$  過程の最適予測は以下の性質をもつ：

1.  $q$  期までの最適予測はすべての観測値  $y_t, y_{t-1}, y_{t-2}, \dots$  に依存する
2.  $q + 1$  期先以上の予測は単に過程の期待値に等しい。
3.  $q$  期までの予測の MSE は予測期間  $k$  が伸びるにつれて単調増加。
4.  $q + 1$  期先以上の予測において MSE は過程の分散に等しい。

無限個の  $y$  の観測値があるという過程は、MA 過程の予測にあたり非常に便利な過程である一方で、実際には現実的ではない。そこで、有限個の  $y$  の観測値しかない  $\Omega_t = \{y_t, y_{t-1}, \dots, y_1\}$  の場合に予測をどのように行うかを議論する。先の  $\text{MA}(2)$  過程を例にとろう。まず、明らかに 3 期先以降の予測に関しては、無限個の観測値がある場合と同じであることは分かるだろう。よって、1 期先・2 期先の予測に関して考える。 $\varepsilon_t, \varepsilon_{t-1}$  の計算が必要となるが、ここで  $\varepsilon$  の初期値を 0 として、 $\varepsilon_1$  から逐次的に  $\varepsilon$  の近似値を計算していく方法を考える。具体的には、モデル式より

$$\varepsilon_t = y_t - \mu - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}$$

という関係が導かれるため、 $\varepsilon_1$  の近似から順番に

$$\begin{aligned}\hat{\varepsilon}_1 &= y_1 - \mu, \\ \hat{\varepsilon}_2 &= y_2 - \mu - \theta_1\hat{\varepsilon}_1, \\ \hat{\varepsilon}_3 &= y_3 - \mu - \theta_1\hat{\varepsilon}_2 - \theta_2\hat{\varepsilon}_1, \\ \hat{\varepsilon}_4 &= y_4 - \mu - \theta_1\hat{\varepsilon}_3 - \theta_2\hat{\varepsilon}_2 \\ &\vdots\end{aligned}$$

という具合に、近似値を求めていき、この  $\hat{\varepsilon}$  を用いて、無限個の観測値がある場合同様に予測を行う。問題はこの手法に妥当性があるのか？という点である。標本期間以前の  $\varepsilon$  を 0 とすることは、標本期間以前の  $y$  を過程の期待値  $\mu$  と仮定することと同じであり、ある程度正当化できる。 $\varepsilon$  は (7.26) 式によって計算することができるが、 $\eta_k$  は指数的に減衰する、つまり、過去の  $y$  の影響は急激に小さくなる。そのため、標本数が大きい場合には、標本期間以前の  $y$  を  $\mu$  と仮定しても大きな問題とはならない。一方で、この方法で求めた  $\varepsilon$  はそもそも近似値であり、最適予測やその MSE もまた近似となっていることには注意が必要である。

### 7.4.3 区間予測

これまでの、 $k$  期先の  $y_{t+k}$  の値を 1 つ予測する所謂「点予測」を行ってきた。これに対して、 $y_{t+k}$  をある確率で含むような区間を予測することを**区間予測**という。

$k$  期先の区間予測を構築するには、 $\Omega_t$  を与えたときの  $y_{t+k}$  の条件付き分布を求める必要がある。

#### AR 過程の区間予測

定常 AR( $p$ ) 過程

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2) \quad (7.36)$$

に対して条件付き分布と区間予測を考える。このとき、 $y_{t+1}$  は

$$y_{t+1} = c + \phi_1 y_t + \phi_2 y_{t-1} + \cdots + \phi_p y_{t-p+1} + \varepsilon_{t+1} \quad (7.37)$$

と書くことが出来るので、 $\Omega_t$  を所与とすると、確率的な変数は  $\varepsilon_{t+1}$  のみである。 $\varepsilon_t$  は正規分布に従うと仮定されているので、 $\Omega_t$  を所与としたときの  $y_{t+1}$  の条件付き分布もまた、正規分布となる。正規分布は期待値と分散により分布が完全に特定できるが、最適予測は条件付き期待値に等しかったことを思い出そう。また、条件付き分散は  $\text{Var}[\varepsilon_{t+1}] = \sigma^2$  であるが、これは MSE に等しい。つまり、 $y_{t+1}$  の条件付き分布は  $N(\hat{y}_{t+1|t}, \text{MSE}(\hat{y}_{t+1|t}))$  となる。これにより、 $y$  について 95% 区間予測は

$$(\hat{y}_{t+1|t} - 1.96\sqrt{\text{MSE}(\hat{y}_{t+1|t})}, \hat{y}_{t+1|t} + 1.96\sqrt{\text{MSE}(\hat{y}_{t+1|t})})$$

となることがわかる。これを踏まえれば 2 期以降先の  $k$  期先区間予測についても

$$(\hat{y}_{t+k|t} - 1.96\sqrt{\text{MSE}(\hat{y}_{t+k|t})}, \hat{y}_{t+k|t} + 1.96\sqrt{\text{MSE}(\hat{y}_{t+k|t})}) \quad (7.38) \quad \text{eq: int\_for}$$

で与えられる。しかし、AR( $p$ ) 過程の  $k$  期先予測の MSE を求めることは難しいため、MSE の値を使ったこの式により区間予測を求めることは難しい。そこで、Kalman Filter などを使って最適予測を行う手法などが提案されたりする。

#### MA 過程の区間予測

AR 過程と同様の議論が成立するので、<sup>eq:MA2\_1f</sup> (7.31) 式と <sup>eq:MA2\_1f</sup> (7.32) 式を <sup>eq: int\\_for</sup> (7.38) へ代入すれば求めることが出来る。

### 7.4.4 AR モデルの予測の一例

以下、AR モデルの予測例を扱いこれまでの内容を振り返るとともに、補足的な情報を与える。

マクロ経済活動を示す四半期別の GDP<sup>5</sup>は重要な経済変数の観測値である。そして生産要素を平均的に投入した場合に産出できる供給量を表す潜在 GDP と実際の GDP の乖離である GDP

<sup>5</sup>GDP：国内総生産。一定期間内に国内で新たに生み出された物やサービスの付加価値のこと。Gross Domestic Product の略。



図 7.1: 内閣府より取得した GDP ギャップ (1980 年から 2022 年)

GDPGAP\_fig1

ギャップが公表<sup>6</sup>されている。GDP ギャップは景気の現状判断のための重要な指標として知られており、これを AR( $p$ ) モデルを使って表現しようとする。

まずは、データの整理から始めよう。次に示すグラフ (図 7.1) は、GDP ギャップの四半期ごとのデータを 1980 年第 1 四半期より 2022 年第 4 四半期までプロットしたものである。これをみると、バブル景気 (1980 年代後半から 1990 年代前半) や、所謂「いざなぎ景気」の後期に GDP ギャップの値が正值、2008 年の世界金融危機や、コロナ禍には大幅に負値となっていることが見て取れる。

まず、手始めに 1980 年第 2 四半期から 2022 年第 4 四半期までの GDP ギャップの系列を用いて、AR(1) モデルを推定すると

$$\hat{y}_t = -0.1741 + 0.8107y_{t-1} \quad (7.39)$$

という結果となった。自己回帰係数は 0.8107 と正值となっており、今期と来期の変数の間に正の関係があることを意味している。

続いて、 $k$  次の自己相関を計算してみよう。 $k$  次の標本自己相関は

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} \quad (7.40)$$

と計算されるのであった。

また、標本自己相関に基づいて、「系列相関がない」という帰無仮説を検定してみよう。次数  $k$  の真の自己相関  $\rho_k = 0$  であるかどうかを調べたいとする。このとき、帰無仮説  $\rho_k = 0$ のもとで、標準誤差

$$SE(\hat{\rho}_k) = \sqrt{\frac{1}{T} \left( 1 + \sum_{s=1}^{k-1} \hat{\rho}_s^2 \right)} \quad (7.41)$$

<sup>6</sup>以下の分析に用いたデータは <https://www5.cao.go.jp/keizai3/getsurei/getsurei-index.html> より取得したもの。

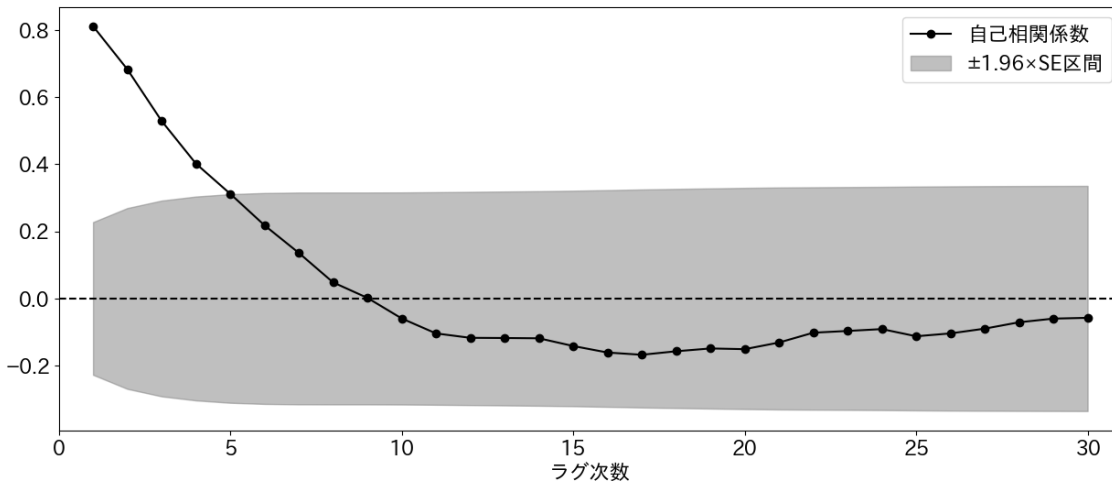


図 7.2: GDP ギャップのコレログラム

GDPGAP\_fig2

として計算された、 $\hat{\rho}_k$  の  $t$  検定統計量  $\hat{\rho}_k / \text{SE}(\hat{\rho}_k)$  は、漸近的に標準正規分布に従う<sup>7</sup>。有意水準 5% の場合、 $t$  検定統計量の絶対値が標準正規分布の臨海値である 1.96 を越えていれば帰無仮説  $\rho_k$  を棄却する。

ここで、系列相関がないためには、特定の次数だけではなく、すべての次数に関して自己相関が 0 であることが必要となる点に注意が必要である。このため、他の次数についても検討しなければならない。有限個の観測値からすべての次数の標本自己相関を計算することは出来ない。そのため、分析者は、選択した最大の次数  $M$  までの標本自己相関について次のような**逐次検定**を行う。まずはじめに、 $k = 1$  として、1 次の標本自己相関の  $t$  検定統計量の絶対値が、有意水準 5% の臨海値 1.96 を越えていれば、帰無仮説  $\rho_1 = 0$  を棄却し、系列相関があると判断する。もし、臨海値を下回っており、帰無仮説が棄却できなければ続いて  $k + 1$  次の標本自己相関について同様の検定を行う。この手続きを  $k = M$  となるまで繰り返し、すべての検定で棄却できなければ、系列相関がないと判断する。

逐次検定を視覚化するため、次数に対する自己相関の推移に加えて、区間  $\pm 1.96 \times \text{SE}(\hat{\rho}_k)$  をグレーでハイライトしたものが図 7.2 である。グラフの見方を説明すると、いずれかの次数  $k$  において標本自己相関がそのハイライトされた区間の外側にあれば、系列相関があると判断することが出来る。具体的には、1 次から 4 次間は有意に正の値となっている。また、5 次以上のラグ次数については、自己相関が 0 という帰無仮説を棄却できないということがわかる。

以上が逐次検定やコレログラムの区間によって系列相関の有無を判断する方法の概略であるが、標本自己相関の数があまりに大きくなると**多重検定**の問題が生じ、逐次検定は好ましくない。これは、 $t$  検定を複数回行うと、設定された有意水準よりも高い頻度で第 1 種の誤りが生じてしまうという問題で、実際に系列相関がないにもかかわらず、どこかの次数で帰無仮説を棄却してしまう確率が系列相関の検定の繰り返し回数とともに増加してしまう。この問題を回避するためには、「複数の自己相関が 0 である」という**結合帰無仮説**

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_M = 0$$

を一度に検定すればよい。このとき、対立仮説は「少なくとも 1 つの自己相関は 0 でない」であ

<sup>7</sup> 詳しい理論的な説明は統計学の教科書を参照されたい。

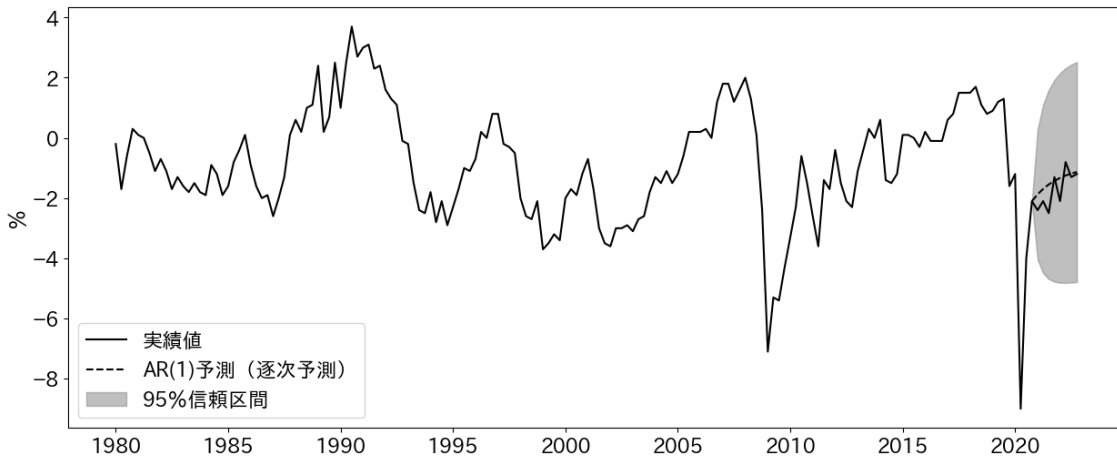


図 7.3: GDP ギャップの AR(1) による逐次予測と実績値 (その 1)

GDPGAP\_fig3

るから、1 から  $M$  の範囲で  $\rho_k \neq 0$  となる次数  $k$  が存在することを確認することになる。次で定義される  $Q$  検定統計量 ( $Q$ ) および修正  $Q$  検定統計量 ( $Q^*$ ) は、帰無仮説  $\rho_1 = \rho_2 = \dots = \rho_M = 0$  のもとで、自由度  $M$  の  $\chi^2$  分布に従う：

$$Q := T \sum_{k=1}^M \hat{\rho}_k^2, \quad (7.42)$$

$$Q^* := T(T+2) \sum_{k=1}^M \frac{\hat{\rho}_k^2}{T-k}. \quad (7.43)$$

これらの検定は複数の自己相関の仮説を一纏めにして行うことから、**かばん検定**などと呼ばれている。

有限標本でより良い性質を持つとされる修正  $Q$  検定を用いて、GDP ギャップの系列相関が統計的に検出されるかを確認したものが次の表であり、 $Q^*$  の値が臨海値を越えていれば帰無仮説は棄却される。 $M = 10$  の場合、 $Q^* = 319.7$  であり、5% 臨海値 15.51 よりも大きいことから、「系列相関がない」という帰無仮説は棄却されるということになる。

$M$	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_M$	0.811	0.682	0.528	0.401	0.311	0.218	0.136	0.048	0.002	-0.061
$Q^*$	115.0	197.5	248.0	278.2	297.2	307.5	312.7	315.1	317.0	319.7
臨海値	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92	18.31

図7.1を見る限り、1 次の自己相関が高いため、AR(1) モデルの予測を試みる。パラメータが道であるため、推定値を置き換えて計算していく。2020 年度までのデータをもとに、AR(1) を予測し、2022 年度末までの逐次予測とその 95%信頼区間の結果と、実績値をプロットしたものが次の図7.3である。2019 年度までのデータをもとに、同様に逐次予測した結果をプロットすると図7.4のようになり、コロナ禍を受けて経済活動が落ち込んだ時期の GDP ギャップは AR(1) モデルによる予測の 95%信頼区間の外にあり、予測困難な事象であったことが確認できる。

続けて、図7.4と同じ前提で直接予測を行った結果が図7.5である。

ここまでは AR(1) モデルに限って予測を行ってきた。それでは、AR( $p$ ) モデルの予測ではどうだろうか。AR(1) の場合と異なり、 $p$  をどのように設定するのかという課題が存在する点が大きく異なるため、AIC, BIC を用いた分析にもふれる。

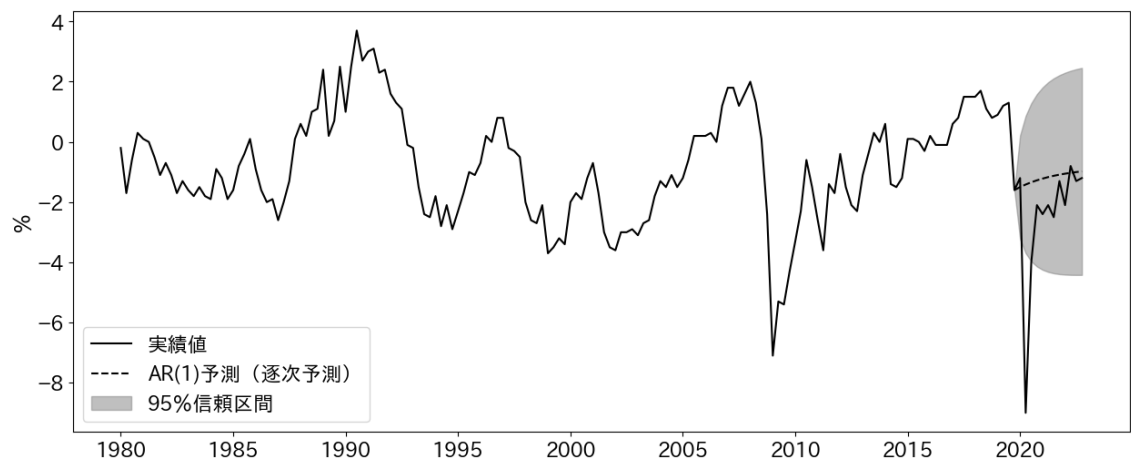


図 7.4: GDP ギャップの AR(1) による逐次予測と実績値（その 2）

GDPGAP\_fig4

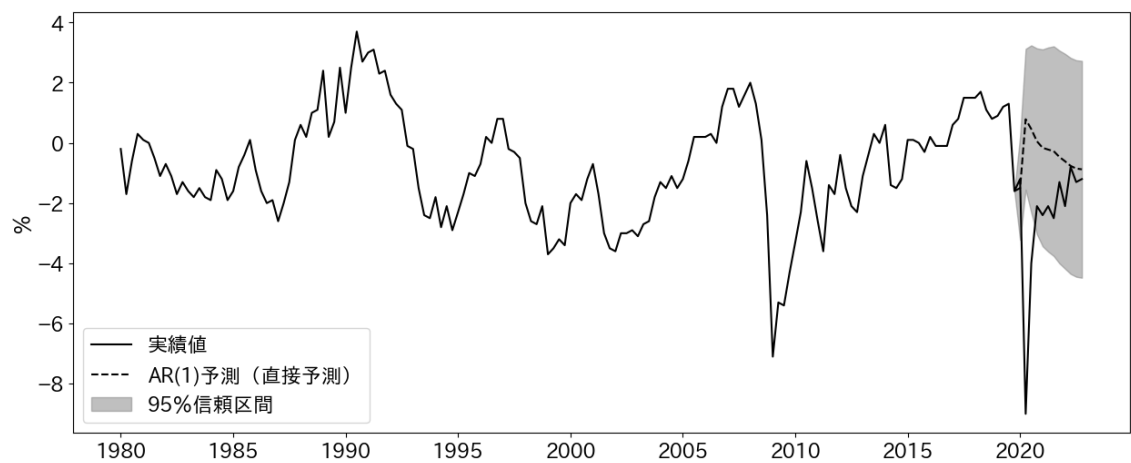


図 7.5: GDP ギャップの AR(1) による直接予測と実績値

GDPGAP\_fig5

$p$  次の自己回帰モデル  $AR(p)$  を考えよう。このとき、 $y_t$  は

$$y_t = \phi_0 + \sum_{i=1}^p \phi_i y_{t-i} + \varepsilon_t \quad (7.44)$$

と表現される。 $AR(p)$  の推定は  $AR(1)$  同様に OLS で推定される。推定に用いるサンプルサイズを  $T$  とするため、データの観測期間を  $t = -p+1, -p, \dots, T$  と設定する。このとき、OLS 推定量は、残差平方和は

$$RSE(p) = \sum_{t=1}^T \varepsilon_p^2 = \sum_{t=1}^T \left[ y_t - \left( \hat{\phi}_0 + \hat{\phi}_1 y_{t-1} + \hat{\phi}_2 y_{t-2} + \dots + \hat{\phi}_p y_{t-p} \right) \right]^2$$

を最小化する  $(\hat{\phi}_0, \hat{\phi}_1, \dots, \hat{\phi}_p)$  である。

$AR(2)$  モデル

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

の場合を考えよう。パラメータ  $\{\phi_0, \phi_1, \phi_2, \sigma\}$  を既知とする。

1 期先予測は、現在と過去の観測値  $\{y_t\}_{t=-1}^T$  のもとでの  $y_{T+1}$  の条件付き期待値から、最適予測と予測誤差はそれぞれ

$$\begin{aligned} y_{T+1|T} &= \mathbf{E}[y_{T+1} | \{y_t\}_{t=-1}^T] = \phi_0 + \phi_1 y_T + \phi_2 y_{T-1}, \\ y_{T+1} - y_{T+1|T} &= (\phi_0 + \phi_1 y_T + \phi_2 y_{T-1} + \varepsilon_{T+1}) - (\phi_0 + \phi_1 y_T + \phi_2 y_{T-1}) = \varepsilon_{T+1} \end{aligned}$$

となり、

$$MSFE = \sigma_y^2(1) = \mathbf{E}[(y_{T+1} - y_{T+1|T})^2] = \sigma^2$$

となることが確認できる。2 期先予測についても最適予測・予測誤差・MSFE はそれぞれ

$$\begin{aligned} y_{T+2|T} &= \phi_0(1 + \phi_1) + (\phi_1^2 + \phi_2)y_T + \phi_1\phi_2 y_{T-1}, \\ y_{T+2} - y_{T+2|T} &= \phi_1 \varepsilon_{T+1} + \varepsilon_{T+2}, \\ \sigma_y^2(2) &= (1 + \phi_1)^2 \sigma^2 \end{aligned}$$

となる。また、3 期先予測についても計算結果だけを述べれば

$$\begin{aligned} y_{T+3|T} &= \phi_0(1 + \phi_1(1 + \phi_1) + \phi_2) + (\phi_1^3 + 2\phi_1\phi_2)y_T + (\phi_1^2\phi_2 + \phi_2^2)y_{T-1}, \\ y_{T+3} - y_{T+3|T} &= (\phi_1^2 + \phi_2)\varepsilon_{T+1} + \phi_1\varepsilon_{T+2} + \varepsilon_{T+3}, \\ \sigma_y^2(3) &= (1 + \phi_1^2 + (\phi_1^2 + \phi_2)^2)\sigma^2 \end{aligned}$$

となる。**Q** 2 期先予測・3 期先予測について、最適予測・予測誤差・MSFE がそれぞれ上述の通りになることを確かめよ。この結果で注意すべきは、予測機関に関わらず、予測値は常に  $y_T, y_{T-1}$  のみに依存していることである。

パラメータ  $\{\phi_0, \phi_1, \phi_2, \sigma\}$  を未知である場合を考えてる。この場合、それぞれのパラメータに  $AR(2)$  の OLS 推定量  $\{\hat{\phi}_0, \hat{\phi}_1, \hat{\phi}_2, \hat{\sigma}\}$  を代入した形で得られる。予測の 95%信頼区間は  $\hat{y}_{T+i|T} \pm 1.96 \cdot \hat{\sigma}_y(i)$  となる。

$AR(p)$  モデルの推定に当たって、ラグ次数  $p$  の選択が必要となる。例えば、自己回帰係数推定値の標準誤差から、次のような逐次検定を用いたラグ次数の選択が出来る。

ラグ次数の最大値の候補として  $p = p_{\max}$  を設定し、 $\text{AR}(p_{\max})$  モデルの  $\phi_{p_{\max}}$  の推定値が有意であればそのモデルを選択し、有意でなければ  $\text{AR}(p_{\max} - 1)$  モデルを推定し  $\phi_{p_{\max}-1}$  の推定値が有意であるかを確認する。この作業を繰り返していく中で、最初に有意となった次数を選択するという具合である。ただし、この逐次検定によるラグ次数選択をした場合、ラグ次数の最大値の候補が大きいほど、多重検定の問題が生じる。

そこで、別のアプローチとして、情報量規準を用いてラグ次数の選択を試みよう。AIC, BIC は

$$\text{AIC}(p) := \ln \left( \frac{\text{RSS}(p)}{T} \right) + \frac{2(p+1)}{T}, \quad (7.45)$$

$$\text{BIC}(p) := \ln \left( \frac{\text{RSS}(p)}{T} \right) + \frac{(p+1) \ln T}{T} \quad (7.46)$$

で定義されるものであった。ここで、 $\text{RSS}(p)$  は、 $\text{AR}(p)$  モデルの残差平方和である。すべてのラグ次数の候補について推定された AR モデルから AIC, BIC を計算し、最小の AIC に対応したラグ次数を選択する。

これまで、最小化する残差平方和に含まれる残差の数をちょうど  $T$  とするために、 $\text{AR}(p)$  モデルの OLS 推定の説明では、データの観測期間を  $t = -p+1$  から  $T$  としていた。しかし、推定する AR モデルのラグ次数が変化しても実際に観測される標本  $\{y_1, y_2, \dots, y_T\}$  のサンプルサイズ  $T$  は不変であるため、情報量規準の計算は注意が必要である。それぞれの AR モデルで利用できる観測値をすべて利用した場合、OLS 推定で最小化される残差平方和の残差の数が増加してしまうからである。このため、ここでは、データの観測値を  $t = 1$  から  $T$  とした場合に実際に計算できる情報量規準について少し議論しておく。

すべてのモデルのサンプルサイズを固定するには、最大のラグ次数  $p_{\max}$  で推定された場合と同じ数の残差に関する残差平方和

$$\text{RSS}(p) = \sum_{t=1+p_{\max}}^T \hat{\varepsilon}_{p,t}^2 = \sum_{t=1+p_{\max}}^T \left[ y_t - \left( \hat{\phi}_0 + \hat{\phi}_1 y_{t+1} + \dots + \hat{\phi}_p y_{t-p} \right) \right]^2$$

を最小化すればよい。この場合には、モデルの推定に用いるサンプルサイズは常に  $T - p_{\max}$  となり、AIC と BIC はそれぞれ

$$\text{AIC}_s(p) = \ln \left( \frac{\sum_{t=1+p_{\max}}^T \hat{\varepsilon}_{p,t}^2}{T - p_{\max}} \right) + \frac{2(p+1)}{T - p_{\max}}, \quad (7.47)$$

$$\text{BIC}_s(p) = \ln \left( \frac{\sum_{t=1+p_{\max}}^T \hat{\varepsilon}_{p,t}^2}{T - p_{\max}} \right) + \frac{(p+1) \ln(T - p_{\max})}{T - p_{\max}} \quad (7.48)$$

と計算される。このように、推定するすべての AR モデルのサンプルサイズが同じになるように調整することで、情報量規準の第 1 項は  $p$  に関する減少関数となることが保証される。

誤差項の標本分散の自由度調整を行う場合には

$$\text{AIC}_f(p) = \ln \left( \frac{\sum_{t=1+p_{\max}}^T \hat{\varepsilon}_{p,t}^2}{T - p_{\max} - p - 1} \right) + \frac{2(p+1)}{T - p_{\max}}, \quad (7.49)$$

$$\text{BIC}_f(p) = \ln \left( \frac{\sum_{t=1+p_{\max}}^T \hat{\varepsilon}_{p,t}^2}{T - p_{\max} - p - 1} \right) + \frac{(p+1) \ln(T - p_{\max})}{T - p_{\max}} \quad (7.50)$$



を用いればよく、それぞれの AR モデルで利用できる観測値をすべて利用した場合は

$$\text{RSS}(p) = \sum_{t=1+p}^T \hat{\varepsilon}_{p,t}^2 = \sum_{t=1+p_{\max}}^T \left[ y_t - \left( \hat{\phi}_0 + \hat{\phi}_1 y_{t+1} + \cdots + \hat{\phi}_p y_{t-p} \right) \right]^2$$

を最小化することになり、AIC と BIC は

$$\text{AIC}_a(p) = \ln \left( \frac{\sum_{t=1+p}^T \hat{\varepsilon}_{p,t}^2}{T-p} \right) + \frac{2(p+1)}{T-p}, \quad (7.51)$$

$$\text{BIC}_a(p) = \ln \left( \frac{\sum_{t=1+p}^T \hat{\varepsilon}_{p,t}^2}{T-p} \right) + \frac{(p+1) \ln(T-p)}{T-p} \quad (7.52)$$

として計算される。

サンプルサイズが十分に大きい場合、理論的にどの計算方法を用いても同じ結果が得られるはずであるが、そうでない場合には計算方法に依存して異なるラグ次数が選択される可能性がある。

**注意 15.** AIC と BIC でラグ次数の選択結果が異なる場合もある。BIC は一貫性を持ち、サンプルサイズが十分に大きいときに正しいモデルを必ず選ぶという点で理論的に好ましい。一方で、サンプルサイズが小さい場合には、AIC による選択結果のほうが好ましい場合もある。このため、サンプルサイズがある程度大きいときには BIC、小さいときには AIC の結果を用いるとすることが好ましいかもしれない。

直接予測のラグ次数の選択においても情報量規準を用いることが可能である。AR( $p$ ) モデルを用いた  $k$  期先予測では、

$$y_{t+k} = c_0(k) + c_1(k)y_t + \cdots + c_p(k)y_{t-p+1} + u_{t+k}$$

を推定することになる。直接予測の場合の情報量規準の計算には、AR( $p$ ) モデルの残差平方和ではなく、直接予測で用いる回帰式の残差平方和

$$\text{RSS}(p) = \sum_{t=1+p}^T \hat{u}_{p,t+k}^2 = \sum_{t=1+p}^T [y_{t+k} - (\hat{c}_0(k) + \hat{c}_1(k)y_t + \cdots + \hat{c}_p(k)y_{t-p+1})]^2$$

に置き換えて、AIC・BIC の式へ代入すればよい。ここで、 $\hat{c}_i (i = 0, 1, \dots, p)$  は直接予測で用いる回帰式の OLS 推定量である。

GDP ギャップの AR モデルのラグ次数を AIC と BIC を用いて選択してみる。最大のラグ次数を  $p_{\max} = 5$  として、AIC と BIC を夫々計算した結果は次のとおり：

$p$	0	1	2	3	4	5
AIC <sub>s</sub>	1.276	<b>0.203</b>	0.211	0.205	0.215	0.226
BIC <sub>s</sub>	1.295	<b>0.240</b>	0.267	0.279	0.309	0.338

最小の情報量規準を太字で表示した。尚、AIC<sub>f</sub>, AIC<sub>a</sub>, BIC<sub>f</sub>, BIC<sub>a</sub> においても同様に  $p = 1$  が選択されるという傾向がみられる。

AR モデルが正しいラグ次数で推定されている場合、誤差項には系列相関がないはずである。そこで、推定された AR モデルの残差系列に対して、 $Q$  検定や修正  $Q$  検定のかばん検定を応用すれば、誤差項に「系列相関がない」という帰無仮説を検定することが出来る。この様に推定さ

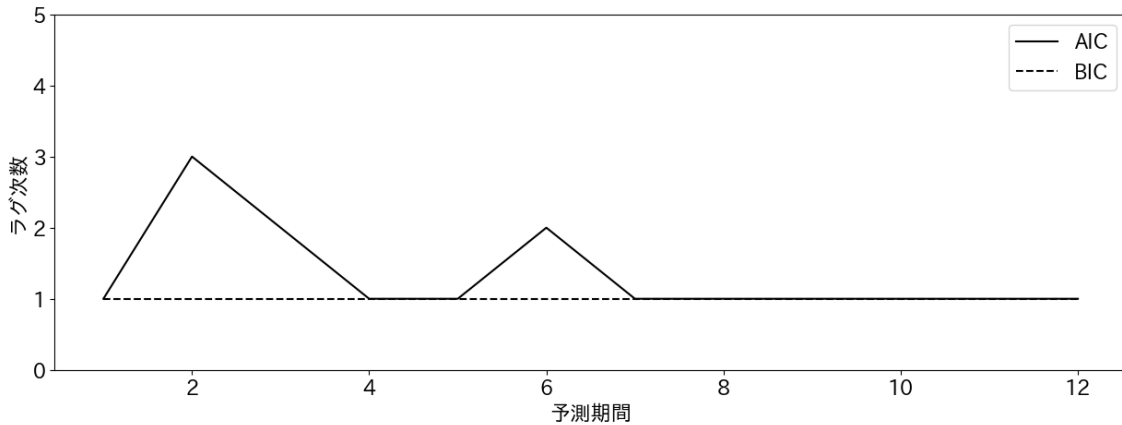


図 7.6: AIC, BIC により選択されたラグ次数

GDPGAP\_fig8

れたモデルを事後的に診断することは、AR モデルが正しく推定されているかを評価するうえで有益であるが、推定された AR モデルの残差系列にかばん検定を用いる場合には、臨海値の選択で自由度調整が必要となる。原系列の検定の場合、帰無仮説  $\rho_1 = \rho_2 = \dots = \rho_M = 0$  のもとで自由度  $M$  の  $\chi^2$  分布に従ったが、AR( $p$ ) モデルの残差系列では自由度  $M - p$  の  $\chi^2$  分布に従う。

修正  $Q$  検定を用いて、GDP ギャップの AR(1) モデルの残差に系列相関がないかを確認する。

$M$	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_M$	-0.053	0.127	-0.011	-0.047	0.040	-0.004	0.029	-0.096	0.037	-0.052
$Q^*$	0.50	3.34	3.38	3.80	4.11	4.14	4.32	6.01	6.30	6.84
臨海値	NaN	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92

$Q^*$  の値はいずれも臨海値を越えていない。 $M = 10$  における  $Q^*$  は 6.85 となっており、有意水準 5% の自由度 9 の  $\chi^2$  分布の臨海値 16.92 より小さいため、「系列相関がない」という帰無仮説を棄却できないことがわかる。

AR(1) 予測の時と同じく、1980 年第 1 四半期から 2019 年第 4 四半期までの観測値を用いて直接予測の AR( $p$ ) モデルを推定し、1 期先である 2020 年第 1 四半期から 12 期先の 2022 年第 4 四半期までの複数気先予測を行う。ラグ次数の選択は AIC・BIC 両方で行い、いずれも AIC<sub>s</sub>, BIC<sub>s</sub> により計算する。これらで計算された AIC・BIC が最小となるようなラグ次数の AR モデルを各予測期間  $k = 1, 2, \dots, 12$  について個別に選択して、予測値とその信頼区間を計算する。

AIC と BIC により選択されるラグ次数は図 7.6 に示された通りで、 $k = 1$  期先予測では、AIC・BIC いずれも AR(1) モデルが選択されている。2 期先予測では AR(3) モデル、3 期先と 6 期先では AR(2) モデルが選択されている。一方で、BIC を用いると、2 期先以降においても AR(1) モデルが選択されている。

AIC を用いた場合の AR( $p$ ) 予測、BIC を用いた場合の AR( $p$ ) 予測はそれぞれ図 7.7、図 7.8 の通りである。どちらの場合においても、2020 年における GDP ギャップの大幅な下落を捉えられていない（95%信頼区間外にある）ことが確認できる。

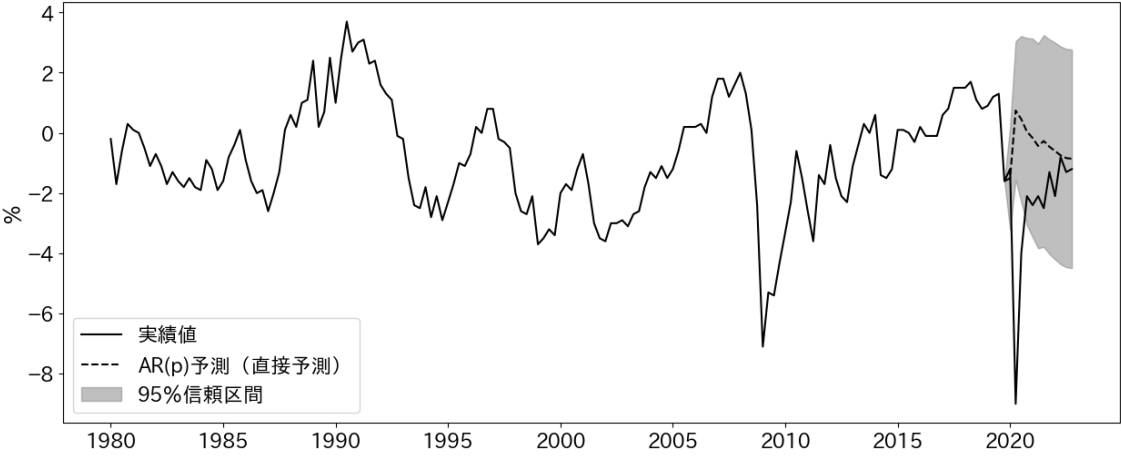


図 7.7: GDP ギャップの AR(p) 予測 (AIC によるラグ次数選択)

GDPGAP\_fig6

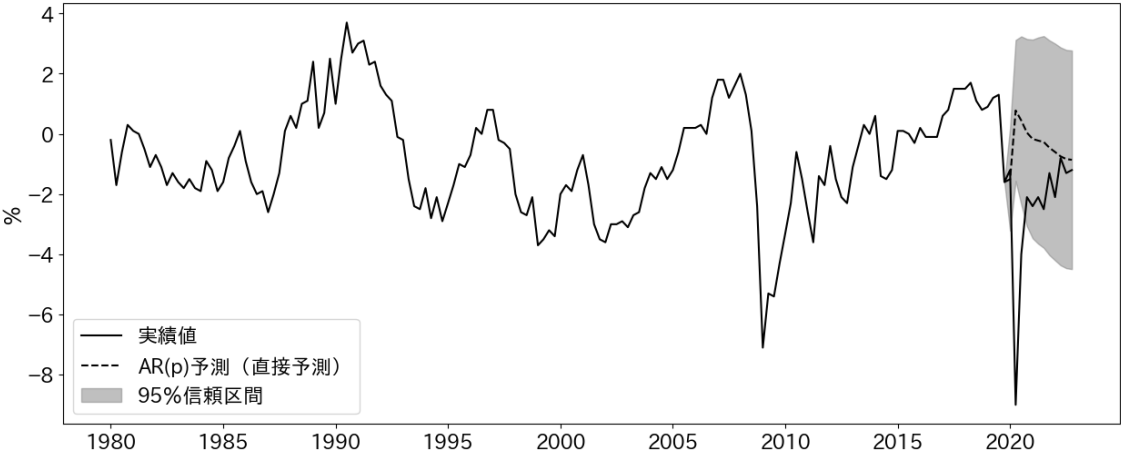


図 7.8: GDP ギャップの AR(p) 予測 (BIC によるラグ次数選択)

GDPGAP\_fig7

## 7.5 VAR モデル

ec:VAR

ベクトル自己回帰モデル (VAR; Vector autoregressive) は, AR モデルを多次元に拡張したものである. 従来の AR モデルでの分析に加えて, 変数間の関係の分析を行うことができるようになる.

例えば, 日本とアメリカの2国の株式指数の関係を知りたい場合に, AR モデルでは夫々を個別にモデル化してきたが,  $2 \times 1$  ベクトル  $\mathbf{y}_t(y_{1,t}, y_{2,t})^\top$  を定めることから始める. ベクトルでこれらを扱うと, 従来使ってきた自己相関や自己共分散, 定常性などの概念を拡張していく必要がある. そこで, まずは, これらの概念の拡張を紹介したうえで, VAR モデルの性質を見ていこう.

ベクトル  $\mathbf{y}_t = (y_{1,t}, y_{2,t}, \dots, y_{n,t})^\top$  に対して, その期待値は

$$\mathbf{E}[\mathbf{y}_t] = (\mathbf{E}[y_{1,t}], \mathbf{E}[y_{2,t}], \dots, \mathbf{E}[y_{n,t}])^\top$$

で定義される.  $k$  次の自己共分散を拡張する場合は,  $n$  個の成分に対し, それぞれ  $k$  次までの自己共分散を表す必要が出てくることから行列を用いて,

$$\begin{aligned} \text{Cov}[\mathbf{y}_t \mathbf{y}_{t-k}] &= [\text{Cov}[y_{i,t}, y_{j,t-k}]]_{ij} \\ &= \begin{pmatrix} \text{Cov}[y_{1,t}, y_{1,t-k}] & \text{Cov}[y_{1,t}, y_{2,t-k}] & \cdots & \text{Cov}[y_{1,t}, y_{n,t-k}] \\ \text{Cov}[y_{2,t}, y_{1,t-k}] & \text{Cov}[y_{2,t}, y_{2,t-k}] & \cdots & \text{Cov}[y_{2,t}, y_{n,t-k}] \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}[y_{n,t}, y_{1,t-k}] & \text{Cov}[y_{n,t}, y_{2,t-k}] & \cdots & \text{Cov}[y_{n,t}, y_{n,t-k}] \end{pmatrix} \end{aligned}$$

で定義される. これを  $k$  次自己共分散行列といい, この行列の  $(i, j)$  成分は  $y_{i,t}$  と  $y_{j,t-k}$  の共分散となっている. このとき, 定義より明らかに, この行列の体積成分は各変数の  $k$  次の自己共分散と等しい. 次の性質は多次元の時系列モデルを数学的に扱う上で, 有用である.

**命題 7.17.**  $k$  次自己共分散行列は  $k = 0$  のとき対称行列となり,  $k \neq 0$  のとき対称行列とならない.

続いて, 定常性を多次元へ拡張する. AR モデルの時と同様, 期待値と自己共分散関数が  $t$  に依存しないとき, ベクトル過程は弱定常と言い, 今後はこれまで同様弱定常正を仮定して議論していく.

**注意 16.** 期待値を  $\boldsymbol{\mu}$ ,  $k$  次の自己共分散行列を  $\Gamma_k$  と書くことにする. 1 変数の場合は, 過程が定常であれば,  $k$  次の自己共分散  $\gamma_k$  は  $\gamma_k = \gamma_{-k}$  を満たしたが, 多次元の場合,  $\Gamma_k = \Gamma_{-k}$  は成り立たず,  $\Gamma_k = \Gamma_{-k}^\top$  となる.

自己共分散行列は異なる時点間の各変数の関係を表すものであるが, 1 変数の時と同様に, 自己共分散行列が単位に依存してしまうことから, これを基準化した自己相関行列を考える.  $k$  次の自己相関行列は

$$\boldsymbol{\rho}_k := \text{Corr}[\mathbf{y}_t, \mathbf{y}_{t-k}] = (\text{Corr}(y_{i,t}, y_{j,t-k}))_{ij}$$

で定義される.

$D$  を  $\mathbf{y}_t$  の分散を対角成分にもつ対角行列であるとする。即ち

$$D := \text{diag}(\text{Var}[y_1], \text{Var}[y_2], \dots, \text{Var}[y_n])$$

とすると、自己相関行列は

$$\boldsymbol{\rho}_k = D^{-1/2} \Gamma_k D^{-1/2}$$

と書くことが出来る。

自己共分散行列と同様、自己相関行列の対角成分は各変数の自己相関となっており  $\boldsymbol{\rho}_k = \boldsymbol{\rho}_{-k}^\top$  が成り立つ。

ホワイトノイズについてもベクトル版で定義しておく。ベクトル過程が  $\boldsymbol{\varepsilon}_t$  がすべての時点  $t$  において、

$$\mathbf{E}[\boldsymbol{\varepsilon}_t] = \mathbf{0}, \mathbf{E}[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_{t-k}^\top] = \begin{cases} \boldsymbol{\Sigma}, & k = 0, \\ \mathbf{0}, & k \neq 0, \end{cases} \quad (7.53)$$

を満たすとき、 $\boldsymbol{\varepsilon}_t$  をベクトルホワイトノイズという。

**系 7.18.** ベクトルホワイトノイズは弱定常過程であり、自己相関を持たない。

$\boldsymbol{\varepsilon}_t$  が平均  $\mathbf{0}$ 、分散共分散行列  $\boldsymbol{\Sigma}$  のベクトルホワイトノイズであることを  $\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma})$  と書いたりする。

### 7.5.1 VAR モデル

ベクトル自己回帰 (VAR; Vector Auto Regressive) モデルは AR モデルの多次元版である。 $n$  変量 VAR( $p$ ) モデルは、 $\mathbf{y}_t$  を定数と、その過程の  $p$  期前の値に回帰したモデルで、

$$\mathbf{y}_t = \mathbf{c} + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\Phi}_2 \mathbf{y}_{t-2} + \dots + \boldsymbol{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}) \quad (7.54)$$

の形で表される。ここで、 $\mathbf{c}$  は  $n \times 1$  の定数ベクトル、 $\boldsymbol{\Phi}_i$  は  $n \times b$  の係数行列である。

$n$  変量 VAR( $p$ ) モデルは、 $n$  本の回帰式からなり、それぞれの回帰式は、各変数を定数と前変数の  $p$  期間の過去の値に回帰した形になっている。パラメータの個数は、1 本の回帰式が定数を含めて  $np + 1$  個の係数を持つため、全体では  $n(np + 1)$  個のパラメータがある。 $\boldsymbol{\varepsilon}_t$  のパラメータは分散共分散行列  $\boldsymbol{\Sigma}$  であり、 $n(n + 1)/2$  個のパラメータがある。纏めると、 $n$  変量 VAR( $p$ ) モデルは合計  $n((np + 1) + (n + 1)/2)$  個のパラメータをもつ大きなモデルとなる。

1 変量 AR( $p$ ) モデルと同様、VAR モデルは必ずしも定常になるとは限らない。定常条件は、AR モデルの条件を行列に拡張したもので与えられ、

$$|\mathbf{I}_n - \boldsymbol{\Phi}_1 z - \boldsymbol{\Phi}_2 z^2 - \dots - \boldsymbol{\Phi}_p z^p| = 0 \quad (7.55)$$

という特性方程式のすべての解の絶対値が 1 より大きいことが VAR モデルの定常条件となる。

**問題 7.1.**  $n$  変量 VAR(1) モデル

$$\mathbf{y}_t = \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$$

について、定常条件は、 $\boldsymbol{\Phi}_1$  のすべての固有値の絶対値が 1 より小さいことである。このことを確認せよ。

VAR モデルの期待値  $\mu$  は、1 変量  $\text{AR}(p)$  の結果の一般化で与えられ

$$\mu = E[y_t] = (I_n - \Phi_1 - \Phi_2 - \cdots - \Phi_p)^{-1} e \quad (7.56)$$

の形になる。自己共分散行列もまた、Yule-Walker 方程式  $\Gamma_k = \Phi_1 \Gamma_{k-1} + \cdots + \Phi_p \Gamma_{k-p}$  により求めることが出来る。

## VAR モデルの推定

VAR モデルの推定は各方程式を個別に OLS で推定するだけで良いため、非常に容易に行うことが出来る。VAR モデルは各方程式は同時点のその他の変数を含む、**同時方程式モデル**ではない。各方程式は誤差項の相関を通じて関係をしているため、**見かけ上無関係な回帰モデル (SUR モデル; Seeming Unrelated Regression model)** と呼ばれる。一般に、SUR モデルを有効的に推定するためには、誤差項の相関を考慮に入れる必要があるため、すべての回帰式を同時に推定する必要がある。しかし、VAR モデルは、すべての回帰式が同一の説明変数を持つため、SUR モデルの中でも特殊な形となっている<sup>8</sup>。この場合には、標準的な SUR モデルの過程の下では、すべての方程式を同時に推定する必要はなく、各方程式を個別に OLS によって推定した係数推定量は漸近性有効性を持つことが知られている。また、 $\varepsilon_t$  が他変量正規分布に従うと仮定した場合、各方程式を個別に OLS によって推定した係数推定量は最尤推定量と一致することも確認できる。

## VAR モデルの次数選択

VAR モデルの次数選択もまた、1 変量の場合と同じように、情報量規準 (IC) を用いて行われることが多い。VAR モデルの場合、ARMA モデルの選択とは異なり、1 つの次数の選択であるため、次数の最大値  $p_{\max}$  以下の VAR モデルをすべて推定し、情報量規準で最適なモデルを選択することが多い。ただし、VAR モデルは多くのパラメータを含むため、有意でないパラメータも多い可能性がある。したがって、AIC や BIC で次数を選択してしまうと、有意でないパラメータの影響を受けて必要以上に小さなモデルを選択してしまう可能性がある。しかし、<sup>okimoto:tsa</sup> 沖本<sup>[4]</sup>によれば、マクロ経済データ間の影響は比較的ゆっくりと伝達することが多いとされ、そのような影響をとらえるためにはある程度の次数が必要となる。そのため、目的に応じては、情報量規準に頼らない経験に基づく定性的な判断がなされることもある。また、マクロ経済データは観測データの数が少ない場合が多いため、モデルのパラメータ数と標本数を比較しながら、一定の推定制度を確保しつつ、出来るだけ大きな次数を選択する場合もある。

### 7.5.2 構造 VAR モデル

VAR モデルは、各方程式は同時点のその他の変数を含む**同時方程式モデル**ではなく、前変数の過去だけを含んでいなかった。これは、同時点の変数が互いに影響を及ぼさないという強い仮定がおかれている。しかし、同時点の変数が互いに影響を及ぼす可能性は大いにあり、**構造 VAR**

<sup>8</sup>Zellner, A. (1962), "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias", Journal of the American Statistical Association, 57, 348-368.

**モデル (structural VAR model)** は、同時点における変数間の関係を明示的に考慮した VAR モデルである。

$n$  変量構造 VAR( $p$ ) モデルは

$$B_0 \mathbf{y}_t = \mathbf{c} + B_1 \mathbf{y}_{t-1} + B_2 \mathbf{y}_{t-2} + \cdots + B_p \mathbf{y}_{t-p} + \tilde{\varepsilon}_t, \tilde{\varepsilon}_t \sim N(\mathbf{0}, D) \quad (7.57) \quad \text{eq:SVAR-p}$$

という形で表される。 $B_0$  は対角成分が 1 の  $n \times n$  行列、 $\mathbf{c}$  は  $n \times 1$  の定数ベクトル、 $B_1, B_2, \dots, B_p$  は  $n \times n$  行列、 $D$  は  $n \times n$  対角行列である。 $\tilde{\varepsilon}_t$  は構造攪乱項、上の (7.57) 式は構造 VAR モデルの構造形と呼ばれる。

構造 VAR モデルと VAR モデルとの違いを確認する。 $\mathbf{y}_t$  に  $B_0$  が掛かっており、これにより各変数の方程式に同時点の他の変数が含まれる。VAR モデルでは  $\varepsilon_t$  の相関によって各変数間の関係を記述していたが、構造 VAR モデルでは同時点の変数間の関係を行列  $B_0$  によって捉えることになるため、構造 VAR モデルでは攪乱項では互いに相関を持たないホワイトノイズになっている<sup>9</sup>。

構造形を用いる上で、構造形が同時方程式モデルであるため、各方程式において説明変数と誤差項が相関を持つために、各方程式を OLS で推定すると、同時方程式バイアスが生じてしまうことがある。従って、構造形は OLS を用いて推定することが出来ない。そこで、構造形の推定に当たっては、各変数の方程式が各変数の過去の値しか含まないように変形した**誘導型 (reduced form)** と呼ばれる形にしてから推定を行う。

具体的には、(7.57) 式の両辺に  $B_0^{-1}$  を左からかけることで次の形で得られる：

$$\begin{aligned} \mathbf{y}_t &= B_0^{-1} \mathbf{c} + B_0^{-1} B_1 \mathbf{y}_{t-1} + B_0^{-1} B_2 \mathbf{y}_{t-2} + \cdots + B_0^{-1} B_p \mathbf{y}_{t-p} + B_0^{-1} \tilde{\varepsilon}_t \\ &= \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \cdots + \Phi_p \mathbf{y}_{t-p} + \varepsilon_t. \end{aligned} \quad (7.58) \quad \text{eq:SVAR-p_reduced}$$

ここで、 $\Phi_0 := B_0^{-1} \mathbf{c}$ ,  $\Phi_i := B_0^{-1} B_i (i = 1, 2, \dots, p)$ ,  $\varepsilon_t := B_0^{-1} \tilde{\varepsilon}_t$  とした。 $\varepsilon_t$  については、 $\tilde{\varepsilon}_t$  と異なり、 $\text{Var}[\varepsilon_t] = B_0^{-1} D (B_0^{-1})^\top$  であるため、 $\varepsilon_t$  の各成分は一般的に相関を持つ。

式 (7.58) をみればわかるように、構造 VAR モデルの誘導形は通常の VAR モデルと同一であり、各方程式に OLS を適用することによって推定ができるようになる。ただし、実際に求めたいのは構造形である式 (7.57) であり、式 (7.58) の推計結果から構造型を復元する必要がある。さて、ここで復元がどのような場合でも出来るならば問題は無いが、誘導形から構造形へ復元できない場合があり、そこには**識別性 (identification)** という問題が存在する。

構造形の式 (7.57) は  $n((n-1) + (np+1) + 1)$  個のパラメータを含む。一方で、誘導形の式 (7.58) は  $n((np+1) + (n+1)/2)$  個のパラメータしか含まないため、パラメータの数が構造形に比べて  $n(n-1)/2$  個少ない。そこで、構造形を識別するためには、構造形に  $n(n-1)/2$  個の制約を課す必要がある。構造形が識別できれば、構造攪乱項が識別できることになるので、誘導形と構造攪乱項を用いて通常の VAR の場合と同様にしてインパルス応答や、分散分解を計算することが出来る。

識別のための手法については、この本稿のレベルを越えるため [hamilton:tsa](#) [3] に譲る。

<sup>9</sup>  $D$  が単なる  $n \times n$  行列ではなく、 $n \times n$  対角行列となっている点に注意

### 7.5.3 各種分析

#### 構造 VAR モデルの場合

構造 VAR モデルでは、攪乱項に互いに相関を持たないホワイトノイズが使われているため、インパルス応答関数や分散分解の計算にあたっては、構造攪乱項  $\varepsilon_t$  を用いることが出来る。

#### TSA1 7章の演習

#### Sec:AR,MA,ARMA 7.2 節の演習

演習 7.1 (AR(1) の性質). 式 (7.4)<sup>eq:AR1</sup>において、パラメータ  $(\phi_0, \phi_1, \sigma)$  を以下のように設定したものをプロットし、その性質を考察してみよう。

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
$\phi_0$	0	2	-2	0	0	0	0	0	0	0
$\phi_1$	0.5	0.5	0.5	-0.5	0.8	0.2	0.5	0.5	1.0	1.1
$\sigma$	1	1	1	1	1	1	0.3	3	1	1

演習 7.2 (MA(1) の性質). 式 (7.9)<sup>eq:MA1</sup>において、パラメータ  $(\theta_0, \theta_1, \sigma)$  を以下のように設定したものをプロットし、その性質を考察してみよう。

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
$\theta_0$	0	2	-2	0	0	0	0
$\theta_1$	0.5	0.5	0.5	0.8	0.2	0.5	0.5
$\sigma$	1	1	1	1	1	3	0.3

#### Sec:ModelChoice 7.3 節の演習

演習 7.3. TOPIX の四半期データをもとに、

1. ラグ  $k = 1, 2, \dots, 20$  の自己相関、偏自己相関を計算し、グラフにプロットせよ。
2. 前問の結果をもとに、候補となるモデルをいくつか選択し、AIC, BIC をそれぞれについて計算することで TOPIX の四半期推移を表現するのに適したモデルを決定せよ。

#### Sec:Prediction 7.4 節の演習

演習 7.4. 前問で選択したいくつかのモデルに対して、10 期までの自己相関をかばん検定で検定せよ。

演習 7.5. 内閣府のホームページ<sup>10</sup>より GDP ギャップのデータを取得し、<sup>Sec:Prediction</sup>7.4 節の例を確認してみよ。

<sup>10</sup><https://www5.cao.go.jp/keizai3/getsurei/getsurei-index.html>



## 第8章 動学的マクロ経済

### 8.1 Solow model

#### 8.1.1 準備・基本性質

##### モデルの設定

$t$  期初の資本ストック  $K_t$ 、 $t$  期の労働投入  $L_t$ 、技術水準  $A_t$  によって  $t$  期の一国全体の生産  $Y_t$  が決まる経済を考える。具体的には、 $\alpha$  を資本分配率として、生産  $Y_t$  は

$$Y_t = K_t^\alpha (A_t L_t)^{1-\alpha}$$

で表現されているとする。家計は、その期の生産  $Y_t$  を消費と投資に割り振るものとする。一国が貿易の無い閉鎖経済を仮定すると、投資と貯蓄は常に等しい。家計の貯蓄率を  $s$  とすると、 $t$  期の投資は  $sY_t$  となり、残余  $(1-s)Y_t$  が消費へ充てられる。 $\delta$  を資本の減耗率として  $K_{t+1}$  は、

$$K_{t+1} = (1-\delta)K_t + sY_t$$

と表される。これら 2 式の方程式からなるモデル：

$$Y_t = K_t^\alpha (A_t L_t)^{1-\alpha}, \quad (8.1)$$

$$K_{t+1} = (1-\delta)K_t + sY_t \quad (8.2)$$

を **Solow model** という。

簡単のため、 $A_t, L_t$  は定率成長するものと仮定し、技術進歩率  $g$  および労働成長率  $n$  を用いて

$$A_{t+1} = (1+g)A_t, \quad (8.3)$$

$$L_{t+1} = (1+n)L_t \quad (8.4)$$

とあらわされているものとする。

Solow model における内生変数は  $(Y_t, K_t)$  の 2 個であり、外生変数は  $\alpha, s, \delta, g, n$  の 5 つである。 $K_1$  が初期値として与えられれば、系列  $\{(Y_t, K_t); t = 1, 2, 3, \dots\}$  を得ることが出来る。

このモデルには、変数がある一定の値のまま時間が経過しても変化しない、所謂「定常状態」がただ一つ存在する。実際、どのような初期値  $K_1$  からスタートしても、時間が十分に経過すると

$$k_t := \frac{K_t}{A_t L_t} \quad (8.5)$$

が変化しない状態に到達する。この  $k_t$  は効率労働当たり資本と呼ばれる物で、定常状態における  $k_t$  の値を  $k^*$  と書くことにする。 $k^*$  は  $k_t$  の定義から

$$(1+g)(1+n)k_{t+1} = (1-\delta)k_t + sk_t^\alpha \quad (8.6) \quad \boxed{\text{eq:11.6}}$$

と展開していくことが出来、 $k_{t+1} = k_t = k^*$  を代入し、 $k^*$  について解けば

$$k^* = \left( \frac{s}{(1+g)(1+n) - (1-\delta)} \right)^{1/(1-\alpha)} \approx \left( \frac{s}{\delta + n + g} \right)^{1/(1-\alpha)} \quad (8.7)$$

を得る。従って、パラメータ  $\alpha, s, \delta, g, n$  が与えられれば、 $k^*$  は一意に定まる。

$\Delta k_{t+1} := k_{t+1} - k_t$  とする。<sup>eq:11.6</sup> (8.6) 式を整理すれば

$$k_{t+1} = \frac{(1-\delta)k_t + sk_t^\alpha}{(1+g)(1+n)} \quad (8.8)$$

両辺から  $k_t$  を差し引くと、

$$\begin{aligned} \Delta k_{t+1} &= \frac{\{(1-\delta) - (1+g)(1+n)\} k_t + sk_t^\alpha}{(1+g)(1+n)} \\ &= \frac{sk_t^\alpha - (\delta + g + n + gn)k_t}{(1+g)(1+n)} \end{aligned} \quad (8.9)$$

となる。分子に注目すれば、 $sk_t^\alpha > (\delta + g + n + gn)k_t$  のとき  $k_t$  は増加、 $sk_t^\alpha < (\delta + g + n + gn)k_t$  のとき  $k_t$  は減少する。定常状態では  $\Delta k_{t+1} = k_{t+1} - k_t = 0$  なので、 $sk_t^\alpha = (\delta + g + n + gn)k_t$  のとき、 $k_t = k^*$  となる。

### 8.1.2 貯蓄と消費の関係

Solow model において、貯蓄率  $s$  を定数として見てきたが、 $s \in (0, 1)$  の範囲で動くとき、定常状態がどのように変化するかを考える。このモデルでは、生産  $Y_t$  のうち  $sY_t$  が投資に充てられ、 $(1-s)Y_t$  が消費に充てられる。効率労働あたり消費として

$$c_t := \frac{(1-s)Y_t}{A_t L_t} = (1-s)k_t^\alpha \quad (8.10) \quad \text{eq:SM}_1$$

と定義する。

定常状態における  $c_t$  を  $c^*$  としたとき、 $s$  と  $c^*$  の関係を調べる。1人あたりの消費は  $C_t/L_t$  であるが、 $A_t$  は外生変数であるから  $c_t$  が最大化されているときには1人あたり消費も最大化されていると考えてよいだろう。<sup>eq:SM\_1</sup> (8.10) 式に  $k_t = k^*$  を代入し、 $k^* =$

## 8.2 Ramsey model

前節の Solow model に効用関数と予算制約を導入しよう。

市場には  $J$  個の家計が存在し、市場から退場することは無いとする。家計  $j$  は以下のような生涯効用関数

$$U_{t,j} := \sum_{i=0}^{\infty} \beta^i \ln(C_{t+i,j}) \quad (8.11) \quad \text{util-req}$$

を持つとする。家計  $j$  は  $t$  期のはじめに資本ストック  $K_{t,j}$  を持ち、それを企業に  $r_t$  という利子率で貸し出すものとする。また、家計は一定の労働を供給することにより賃金  $w_{t,j}$  を得るものと

する。その所得から、家計  $j$  は消費  $C_{t,j}$  を行い、残りは投資するものとする。資本の消耗率は一律で  $\delta$  であるとする、これらを纏めて  $t$  期の予算制約は

$$K_{t+1,j} = r_t K_{t,j} + w_{t,j} - C_{t,j} + (1 - \delta) K_{\delta,j} \quad (8.12)$$

となる。

簡単のため、すべての家計について、その選好や能力はまったく同一であるという「代表的家計」という概念を用いる。つまり、添え字  $j$  により家計の別を表現していたが、これらを考える必要がなくなる。代表的家計における効用最大化問題は

$$\begin{aligned} \max_{\mathbf{C}} U_t &= \sum_{i=0}^{\infty} \beta^i \ln(C_{t+i}) \\ \text{subject to } K_{i+1} &= r_i K_i + w_i - C_i + (1 - \delta) K_i, \quad i \geq t \end{aligned} \quad (8.13)$$

として定式化できる。ここで、 $\mathbf{C} := \{C_{t+i}\}_{i=0}^{\infty}$  とした。資本の初期値  $K_t$  を所与のものとする。Lagrangian は

$$\Lambda = \sum_{i=t}^{\infty} \beta^{i-t} \ln(C_i) + \sum_{i=t}^{\infty} \lambda'_i (r_i K_i + w_i - C_i(1 - \delta) K_i - K_{i+1}) \quad (8.14)$$

とおけるが、ここで  $\lambda'_i = \beta^{i-t} \lambda_t$  と置き換えると、

$$\Lambda = \sum_{i=t}^{\infty} \beta^{i-t} [\ln(C_i) + \lambda'_i (r_i K_i + w_i - C_i(1 - \delta) K_i - K_{i+1})] \quad (8.15)$$

となる。1 階の条件を求めよう。家計がコントロール可能である  $C_t$  ならびに、 $C_t$  のコントロールによって間接的にコントロールされる  $K_{t+1}$  で偏微分すると

$$\frac{1}{C_t} - \lambda_t = 0, \quad (8.16) \quad \boxed{\text{EqEuler-1}}$$

$$\beta(r_{t+1} - \delta + 1)\lambda_{t+1} - \lambda_t = 0 \quad (8.17) \quad \boxed{\text{EqEuler-2}}$$

を得る<sup>1</sup>。

**注意 17.** 代表的家計による議論が成り立つためには、効用関数が式 (8.11) のような相対的危険回避度一定型である必要がある。実際、添え字  $j$  を考えながら式展開をして最後に  $C_t = \sum_{j=1}^J C_{t,j}$  を求めても、最初から  $C_{t,j}$  を  $C_t$  に置き換えて式展開しても同じ結果を得ることが出来る。

続いて、企業に目を向ける。多数の企業が存在し、企業  $k$  の生産関数を

$$Y_{t,k} = A_t (K_{t,k})^\alpha (L_{t,k})^{1-\alpha} \quad (8.18)$$

とする。ただし、以降は家計と同様に、代表的企業という概念を用いることとする。

代表的企業の利潤最大化の一階条件は

$$\alpha A_t \left( \frac{K_t}{L_t} \right)^{1-\alpha} - r_t = 0, \quad (8.19)$$

$$(1 - \alpha) A_t \left( \frac{K_t}{L_t} \right)^\alpha - w_t = 0 \quad (8.20)$$

<sup>1</sup> この関係は、 $i > t$  なる  $C_i, K_{t+i}$  においても同様に成り立つので一般の  $t$  で成り立つ

となる。これが成り立つように  $K_t/L_t, r_t, w_t$  がそれぞれ調整される。

家計により供給される労働力  $L_t$  が1であるとき、これらの条件は

$$\alpha A_t K_t^{1-\alpha} - r_t = 0, \quad (8.21) \quad \text{EqEuler-3}$$

$$(1 - \alpha) A_t K_t^\alpha - w_t = 0 \quad (8.22)$$

となり、市場全体の資本ストックの増分  $\Delta K_{t+1}$  は生産から消費と減耗を差し引いたものに等しいから

$$\begin{aligned} \Delta K_{t+1} &:= K_{t+1} - K_t \\ &= Y_t - C_t - \delta K_t \\ &= A_t K_t^\alpha - C_t - \delta K_t \end{aligned} \quad (8.23)$$

となり、この式から資本の遷移式

$$K_{t+1} = A_t K_t^\alpha + (1 - \delta) K_t - C_t \quad (8.24)$$

を得る。この資本の遷移式と [EqEuler-1](#) 式、[EqEuler-2](#) 式と [EqEuler-3](#) 式より得られる Euler 方程式

$$\frac{C_{t+1}}{C_t} = \beta(\alpha A_{t+1} K_{t+1}^{\alpha-1} - \delta + 1) \quad (8.25)$$

の2式が Ramsey model の核となる部分である。

**問題 8.1.** Ramsey model においても家計の所得  $R_t K_t + w_t$  が生産  $Y_t$  と等しいことを確認せよ。

### 8.3 Real Business Cycle model

Ramsey model には、景気変動の要素を加え実物的景気循環モデル (Real Business Cycle model; RBC model) を紹介する。このモデルを発展させたものが目標である動学的確率的一般均衡モデル (DSGE model) である。

#### 8.3.1 モデルのセットアップ

##### 代表的家計

まずは、代表的家計から準備を始める。

代表的家計の生涯効用関数を

$$U_t = \sum_{i=0}^{\infty} \beta^i (\ln(C_{t+i}) - \mu L_{t+i}^{\gamma+1}) \quad (8.26)$$

とする。ここで、 $\mu L_t^\gamma + 1$  は労働の不効用と呼ばれるものを表しており、 $\mu$  は労働の不効用の相対ウェイト、 $\gamma$  は労働供給の弾力性の逆数とする。代表的家計は  $t$  期のはじめに資本ストック  $K_t$  を有し、それを企業に利子率  $r_t$  で貸し出すものとする。また、労働1単位あたりの賃金を  $w_t$  とし、家計は労働  $L_t$  を供給することで、労働による所得  $w_t L_t$  を得るものとする。その所得のうち、

代表的家計は消費  $C_t$  を行い、残りを投資に回すものとする。資本の減耗率を  $\delta$  とすると、 $t$  期における予算制約は

$$K_{t+1} + C_t = r_t K_t + w_t L_t + (1 - \delta) K_t \quad (8.27)$$

となる。よって、家計の異時点間の最適化問題は

$$\begin{aligned} \max_{C, L} U_t &= \sum_{i=0}^{\infty} \beta^i \left( \ln(C_{t+i}) - \mu L_{t+i}^{\gamma+1} \right), \\ \text{subject to } K_{i+1} + C_i &= r_i K_i + w_i L_i + (1 - \delta) K_i, i \geq t \end{aligned} \quad (8.28)$$

と定式化できる。この最適化問題においては、 $\lambda_t, \lambda_{t+1}, \dots$  を Lagrange 乗数とし、Lagrangean を

$$\Lambda = \sum_{i=t}^{\infty} \beta^{i-t} \left[ \ln(C_i) - \mu L_i^{\gamma+1} + \lambda_i (r_i K_i + w_i L_i + (1 - \delta) K_i - K_{i+1} - C_i) \right] \quad (8.29)$$

とする Lagrange の未定乗数法で最適化の 1 階条件を求めることが出来る。 $\Lambda$  を  $C_t, L_t, K_{t+1}$  で偏微分した結果が 0 となることから条件は

$$\frac{1}{C_t} - \lambda_t = 0, \quad (8.30)$$

$$w_t \lambda_t - (\gamma + 1) \mu L_t^{\gamma} = 0, \quad (8.31) \quad \text{eq46}$$

$$\beta(r_{t+1} - \delta + 1) \lambda_{t+1} - \lambda_t = 0 \quad (8.32) \quad \text{eq47}$$

である。 $\lambda_t = 1/C_t$  であるから、(8.31) 式と (8.32) 式から  $\lambda_t, \lambda_{t+1}$  を消して整式すれば

$$\frac{w_t}{C_t} = (\gamma + 1) \mu L_t^{\gamma}, \quad (8.33)$$

$$\frac{C_{t+1}}{C_t} = \beta(\gamma_{t+1} - \delta + 1) \quad (8.34)$$

を得る。2 式目は Ramsey model でも登場した Euler 方程式で、RBC model においては、労働を内製化していくため、消費と労働の関係を表す 1 式目が新たに登場する。

### 代表的企業

続いて、代表的企業の準備に掛かろう。

代表的企業の生産関数を

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha} \quad (8.35)$$

とする。このとき、利潤最大化の一階条件は、

$$r_t = \alpha A_t K_t^{\alpha-1} L_t^{1-\alpha}, \quad (8.36)$$

$$w_t = (1 - \alpha) A_t K_t^{\alpha} L_t^{-\alpha} \quad (8.37)$$

となる。

### 財市場における均衡と技術水準 $A_t$ の遷移

繰り返しになるが、市場全体の資本ストックの増分は、生産から消費と減耗を差し引いた分に等しいので Ramsey model の時同様

$$K_{t+1} = Y_t + (1 - \delta)K_t - C_t \quad (8.38)$$

という資本の遷移式が得られる。技術水準  $A_t$  は、技術ショック  $\varepsilon_t$  を外生変数、外生ショックの持続性に関するパラメータを  $\rho$  として、以下のような AR(1) 過程に従うものとする：

$$\ln(A_{t+1}) = \rho \ln(A_t) + \varepsilon_{t+1}, \rho \in [0, 1). \quad (8.39)$$

**問題 8.2.** 先の技術水準  $A_t$  に関する AR(1) 過程について、定常状態における  $A_t$  の値  $A^*$  が  $A^* = 1$  であることを確認せよ。

パラメータ  $\rho$  の役割については、図示すればより明快となるが、時系列分析の基本的な内容であることからここでは式だけを見て性質を説明する。 $\rho = 0$  のケースは、技術のショックはその期の技術水準はその期の技術水準のみに影響を与え、その後の技術水準には全く影響を与えない。 $\rho = 0.5$  のとき、技術ショックの影響は1期ごとに半減していく。 $\rho$  が1に近づくほど、この影響は長期化していく。 $\rho = 1$  の場合には、一時の技術ショックの影響がいつまでも消えることなく技術水準  $A_t$  に影響を与え続けるため、 $A_t$  の定常状態は存在しなくなること注意到しておく。

**問題 8.3.** RBC モデルにおいても  $t$  期における家計の所得  $r_t K_t + w_t L_t$  と生産  $Y_t$  が等しいことを確認せよ。

RBC model の全容を纏めておく。RBC model は

$$\frac{w_t}{C_t} = (\gamma + 1)\mu L_t^\gamma, \quad (8.40)$$

$$\frac{C_{t+1}}{C_t} = \beta(r_{t+1} - \delta + 1), \quad (8.41)$$

$$Y_t = A_t K_t^\alpha L_t^{1-\alpha}, \quad (8.42)$$

$$w_t = (1 - \alpha)A_t K_t^\alpha L_t^{-\alpha}, \quad (8.43)$$

$$K_{t+1} = Y_t + (1 - \delta)K_t - C_t, \quad (8.44)$$

$$r_t = \alpha A_t K_t^{\alpha-1} L_t^{1-\alpha}, \quad (8.45)$$

$$\ln(A_{t+1}) = \rho \ln(A_t) + \varepsilon_{t+1} \quad (8.46)$$

という7本の方程式からなる。

ここで内生変数は  $C_t, L_t, K_t, Y_t, w_t, r_t, A_t$  の7つで、Ramsey model から賃金  $w_t$  と労働  $L_t$  が新たに内生化されたことになる。家計・企業が間接的にしか制御できないか、全く制御できない変数は  $K_t, A_t$  である。

モデルのパラメータは、割引因子  $\beta$ 、労働の不効用の相対ウェイト  $\mu$ 、労働供給の弾力性の逆数  $\gamma$ 、資本分配率  $\alpha$ 、減耗率  $\delta$ 、技術水準における AR(1) モデルの係数  $\rho$  の6つであり、Ramsey model と比較すると賃金と労働を新たに内生化したことによって、 $\mu$  と  $\gamma$  が新たに追加された。

### 8.3.2 Impulse 応答

モデルがどのような特徴を持っているかを確認すべく、 $t = 1$  期において、+1% 技術ショックが加わった場合の反応を計算してみよう。

### 8.3.3 RBC model から New Keynesian model へ

RBC model が与える示唆はラフに纏めれば

1. 一時的な技術ショックに対して、生産・消費・投資・労働・賃金・利子率といった変数が同じ方向へ動くということ
2. 消費が生産や投資といったほかの変数よりも滑らかに動くという予想を与える

これらは、どちらも実際に観察される現象と一致する結果を与えているとされている。技術的ショックを導入することにより、これらの現象を整合的に説明するという意味で、RBC model は魅力的な理論モデルである一方、不十分性が指摘されている点も存在する。

1. 技術水準  $A_t$  は生産関数から計算すれば Solow 残差と呼ばれるものに一致するが、ソロー残差が技術水準だけを表しているかという疑問が生じる。
2. RBC model の想定では、実体経済に何らかのショックが起きたとき、中央銀行が何ら対処しないか、あるいはその対処は全く効果がないということになり、金融政策が実態経済にどのような影響を与えるのか・与えたのかを説明することができない。

これらの疑問は、RBC model が景気循環を技術ショックだけで説明しようとすることに起因している。RBC model にインフレ率や政策金利・名目金利を導入すれば解決するのではないかということは既に発想されており、一定の成果を収めている。比較的単純な手法で Cash In Advance(CIA) という方法があり、さらに手の込んだインフレ率の導入法として、価格の硬直性・非伸縮性をどうぬ輸する方法もある。RBC モデルに Calvo 型価格設定と呼ばれる硬直価格理論を導入した New Keynesian model と呼ばれるものは、現代のマクロ経済学における代表的な分析ツールの 1 つとなっている。

## 8.4 New Keynesian model

RBC model の変数はすべて実物変数で、インフレ率や政策金利といった名も高変数はそこに含まれなかったが、そのような変数を内生化するようにモデルに工夫をしていく。

### 8.4.1 導入：CIA model

インフレ率を内生化する 1 つの方法として Cash in advance model(CIA model) を説明する。CIA model は、RBC model での家計の予算制約とは別に、CIA 制約と呼ばれる制約条件を加えることで構成できる。単純化のため、資本ストックの無い消費財のみが存在するモデルを想定して話を薦める。

## セッティング

代表的家計の生涯効用関数を RBC model 同様

$$U_t := \sum_{t=0}^{\infty} \beta' \left\{ \ln(C_{t+i} - \mu L_{t+i}^{\gamma+1}) \right\} \quad (8.47)$$

とする。労働 1 単位当たりの名目賃金を  $W_t$  とし、家計は労働  $L_t$  を供給することで、労働所得  $W_t L_t$  を得るものとする。家計は、その所得と資産を名目消費  $P_t C_t$  と貨幣保有  $M_t$ 、債券保有  $B_t$  のいずれかに振り分けるものとする。債券保有  $B_t$  は名目建てで、名目利子率  $i_t$  の利回りとし、名目貨幣保有  $M_t$  には利子がつかないものとする。家計の  $t$  期における予算制約は

$$P_t C_t + M_t + B_t = W_t L_t + M_{t-1} + (1 + i_{t-1}) B_{t-1} \quad (8.48)$$

となるが、これに加えて新たに

$$P_t C_t = M_{t-1} \quad (8.49)$$

なる CIA 制約<sup>2</sup>がかかるものとする。

この場合の Lagrangean は、2 つの乗数  $\tilde{\lambda}_t, \tilde{\xi}_t$  を用いて

$$\Lambda = \sum_{i=t}^{\infty} \beta^{i-t} \left[ \ln(C_i) - \mu L_t^{\gamma+1} + \tilde{\lambda}_i \{ W_i L_i + M_{i-1} + (1 + i_{i-1}) B_{i-1} - P_i C_i - M_i - B_i \} + \tilde{\xi}_i (M_{i-1} - P_i C_i) \right] \quad (8.50)$$

と×、この  $\Lambda$  を  $C_t, L_t, M_t, B_t$  で偏微分し、最適化の 1 階条件

$$\frac{1}{C_t} - P_t \tilde{\lambda}_t - P_t \tilde{\xi}_t = 0, \quad (8.51) \quad \text{eq:CIA_Lag1}$$

$$W_t \tilde{\lambda}_t - \mu L_t^{\gamma} (\gamma + 1) = 0, \quad (8.52) \quad \text{eq:CIA_Lag2}$$

$$\beta \tilde{\lambda}_{t+1} + \beta \tilde{\xi}_{t+1} - \tilde{\lambda}_t = 0, \quad (8.53) \quad \text{eq:CIA_Lag3}$$

$$\beta(1 + i_t) \tilde{\lambda}_{t+1} - \tilde{\lambda}_t = 0 \quad (8.54) \quad \text{eq:CIA_Lag4}$$

を得ることが出来る。

続いて、企業は労働投入に比例する消費財を生産するものとして、代表的企業の生産関数を

$$C_t = A_t L_t \quad (8.55) \quad \text{eq:CIA_principal}$$

とする。ここで  $A_t$  は技術水準を意味する。消費財  $C_t$  の 1 単位あたりの価格を  $P_t$  で表せば、企業の利潤  $\Pi_t$  は

$$\Pi_t = P_t C_t - W_t L_t \quad (8.56)$$

となる。従って、利潤最大化の 1 階条件は

$$P_t A_t - W_t = 0 \quad (8.57)$$

<sup>2</sup>すなわち、家計が当期に  $P_t C_t$  だけ消費するには、前期末にそれと動学の名目貨幣  $M_{t-1}$  を保有しないといけないという制約。



である。技術水準  $A_t$  は、技術ショック  $\varepsilon_t$  を外生変数とし、RBC model 同様

$$\ln(A_{t+1}) = \rho_A \ln(A_t) + \varepsilon_{t+1} \quad (8.58)$$

に従うものとする。

中央銀行による名目貨幣供給は、

$$M_t = (1 + \zeta_t)M_{t-1} \quad (8.59)$$

という過程に従うものとする。ここで、 $\zeta_t$  は名目貨幣供給の増加率を意味する。 $\zeta_t$  は  $\zeta$  を定常値とする AR(1) 過程

$$\zeta_{t+1} - \zeta = \rho_\zeta(\zeta_t - \zeta) + \varepsilon'_{t+1} \quad (8.60)$$

に従うものとする。ここで、 $\varepsilon'_t$  は貨幣供給ショックを表す。

### モデルの全体像

CIA model では、インフレ率  $\pi_t$  が  $\pi_t = P_t/P_{t-1} - 1$  のときのみ定常値をもち、 $P_t$  が定常値を持たないため、モデル全体を  $P_t$  を持たない形に書き換えられる必要がある。具体的には、<sup>eq:CIA\_Lag1</sup> (8.51) 式は、 $\lambda := P_t \tilde{\lambda}_t, \xi := P_t \tilde{\xi}_t$  とすれば

$$\frac{1}{C_t} - \lambda_t - \xi_t = 0, \quad (8.61)$$

<sup>eq:CIA\_Lag2</sup> (8.52) 式は、 $L_t := C_t/A_t, W_t := P_t A_t$  とすれば

$$A_t \lambda_t - \mu \left( \frac{C_t}{A_t} \right)^\gamma (\gamma + 1) = 0, \quad (8.62)$$

<sup>eq:CIA\_Lag3</sup> (8.53), <sup>eq:CIA\_Lag4</sup> (8.54) 式についても  $\pi_t$  を用いて

$$\beta \lambda_{t+1} + \beta \xi_{t+1} - (1 + \pi_{t+1}) \lambda_t = 0, \quad (8.63)$$

$$\beta(1 + i_t) \lambda_{t+1} - (1 + \pi_{t+1}) \lambda_t = 0 \quad (8.64)$$

を得る。

続いて、実質貨幣保有  $m_t := M_t/P_t$  を用いれば、CIA 制約 <sup>eq:CIA\_principal</sup> (8.55) 式と名目貨幣供給  $M_t$  は

$$C_t = \frac{m_{t-1}}{1 + \pi_t}, \quad (8.65)$$

$$(1 + \pi_t)m_t = (1 + \xi_t)m_{t-1} \quad (8.66)$$

と書ける。技術水準  $A_t$  および名目貨幣供給の増加率  $\zeta_t$  の過程を含めた 8 本の式をモデル方程式として用いる。内生変数は、消費 ( $C_t$ )、名目利子率 ( $i_t$ )、技術水準 ( $A_t$ )、インフレ率 ( $\pi_t$ )、実質貨幣保有 ( $m_t$ )、名目貨幣供給の増加率 ( $\zeta_t$ )、ラグランジュ乗数 (予算制約、 $\lambda_t$ )、ラグランジュ乗数 (CIA 制約、 $\xi_t$ ) の 8 つであり、うち  $A_t, \zeta_t$  の 2 個が状態変数、技術ショック ( $\varepsilon_t$ ) と貨幣供給ショック ( $v_t$ ) の 2 つが外生変数である。パラメータは割引因子 ( $\beta$ )、労働の不効用の相対ウェイト ( $\mu$ )、労働供給の弾力性の逆数 ( $\gamma$ )、技術水準の AR(1) モデルにおける係数  $\rho_A$ 、名目貨幣供給の AR(1) モデルにおける係数  $\rho_\zeta$ 、 $\zeta_t$  の定常値 ( $\zeta$ ) の 6 つである。

定常状態における内生変数それぞれの値は  $C_t = C_{t+1} = C^*$  として置き換えることで得られる連立方程式を解くことで

$$A^* = 1, \quad (8.67)$$

$$\pi^* = \zeta^* = \zeta, \quad (8.68)$$

$$C^* = \left( \frac{\beta}{\mu(\gamma+1)(1+\zeta)} \right)^{\frac{1}{\gamma+1}}, \quad (8.69)$$

$$m^* = (1+\zeta)C^*, \quad (8.70)$$

$$\lambda^* = \frac{\beta}{(1+\zeta)C^*}, \quad (8.71)$$

$$\xi^* = \frac{1}{C^*} \left( \frac{1+\zeta-\beta}{1+\zeta} \right) \quad (8.72)$$

となる。

#### 8.4.2 New Keynesian model

CIA model では、貨幣をモデルに導入することで、インフレ率を内生化してみせた。次に紹介する New Keynesian model では、独占的競争モデルと Calvo 型価格設定モデルを組み合わせることで貨幣をモデルに導入しなくてもインフレ率を内生化することが出来るようになる。

##### 独占的競争モデル

価格の粘着性をモデルに導入するためには、独占的競争モデルが必要となる。このモデルは、財の供給者は多数存在するが、それぞれ他社と異なる材を生産するために、供給側が自由に価格を決定することが出来るとするモデルである。

家計の効用関数に CES 型の効用関数<sup>3</sup>を仮定、具体的には

$$C = \left( \int_0^1 c_j^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}} \quad (8.73)$$

とする。経済には  $n$  種類の財  $\{c_1, c_2, \dots, c_n\}$  が存在するもの俊、積分区間  $[0, 1]$  の分割を  $\{0, 1/n, 2/n, \dots, (n-1)/n, 1\}$  とする。このとき、先の  $C$  は極限を用いて

$$C = \left( \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n c_j^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \quad (8.74)$$

と表現することが出来る。

財  $c_j$  の価格を  $p_j$  で表すことにすると、家計の効用最大化問題は費用最小化問題として

$$\min_{\{c_j\}} \int_0^1 p_j c_j dj \text{ subject to } C - \left( \int_0^1 c_j^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}} = 0 \quad (8.75)$$

<sup>3</sup>CES; Constant Elasticity of Substitution: だいたいの弾力性一定。経済には無限に多くの異なる種類の剤が存在することを仮定している。

と定式化できる。Lagrange 乗数を  $\psi$  とすると、Lagrangean は

$$\Psi = \int_0^1 p_j c_j dj + \psi \left[ C - \left( \int_0^1 c_j^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}} \right] \quad (8.76)$$

と定義できる。最適化の 1 階条件は  $\partial \Psi / \partial c_j = 0$  であるから、

$$\frac{\partial}{\partial c_j} \int_0^1 p_k c_k dk - \psi \frac{\partial}{\partial c_j} \left( \int_0^1 c_k^{\frac{\eta-1}{\eta}} dk \right)^{\frac{\eta}{\eta-1}} = 0 \iff c_j = \left( \frac{p_j}{\psi} \right)^{-\eta} C \quad (8.77) \quad \boxed{\text{eq: CES\_Lag}}$$

を得る。この条件を家計の効用関数へ代入すれば、

$$\begin{aligned} C &= \left[ \int_0^1 \left\{ \left( \frac{p_j}{\psi} \right)^{-\eta} C \right\}^{\frac{\eta-1}{\eta}} dj \right]^{\frac{\eta}{\eta-1}} \\ &= \left( \frac{1}{\psi} \right)^{-\eta} \left( \int_0^1 p_j^{1-\eta} dj \right)^{\frac{\eta}{\eta-1}} C \end{aligned} \quad (8.78)$$

を得る。これを  $\psi$  について解くと

$$\psi = \left( \int_0^1 p_j^{1-\eta} dj \right)^{\frac{1}{1-\eta}} \quad (8.79)$$

となる。

名目消費は、 $\int_0^1 p_j c_j dj$  で定義でき、上の 1 階条件と  $\psi$  を用いれば

$$\begin{aligned} \int_0^1 p_j c_j dj &= \int_0^1 \left\{ \left( \frac{p_j}{\psi} \right)^{-\eta} C \right\} p_j dj \\ &= C \psi^\eta \int_0^1 p_j^{1-\eta} dj \\ &= \psi^\eta \psi^{1-\eta} C \\ &= \psi C \end{aligned} \quad (8.80)$$

となるので、Lagrange 乗数  $\psi$  は一般物価を表すことがわかる。また、集計された実質消費  $C$  の定義も最初の家計効用関数で良いことがわかる。

一方、1 階条件の  $\boxed{\text{eq: CES\_Lag}}$  (8.77) 式の両辺を自然対数をとると

$$\ln(c_j) = -\eta \ln \left( \frac{p_j}{\psi} \right) + \ln(C) \quad (8.81)$$

で、 $C$  が一定として、 $c_j, p_j, \psi$  がそれぞれ、 $c'_j, p'_j, \psi'$  に変化したとき、差分を取れば

$$\ln(c'_j) - \ln(c_j) = -\eta \left\{ \ln \left( \frac{p'_j}{\psi'} \right) - \ln \left( \frac{p_j}{\psi} \right) \right\} \quad (8.82)$$

である。これを見れば、財  $c_j$  に対する需要の価格弾力性が  $\eta$  であることが分かる。

**問題 8.4.**  $\boxed{\text{eq: CES\_Lag}}$  (8.77) を確認せよ。

### New Keynesian model のセッティング

例のごとく、代表的家計から始める。代表的家計の生涯効用関数を RBC model の効用関数と、独占的競争モデルの効用関数を組み合わせて；

$$U_t := \sum_{i=0}^{\infty} \beta^i \left\{ \ln(C_{t+i}) - \mu L_{t+i}^{\gamma+1} \right\}, \quad (8.83)$$

$$C_t := \left( \int_0^1 c_{j,t}^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}} \quad (8.84)$$

と定める。このとき、代表的家計の最適化行動は2段階に分かれる。[1] 家計は当期の消費  $C_t$  の大きさに依存することなく費用を最小化するように、当期の消費に占める企業  $j$  の生産する  $c_{j,t}$  の割合を決め、[2] 異時点の予算制約に基づき当期の消費  $C_t$  を決定する。

**Step.1** 費用最小化行動は、財  $c_{j,t}$  の価格を  $p_{j,t}$  とすると、

$$\begin{aligned} \min_{\{c_{j,t}\}} & \int_0^t p_{j,t} c_{j,t} dj, \\ \text{subject to } & C_t - \left( \int_0^1 c_{j,t}^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}} = 0 \end{aligned} \quad (8.85)$$

と定式化できる。先ほど得られた結果を用いれば、一般物価  $P_t$  と財  $c_j$  に対する  $t$  期の需要  $c_{j,t}$  は、それぞれ

$$P_t = \left( \int_0^1 p_{j,t}^{1-\eta} dj \right)^{\frac{1}{1-\eta}}, \quad (8.86)$$

$$c_{j,t} = \left( \frac{p_{j,t}}{P_t} \right)^{-\eta} C_t \quad (8.87)$$

と表される。労働1単位あたりの名目賃金を  $W_t$  とし、家計は労働  $L_t$  を供給することで、労働所得  $W_t L_t$  を得て、その所得俊さんを消費  $C_t$  と債券保有  $B_t$  のいずれかに振り分けるものとする。債券保有  $B_t$  は名目建てで、名目利子率  $i_t$  の利回りとする。名目利子率  $i_t$  は、中央銀行の金融政策により決定されるものとする。このとき、 $t$  期の予算制約は

$$P_t C_t + B_t = W_t L_t + (1 + i_{t-1}) B_{t-1} + \tau_{j,t} \quad (8.88)$$

と表される。RBC model や CIA model の場合とは異なり、独占的競争によって生じる企業の利潤の家計への移転  $\tau_{j,t}$  が加わる。企業が家計によって所有されていると仮定されているためである<sup>4</sup>。

**Step.2** 家計の効用最大化は

$$\begin{aligned} \max_{C,L} & \sum_{i=0}^{\infty} \beta^i \left\{ \ln(C_{t+i}) - \mu L_{t+i}^{\gamma+1} \right\} \\ \text{subject to } & C_i + \frac{B_i}{P_i} = \frac{1}{P_t} \{ W_i L_i + (1 + i_{i-1}) B_{i-1} + \tau_{j,i} \}, \quad i \geq t \end{aligned} \quad (8.89)$$

<sup>4</sup>Step.2 より分かるように家計の最適化行動には影響しない。

と定式化できる。予算制約は両辺を  $P_i$  で割ることで実質化した。 $\lambda_t$  をラグランジュ乗数とすると、ラグランジアンは

$$\Lambda = \sum_{i=t}^{\infty} \beta^{i-t} \left[ \ln(C_i) - \mu L_i^{\gamma+1} + \lambda_i \left\{ \frac{W_i L_i + (1 + i_{i-1}) B_{i-1} + \tau_{j,i} - B_i}{P_i} - C_i \right\} \right] \quad (8.90)$$

と定義できる。ラグランジアンを  $C_t, L_t, B_t$  を偏微分し最適化の1階条件

$$\frac{1}{C_t} - \lambda_t = 0, \quad (8.91)$$

$$\frac{W_t \lambda_t}{P_t} - \mu(\gamma + 1) L_t^\gamma = 0, \quad (8.92) \quad \text{eq:NKM\_3}$$

$$\frac{\beta(i_t + 1) \lambda_{t+1}}{P_{t+1}} - \frac{\lambda_t}{P_t} = 0 \quad (8.93)$$

を得る。

### Calvo 型の価格設定モデル

財  $c_{j,t}$  を生産する企業  $j$  の生産関数を

$$c_{j,t} = A_t L_{j,t} \quad (8.94) \quad \text{eq:NKM\_6}$$

とする。 $A_t$  は各企業共通の技術水準で、これを所与とし、企業  $j$  は労働投入  $L_{j,t}$  に比例する生産  $c_{j,t}$  を行う。労働投入1単位に対する名目賃金は  $W_t$  なので、費用最小化行動は

$$\begin{aligned} \min_{L_{j,t}} & \frac{W_t L_{j,t}}{P_t} \\ \text{subject to } & c_{j,t} = A_t L_{j,t} \end{aligned} \quad (8.95)$$

と定式化できる。 $\varphi_t$  をラグランジュ乗数とすると、ラグランジアンは

$$\Phi_t = \frac{W_t L_{j,t}}{P_t} + \varphi_t (c_{j,t} - A_t L_{j,t}) \quad (8.96)$$

と定義出来て、一回の条件はこれを  $L_{j,t}$  で偏微分することで

$$\varphi_t = \frac{W_t}{A_t P_t} \quad (8.97)$$

となる。ここで、生産1単位あたりの名目平均費用は

$$\frac{W_t L_{j,t}}{c_{j,t}} = \frac{W_t}{A_t} \quad (8.98)$$

で表される。 $\varphi_t$  は生産1単位あたりの実質の平均費用を表す。

独占的競争のもとで、企業は他社と異なる財を生産するため価格を自由に決定できるが、每期価格改定できるわけではなく、ある一定の確率で  $t$  期に価格改定できるものとする。価格改定できない確率を  $\nu$ 、価格改定できる確率を  $1 - \nu$  とする。企業は、もし当期に価格改定できるとすると、次に価格改定できるまでの利潤の割引現在価値を最大にするように当期の価格を決定する

と考えられる。次に価格改定できるまでの利潤の  $t$  期における割引現在価値は、当期の企業の販売価格を  $p_{j,t}$  とすると、 $\prod_{k=1}^0 (1 + i_{t+k-1}) = 1$  として

$$\Theta_{j,t} = \sum_{i=0}^{\infty} \frac{\nu^i}{\prod_{k=1}^i (1 + i_{t+k-1})} (p_{j,t} - P_{t+i} \varphi_{t+i}) c_{j,t+i} \quad (8.99)$$

を最大化することによって最大になる。 $p_{j,t}$  が販売価格、 $P_{t+i} \varphi_{t+i}$  が  $t+i$  期における名目平均費用、 $c_{j,t+i}$  が販売数量であるため、 $(p_{j,t} - P_{t+i} \varphi_{t+i}) c_{j,t+i}$  は  $t$  期から  $t+i$  期まで企業が価格を開延できなかった場合の  $t+i$  期の利潤を表す。 $i=1$  のとき、

$$\nu \frac{(p_{j,t} - P_{t+1} \varphi_{t+1})}{1 + i_t} \left( \frac{p_{j,t}}{P_{t+1}} \right)^{-\eta} C_{t+1} + (1 - \nu) \frac{(p_{j,t} - P_{t+1} \varphi_{t+1})}{1 + i_t} \left( \frac{p_{j,t+1}}{P_{t+1}} \right)^{-\eta} C_{t+1} \quad (8.100)$$

が  $t+1$  期の利潤の現在価値の期待値である。第2項は、来期に価格改定できた場合の現在価値を表しており、 $p_{j,t+1}$  は企業が投機に  $p_{j,t}$  をいくらにするかによらず、自由に決めることが出来る。すなわち、第2項は企業が当期に価格  $p_{j,t}$  をいくらに定めるかは無関係であるため、無視することが出来る。 $t+2$  期、 $t+3$  期、 $\dots$  と同様に考えれば、 $\Theta_{j,t}$  を最大化することによって企業の利潤の割引現在価値が最大になる。

$\Theta_{j,t}$  に  $c_{j,t}$  を代入すれば

$$\Theta_{j,t} = \sum_{i=0}^{\infty} \frac{\nu^i P_{t+i}}{\prod_{k=1}^i (1 + i_{t+k-1})} \left( \frac{p_{j,t}}{P_{t+i} - \varphi_{t+i}} \right) \left( \frac{p_{j,t}}{P_{t+i}} \right)^{-\eta} C_{t+i} \quad (8.101)$$

となる。最適化の1階条件から

$$\frac{1}{1 + i_t} \frac{P_{t+1}}{P_t} = \beta \left( \frac{C_t}{C_{t+1}} \right) \quad (8.102)$$

なので（オイラー方程式）、両辺を  $t$  から  $t+i-1$  までかけ合わせることで

$$\frac{1}{\prod_{k=1}^i (1 + i_{t+k-1})} \frac{P_{t+i}}{P_t} = \beta^i \left( \frac{C_i}{C_{t+i}} \right) \iff \frac{P_{t+i}}{\prod_{k=1}^i (1 + i_{t+k-1})} = \beta^i \left( \frac{C_i}{C_{t+i}} \right) P_t \quad (8.103)$$

が得られる。従って、 $\Theta_{j,t}$  は

$$\Theta_{j,t} = P_t C_t \sum_{i=0}^{\infty} \nu^i \beta^i \left\{ \left( \frac{p_{j,t}}{P_{t+i}} \right)^{1-\eta} - \varphi_{t+i} \left( \frac{p_{j,t}}{P_{t+i}} \right)^{-\eta} \right\} \quad (8.104)$$

と書き換えられる。 $\Theta_{j,t}$  を  $p_{j,t}$  で偏微分し1階の条件

$$P_t C_t \sum_{i=0}^{\infty} \nu^i \beta^i \left\{ (1 - \eta) \frac{p_{j,t}^{-\eta}}{P_{t+1}^{1-\eta}} + \eta \varphi_{t+i} \frac{P_{j,t}^{-\eta-1}}{P_{t+i}^{-\eta}} \right\} = 0 \quad (8.105)$$

を得る。この条件は、当期に価格改定ができる企業にとって共通の条件であるため、 $p_{j,t}$  の添え字  $j$  を外して、 $p_t^*$  を最適な価格として整理すると

$$\begin{aligned} \sum_{i=0}^{\infty} \nu^i \beta^i (\eta - 1) \frac{(p_t^*)^{-\eta}}{P_{t+i}^{1-\eta}} &= \sum_{i=0}^{\infty} \nu^i \beta^i \eta \varphi_{t+i} \frac{(p_t^*)^{-\eta-1}}{P_{t+i}^{-\eta}} \\ \iff (\eta - 1) \sum_{i=0}^{\infty} \nu^i \beta^i P_{t+i}^{\eta-1} &= \frac{\eta}{p_t^*} \sum_{i=0}^{\infty} \nu^i \beta^i \varphi_{t+i} P_{t+i}^{\eta} \\ \iff p_t^* \sum_{i=0}^{\infty} \nu^i \beta^i P_{t+i}^{\eta-1} &= \frac{\eta}{\eta - 1} \sum_{i=0}^{\infty} \nu^i \beta^i \varphi_{t+i} P_{t+i}^{\eta} \end{aligned} \quad (8.106)$$

eq: NKM\_2

となり、ここでさらに  $P_t^\eta$  で割れば

$$\begin{aligned} \Leftrightarrow \frac{p_t^*}{P_t} \sum_{i=0}^{\infty} \nu^i \beta^i \left( \frac{P_{t+i}}{P_t} \right)^{\eta-1} &= \frac{\eta}{\eta-1} \sum_{i=0}^{\infty} \nu^i \beta^i \varphi_{t+i} \left( \frac{P_{t+i}}{P_t} \right)^{\eta} \\ \Leftrightarrow \frac{p_t^*}{P_t} &= \frac{\eta}{\eta-1} \frac{\sum_{i=0}^{\infty} \nu^i \beta^i \varphi_{t+i} (P_{t+i}/P_t)^{\eta}}{\sum_{i=0}^{\infty} \nu^i \beta^i (P_{t+i}/P_t)^{\eta-1}} \end{aligned} \quad (8.107) \quad \boxed{\text{eq:NK1}}$$

となる。ここで、

$$F_t := \sum_{i=0}^{\infty} \nu^i \beta^i \varphi_{t+i} \left( \frac{P_{t+i}}{P_t} \right)^{\eta} \quad (8.108)$$

と定義すると

$$\begin{aligned} F_t &= \varphi_t + \sum_{i=1}^{\infty} \nu^i \beta^i \varphi_{t+i} \left( \frac{P_{t+i}}{P_t} \right)^{\eta} \\ &= \varphi_t + \sum_{i=0}^{\infty} \nu^{i+1} \beta^{i+1} \varphi_{t+i+1} \left( \frac{P_{t+i+1}}{P_t} \right)^{\eta} \\ &= \varphi_t + \nu \beta \left( \frac{P_{t+1}}{P_t} \right)^{\eta} \sum_{i=0}^{\infty} \nu^i \beta^i \varphi_{t+i+1} \left( \frac{P_{t+i+1}}{P_{t+1}} \right)^{\eta} \\ &= \varphi_t + \nu \beta \left( \frac{P_{t+1}}{P_t} \right)^{\eta} F_{t+1} \end{aligned} \quad (8.109)$$

という漸化式を得る。同様に

$$D_t := \sum_{i=0}^{\infty} \nu^i \beta^i \left( \frac{P_{t+i}}{P_t} \right)^{\eta-1} \quad (8.110)$$

と定義すると、

$$D_t = 1 + \nu \beta \left( \frac{P_{t+1}}{P_t} \right)^{\eta-1} D_{t+1} \quad (8.111)$$

を得る。よって、<sup>eq:NK1</sup>(8.107) 式は

$$\frac{p_t^*}{P_t} = \frac{\eta}{\eta-1} \frac{F_t}{D_t} \quad (8.112)$$

となる。一方で、一般物価  $P_t$  の式より

$$P_t^{1-\eta} = \int_0^1 p_{j,t}^{1-\eta} dj \quad (8.113)$$

であるが、当期において  $p_t^*$  に価格改定できる企業の割合が  $1-\nu$ 、前期のままである割合が  $\nu$  であること、積分を和の極限で表現できることから

$$\begin{aligned} P_t^{1-\eta} &= \int_0^1 \left\{ (1-\nu)(p_t^*)^{1-\eta} + \nu p_{j,t-1}^{1-\eta} \right\} dj \\ &= (1-\nu) \int_0^1 (p_t^*)^{1-\eta} dj + \nu \int_0^1 p_{j,t-1}^{1-\eta} dj \\ &= (1-\nu)(p_t^*)^{1-\eta} + \nu P_{t-1}^{1-\eta} \end{aligned} \quad (8.114)$$

と表すことが出来る。

### 金融政策ルール

名目利子率が、Taylor ルールに従うものだとする。金融政策ルールとは、インフレ率や経済成長の安定化を目的としてマクロ経済の変動に応じて機械的に金融政策を運営する方式を数式により表現したものである。 $i_t$  を名目利子率、 $i^*$  を名目利子率の定常値、 $\pi_t = \frac{P_t}{P_{t-1}} - 1$  をインフレ率、 $\pi^T$  をターゲットインフレ率、 $Y_t, Y_t^f$  をそれぞれ実質 GDP と景気中立的な GDP の水準とすると、Taylor ルールのもとでは、中央銀行は名目利子率を政策金利として

$$i_t = i^* + \varphi_\pi(\pi_t - \pi^T) + \phi_y \ln \left( \frac{Y_t}{Y_t^f} \right) \quad (8.115) \quad \text{eq:NKM\_5}$$

によって決定する。Taylor ルールの係数は、中央銀行が何を目的関数として金融政策を実行するかによって変わり得る。 $\varphi_\pi \pi_t$  の項は、足元インフレ率の大小に応じて政策金利を定常値  $i^*$  からどの程度引き上げるか、引き下げるかを表している。第3項は

$$\ln \left( \frac{Y_t}{Y_t^f} \right) = \ln \left( 1 + \frac{Y_t - Y_t^f}{Y_t^f} \right) \approx \frac{Y_t - Y_t^f}{Y_t^f} \quad (8.116)$$

で、GDP の景気中立的な水準からの乖離率、つまり GDP ギャップの大小に応じて、政策金利をどの程度調整するかを表している。

ここで、景気中立的な GDP の水準  $Y_t^f$  がどこであるかが問題になるのだが、ここでは仮に価格がフレキシブル、つまり  $\nu = 0$  の場合の GDP の水準を景気中立的な GDP の水準とする。 $\nu = 0$  のとき、<sup>eq:NKM\_2</sup>(8.106)式より、

$$p_t^* P_t^{\eta-1} = \frac{\eta}{\eta-1} \varphi_t P_t^\eta \quad (8.117)$$

が企業にとっての利潤最大化になる。価格がフレキシブルな場合には、每期すべての企業が同じ価格に設定することから  $p_t^* = P_t$  であり、このとき

$$\varphi_t = \varphi^f = \frac{\eta-1}{\eta} \quad (8.118)$$

が利潤最大化条件となる。

$W_t/P_t = A_t \varphi_t$ 、生産関数より  $L_t = C_t/A_t$  なので、これらを <sup>eq:NKM\_3</sup>(8.92)式へ代入すれば

$$\frac{A_t \varphi_t}{C_t} - \mu(\gamma+1) \left( \frac{C_t}{A_t} \right)^\gamma = 0 \quad (8.119) \quad \text{eq:NKM\_4}$$

という消費、技術水準と平均費用の関係を表す式が得られる。 $\varphi_t = \varphi^f$  のときの  $C_t$  を  $C_t^f$  と書くことにすると、<sup>eq:NKM\_4</sup>(8.119)より

$$C_t^f = A_t \left( \frac{\varphi^f}{\mu(\gamma+1)} \right)^{\frac{1}{\gamma+1}} = A_t C^* \quad (8.120)$$

を得る。ここで、 $C^*$  は  $C_t$  の定常値とする<sup>5</sup>。消費財のみ存在するこのモデルでは、GDP は消費に等しい（つまり  $Y_t = C_t$ ）ため、この  $C_t^f$  が価格が伸縮的な場合の GDP の水準となる。

<sup>5</sup>後述



(8.115)式<sup>eq:NKM\_5</sup>の  $Y_t, Y_t^f$  をそれぞれ、 $C_t, C_t^f = A_t C^*$  で置き換え、さらに金融政策ルールの誤差項  $v_t$  を付加した

$$i_t = i^* + \varphi_\pi(\pi_t - \pi^T) + \phi_y \ln\left(\frac{C_t}{A_t C^*}\right) + v_t \quad (8.121) \quad \text{eq:NKM_6}$$

をモデル式として用いることにする<sup>6</sup>。金融政策ルールの誤差項は、金融政策ショック  $z_t$  を外生変数として AR(1) 過程

$$v_{t+1} = \rho_v v_t + z_{t+1} \quad (8.122) \quad \text{eq:NKM_7}$$

に従うものとする。

マクロ生産関数は、(8.121)式<sup>eq:NKM\_6</sup>より、

$$C_t = A_t L_t \quad (8.123)$$

である。技術水準  $A_t$  は技術ショック  $\varepsilon_t$  を外生変数とし、

$$\ln(A_{t+1}) = \rho_A \ln(A_t) + \varepsilon_{t+1} \quad (8.124) \quad \text{eq:NKM_8}$$

という過程に従うものとする。

### モデルの全体と定常値

CIA model の時と同じく、 $P_t$  はそのままでは内生変数として取り扱えないため、先に定義したインフレ率  $\pi_t$  をモデル変数とする。オイラー方程式を  $\pi_t$  を使って表現すると

$$(1 + \pi_{t+1}) \frac{C_{t+1}}{C_t} = \beta(1 + i_t) \quad (8.125)$$

である。同じ理由で  $p_t^*$  についても  $\pi_t^* := \frac{p_t^*}{P_{t-1}} - 1$  をモデル変数とする。インフレ率を決める式は、

$$\frac{1 + \pi_t^*}{1 + \pi_t} = \frac{\eta}{\eta - 1} \frac{F_t}{D_t} \quad (8.126)$$

と  $F_t, D_t$  の漸化式

$$F_t = \varphi_t + \nu \beta (1 + \pi_{t+1})^\eta F_{t+1}, \quad (8.127)$$

$$D_t = 1 + \nu \beta (1 + \pi_{t+1})^\eta D_t \quad (8.128)$$

と  $P_t^{1-\eta}$  の変数を置き換えた

$$(1 + \pi_t)^{1-\eta} = (1 - \nu)(1 + \pi_t^*)^{1-\eta} + \nu \quad (8.129)$$

の4本である。名目利子率を決める Taylor ルールとその誤差項の式は (8.121)式<sup>eq:NKM\_6</sup>、(8.122)式<sup>eq:NKM\_7</sup>。技術の水準の遷移式は (8.124)式<sup>eq:NKM\_8</sup>であり、消費、技術水準と平均費用の関係式を変形すれば

$$\varphi_t = \mu(\gamma + 1) \left(\frac{C_t}{A_t}\right)^{\gamma+1} \quad (8.130)$$

<sup>6</sup>以下、簡単のため  $\pi^T = 0$  とする。

である。この New Keynesean model は上述の 9 本の方程式からなる。

定常状態における内生変数それぞれの値（定常値）は、 $C_t = C_{t+1} = C^*$  として置き換えることによって得られる 9 本の連立方程式を解くことによって得られ、具体的には

$$\pi^* = \pi^o = v^* = 0, i^* = \frac{1-\beta}{\beta}, D^* = \frac{1}{1-\nu\beta}, F^* = \frac{\eta-1}{\eta(1-\nu\beta)}, \varphi^* = \frac{\eta-1}{\eta}, A^* = 1, C^* = \left( \frac{\eta-1}{\mu(\gamma+1)\eta} \right)^{\frac{1}{\gamma+1}}$$

である。

## 8.5 金利の期待仮説とイールドカーブ

金融市場では、さまざまな満期の債券が取引されている。これら異なる満期の再建金利の関係をみるときに使われるのが、イールドカーブ（利回り曲線）である。これは、横軸に満期、縦軸に金利をとり、角時点における金利と満期の関係をグラフ化したものになる。そして、イールドカーブは、通常の場合、満期が長くなるにつれて金利水準が上昇する右上がりの形状となる。このため、右上がりのイールドカーブの形状を**順イールド**、右下がりの形状を**逆イールド**とよぶ。

### 8.5.1 金利の期待仮説

イールドカーブの形状を説明する基本的な考え方は、**金利の期待仮説**と呼ばれ、異なる満期の金利はすべて、市場での取引を通じて価格裁定が働いていると考える。

**例 8.1.** 2つの市場 A,B が存在しており、市場 A と市場 B を比較すると、市場 A は資金需要が旺盛で、市場金利の水準が高い状況  $i_A(0) < i_B(0)$  からスタートする。

ある投資家が、市場 A で資金を調達し、市場 B で資金を供給することを考える。市場 B における供給金利と市場 A での調達金利の差（＝利ざや）を利益として獲得する。この結果、市場 A の金利は上昇、市場 B の金利は低下し、最終的に市場金利が一致（ $i_A(T) = i_B(T)$ ）するまでこの行動が継続する。こうした投資家の行動を**金利裁定**という。

金利の期待仮説においては、こういった金利裁定が、次の例にあげるような異なる満期間で行われると考える。

**例 8.2.** 市場に2年先までの取引が存在するとし、現時点  $t$  において取引される金利は1年物金利  $i_1(t)$  と、2年物金利  $i_2(t)$  であるとする。また、現時点では取引されないが、1年後からの1年物金利の予想を  $i_1(t+1)$  で表す<sup>7</sup>。

ここで、2つの投資行動を考える。投資行動 A では、現時点  $t$  において2年物金利で運用することを考える。投資行動 B では、現時点  $t$  において1年もの金利で運用し、1年後に再度その1年もの金利に再投資することを考える。このとき、A,B によって2年後の成果に差があれば、裁定機会となってしまいうため、A,B の成果は等しくなる。

### 8.5.2 タームプレミアム

イールドカーブについて、通常、満期が長くなるにつれて金利が上昇する右上がりの形状（順イールド）をしていることを説明するためには、金利の期待仮説に加えて、満期が長くなるにつれて、より長期間にわたって資金を固定することへのリスクプレミアムである**タームプレミアム**を考える必要がある。

仮に将来の政策金利が一定であると予想されている場合、金利の期待仮説に基づくと、イールドカーブは水平となるが、これにタームプレミアムを加味すれば、イールドカーブは右上がりとなる。

<sup>7</sup>復習：現時点で取引されている金利をスポット金利、1年先からスタートする取引の金利をフォワード金利と呼ぶのであった

例えば、経済が均衡状態にあり、将来にわたって均衡状態が継続される可能性が高いと予想されているとすると、このとき、将来にわたる短期金利の予想経路は、均衡水準の短期金利で一定になると考えられる。実際、観察されるイールドカーブは、この期待仮説に基づくイールドカーブにタームプレミアムを可算したものになる。上述の通り、タームプレミアムは満期が長くなるにつれて上昇するため、イールドカーブは右上がりの形状となる。

### 8.5.3 フォワードレートカーブ

通常、イールドカーブは、各満期時点におけるスポットレートを使って描くことが多い。しかし、金融政策分析を行う際は、イールドカーブを、フォワードレートを用いて書くことが有効となる。これは、スポットレートから逆算した将来のフォワードレートは、市場が予想する政策金利の将来経路に関する情報を包含しているためである。ただし、市場で観察されるスポットレートから逆算されたフォワードレートは、タームプレミアムを含めた金利となっており、完全なリスクフリーレートの予測値となっているわけではない。

## 8.6 イールドカーブの変動メカニズム

異なる満期の金利の関係はどのように決定されるのか？この議論を始めるにあたって必要となる、「短期金利・長期金利・物価上昇率の関係」について確認する。

### 8.6.1 長期金利とインフレ率

金融政策 p.120

### 8.6.2 Nelson-Siegel モデル

長期金利が満期時点までの予想短期金利の平均である関係をもとに、満期を連続変数で考えると、満期  $m$  のスポットレート・フォワードレートをそれぞれ  $r(m)$ ,  $f(m)$  と表すと、

$$r(m) = \frac{1}{m} \int_0^m f(s) ds$$

という関係を得る。すなわち、フォワードレートについて、満期を説明変数として関数化すれば、この関数を 0 から  $m$  まで積分し、 $m$  で割ることによってスポットレートが得られることになる。これは、満期  $m$  時点までのフォワードレートカーブの下側の面積を計算し、それを  $m$  で序すことで平均的な金利水準を求めるということを連続の世界でやっていることに他ならない。

Nelson-Siegel モデルでは、フォワードレートの関数系として次式のような式を考える：

$$f(m) = L + Se^{-\lambda m} + C\lambda me^{-\lambda m}$$

ここで、 $L$  はフォワードレートの長期的な水準 (level)、 $S$  は傾き (slope)、 $C$  は曲率 (curvature) を捉えるパラメータで、 $\lambda$  は長期水準への終息の速さをコントロールするパラメータである<sup>8</sup>。

<sup>8</sup>一般に、複数の満期の金利データに対して、主成分分析をすると、水準・傾き・曲率の3つの要素に対応する3つの主成分が優位に検出される。その意味で、この Nelson-Siegel モデルは、比較的勘弁な関数系でイールドカーブの変動を捉えることが出来るため、金融政策分析において利用されている。

上の2式より、スポットレート関数系は

$$r(m) = L + S \frac{(1 - e^{-\lambda m})}{\lambda m} + C \left( \frac{1 - e^{-\lambda m}}{\lambda m} - e^{-\lambda m} \right)$$

で表される。

$m \downarrow 0, m \uparrow +\infty$  を計算すれば、 $r(0) = f(0) = L + S, r(+\infty) = f(+\infty) = L$  という結果が得られ、スポットレート・フォワードレートの満期が0および $+\infty$ に近づくとき、極限值がそれぞれ  $L + S, L$  となり、関数のパラメータによって表現できることになる。

**問題 8.5.** Nelson-Siegel モデルを用いて、順イールド・逆イールドの時のイールドカーブ（スポットレートとフォワードレートが併記されたもの）を作成せよ。

順： $L = 3, S = -1, C = 0$  逆： $L = 3, S = 1, C = 0$  として書いてみればよい。

## 8.7 政策金利の指針

中央銀行は、金融政策の運営にあたり、イールドカーブの中で最も短い満期の翌日物金利を操作することで、金融政策を運営している。例えば、日本銀行は、無担保コールレートという短期資金市場の金利、アメリカ連邦準備制度 (FRB) はフェデラルファンドレート (FF レート) という短期資金市場金利を操作している。欧州中央銀行 (ECB) の政策金利はリファイナンスレートと呼ばれる ECB の貸出金利だが、金利市場調整は ESTR (European Short-term Rate) という短期資金市場金利を安定化するように運営されている。

### 8.7.1 Taylor ルール

政策金利の誘導水準に関する指針となる金融政策ルールのうち、Taylor ルールについて説明する。

Taylor ルールは、John Taylor にいより提唱された金融政策ルールで、政策金利をインフレ率や産出量ギャップといったマクロ経済変数の変動にどう反応させるかを表した関係式で表現される。

#### 基本形

Taylor ルールは、政策金利の誘導水準をインフレ率と産出量ギャップに結びつける政策反応関数として定義される。この Taylor ルールは、New Keynesian model を使った金融政策分析の中で広く用いられており、実践的な政策指針ではあるものの、一切の裁量を排除する厳格なルールではない。Taylor ルールからの逸脱にも意味がある場合や、金融政策の将来経路に関するコミュニケーションの基盤としての役割もはたしている。

Taylor ルールの基本形は

$$i = r^* + \pi^* + \alpha(\pi - \pi^*) + \beta(y - y^*)$$

で表現される。ここで、 $i$  は名目政策金利、 $r^*$  は均衡実質金利、 $\pi$  はインフレ率、 $\pi^*$  はインフレ目標、 $\pi$  はインフレ率、 $y$  は対数変換された産出量、 $y^*$  は対数変換された潜在産出量である。この式を少し書き換えると、

$$i = r^* + \pi + (\alpha - 1)(\pi - \pi^*) + \beta(y - y^*)$$

の形になる。政策金利は、最も満期の短い翌日物金利であるため、現在のインフレ率を控除したものを実質政策金利と考えることができる。このため、2つめの式の右辺にある  $\pi$  を移項し

$$i - \pi = r^* + (\alpha - 1)(\pi - \pi^*) + \beta(y - y^*)$$

と書き換えれば、実質政策金利の水準を示す式としても解釈できる。

Taylor ルールにいて望ましい政策金利の水準を決める要因は大きく3つ

- (1) **実際に観察されるインフレ率とその目標水準の差** インフレ率  $>$  インフレ目標ならば、政策金利を引き上げ、インフレ率  $<$  インフレ目標ならば、政策金利を引き下げる
- (2) **GDP ギャップ**<sup>9</sup> GDP ギャップがプラス方向に大きくなっていくと景気が改善から加熱に向かっており、マイナス方向に動くと景気が後退していくことを意味する。
- (3) **定数項** インフレ率が目標水準と一致し、GDP ギャップが0であるときの政策金利の水準（つまり、経済が均衡状態にあるときの政策金利）。この定数項は、名目金利と実質金利をインフレ予想に分解する Fischer 方程式を使い、均衡実質金利と長期インフレ率の2つに分解することが出来る。均衡実質金利は、長期的な潜在成長率と関係しており、経済成長トレンドが高ければ、均衡実質金利の水準も高くなる。長期インフレトレンド予想は、金融政策に対する信認が確立されていれば、中央銀行の設定するインフレ目標近くにアンカリングされている。このため、定数項は、実質金利とインフレ目標の和であると想定される。

Taylor ルールにおいて重要なポイントは、マクロ経済の安定化のために、インフレ率にかかる係数  $\alpha$  が1よりも大きな値をとる必要がある点であり、これを Taylor 原理という。

簡単のため、GDP ギャップが不変であるとする。このとき、 $\alpha \geq 1$  であれば、インフレ率が1%上昇すると、政策金利が1%以上引き上げられることになる。結果として、実質政策金利（名目政策金利からインフレ率を差し引いたもの）も必ず上昇することになる。

金利のマクロ経済への影響を関がえるにあたっては、実質金利（名目金利からインフレ率を除いたもの）を考慮する必要がある。このため、同じ水準の名目金利であっても、インフレ率いかんによって、借入時の実質的な負担が異なることになる。

$$\text{実質金利} = \text{名目金利} - \text{インフレ率}$$

このことを考えると、景気を抑制してインフレ率を低下させるためには、実質金利を上昇させる必要があることになる。逆に、景気を刺激してインフレ率を上昇させるためには、実質金利を引き下げる必要がある。名目・実質の両方を同じ方向に動かすには、インフレ率の変動以上の名目金利の変動が必要になる。これがテイラー原理が意味することになる。

**問題 8.6.** GDP ギャップを無視して、 $\pi^*, r^*$  を適切において、 $\alpha = 0.751.5$  の2つのパターンで政策金利とインフレ率の関係をグラフにプロットせよ。

**注意 18.** ELB というのがある。

日本における Taylor ルール適合性

## 第9章 数値計算法の基礎

### 9.1 乱数生成

#### 9.1.1 \*一様乱数とメルセンヌツイスター

乱数生成法に関して厳密な定義を与えることは難しく、「 $[0, 1]$  区間の中から」一様かつ独立にデタラメな数を発生させる方法」としておく。これにより得られた数列を乱数列と呼ぶ。サイコロは  $\{1, 2, 3, 4, 5, 6\}$  の中からデタラメな数を発生させる乱数生成法であるし、コイントスはコインが表の時を 1、裏の時を  $-1$  と対応させれば  $\{1, -1\}$  の中からデタラメな数を発生させる乱数発生法である。後に述べる、モンテカルロ法は確率的現象を含んだあらゆる現象のシミュレーションに用いるが、ここで乱数が必要になる。

決定的な動作しかできない計算機では、「真にデタラメ」な数を発生することが出来ない。レイコロやコイントスなどから乱数を発生することは出来るものの、コストの観点・スピードの観点から現実的ではなく、何よりも再現不能であるという問題から追試・検証に使うことが出来ない。

そこで、計算機において疑似的な乱数を生成する方法が様々考えられた。一つは線形合同法 (Linear Congruential Generator; LCG) である。これは、ある整数  $x_0$  をシードとして選び、漸化式によって  $x_1, x_2$  を次々に生成するものである。例えば、漸化式を

$$x_n = 1103515245x_{n-1} + 12345 \pmod{2^{32}}.$$

によって定めると、 $x_0 = 3$  のとき

$$\begin{aligned} x_1 &= 3,310,558,080 & (3 \times 1103515245 + 12345 = 3310558080) \\ x_2 &= 465,823,161 & (3310558080 \times 1103515245 + 12345 = 3653251310737941945 \equiv 465823161) \\ x_3 &= 679,304,702 & (465823161 \times 1103515245 + 12345 = 514042959637601790 \equiv 679304702) \\ &\vdots & \vdots \end{aligned}$$

のように得られる。

このような手法によって得られる疑似乱数のメリットは、漸化式と初期シードを記録しておけば、誰でも同じ数列を高速かつ低コストで再現できる。しかし、このような数列を乱数と呼んでよいのかという疑問が生じる。実際、von Neumann は “*Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin.*” と述べている。

例えば、先の線形合同法による数列の周期は、シードによらず  $2^{32}$  となる。1970 年～80 年代にかけて ANSI-C などの標準疑似乱数にも用いられていたが、現代では、 $2^{32} \approx 43$  億個の乱数などすぐに使い切ってしまう。また、生成される数列は、局所的には乱数のようなふるまいに見えるが、数千万個もの出力を眺めると、非乱数性 ( $\approx$  規則性) が分かってしまう。

そこで、疑似乱数に対して求められるスペックについて考えて見ると

- 出力に当たって高速であること

- 乱数性

疑似乱数は、所詮漸化式による数列にすぎない。「乱数性」について、実証可能な定義がないことから、「乱数であること」を示すことができない。そこで、

- 周期
- 分布

に着目して、出来るだけ良いものを使おうという発想に行きつく。

周期の観点から、線形合同法を見つめなおすと

$$x_n = \alpha x_{n-1} + C \pmod{M}$$

という数列の周期は高々  $M$  であり、乱数のスペックを周期・分布で捉える上では分かり易いものの、周期を大きくするには  $M$  を大きくするしかなく、大きなケタによる掛け算や割り算が必要となるためスピードを落としかねないというデメリットが存在する。

そこで、松本 眞と西村 拓士はメルセンヌツイスターという手法を開発した<sup>1</sup>。この手法は数学における、 $\mathbb{F}_2[t]$  ( $\mathbb{F}_2 = \{0, 1\}$  は二元体) 多項式を用いて、周期  $2^{607} - 1$  (187 ケタ) で、607 次元菌等分布性という性質を持つという点で、疑似乱数としてかなり良い性質を持つものである。

### 9.1.2 様々な確率分布に従う確率変数を生成する

一様乱数を用いて、与えられた確率分布を持つ確率変数を生成する方法を述べる。

#### 逆関数法

逆関数法は、一様分布に従う確率変数  $U$  から密度関数  $f(x)$  を持つ確率変数  $X$  を作る方法である。 $X$  の分布関数を  $F(x) = P(X \leq x)$  とする。一般には分布関数は単調非減少関数であるが、特に単調増加を仮定する。すると、 $F$  の逆関数が存在し、 $y = F^{-1}(x)$  とする。

**定理 9.1.**  $X = F^{-1}(U)$  とすると、 $X$  の分布関数は  $F(x)$  となる。

*Proof.*  $\mathbf{P}(X \leq x) = \mathbf{P}(F^{-1}(U) \leq x) = \mathbf{P}(U \leq F(x)) = F(x).$  □

**例 9.1** (一様乱数から指数分布に従う乱数の生成).  $y = 1 - e^{-\lambda x} \iff x = -\frac{\ln(1-y)}{\lambda}$  より、

$$-\frac{\ln(1-U)}{\lambda} \sim -\frac{\ln(U)}{\lambda} \sim \text{Exp}(\lambda).$$

離散型の確率分布についても、確率分布をもとに区間を指定することで生成することができる。

**例 9.2** (離散型の逆関数法).  $\mathbf{P}(X = x_1) = p_1, \mathbf{P}(X = x_2) = p_2, \mathbf{P}(X = x_3) = p_3, p_1 + p_2 + p_3 = 1$  となる確率分布を作りたければ

$$\begin{aligned} U \in [0, p_1] &\longrightarrow X = x_1, \\ U \in (p_1, p_1 + p_2] &\longrightarrow X = x_2, \\ U \in (p_1 + p_2, 1] &\longrightarrow X = x_3 \end{aligned}$$

のように  $X$  の値を定めればよい。

<sup>1</sup>この疑似乱数生成法のエンジンは松本眞氏のホームページなどで公開されている。



## 棄却法

## 合成法

## 9.1.3 Cholesky 分解

$n$  コのデータ系列があったとき、 $i$  番目と  $j$  番目の相関係数  $\rho_{ij}$  で表す。この相関係数  $\rho_{ij}$  を  $(i, j)$  成分とする  $n \times n$  対称行列を**相関行列**という。この相関行列をインプットとして、与えられた相関を満たすような  $n$  この乱数を生成する方法を考えよう。

ここで使われるのは、線形代数における Cholesky 分解である。

**定義 9.1** (Cholesky 分解). 正定値 Hermite 行列  $A$  に対して、下三角行列  $L$  が存在して、

$$A = LL^*$$

が成り立つ。ここで、 $L^*$  は  $L$  の共役転置を指す。これを  $A$  の **Cholesky 分解**という。

平均0、分散が  $\sigma^2$  である  $n$  コの独立な乱数  $X_1, X_2, \dots, X_n$  を並べたベクトルを  $X = (X_1 X_2 \dots X_n)^\top$  とする。この  $X_1, X_2, \dots, X_n$  の相関行列を  $R$  で表す。 $R$  は上の相関係数  $\rho_{ij}$  によって

$$R = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & & & \\ \vdots & & \ddots & \vdots \\ \rho_{n1} & & \cdots & \rho_{nn} \end{pmatrix} \quad (9.1)$$

と表される。我々が考えたいのは、 $X_1, X_2, \dots$  などと同じ分散を持っていて相関行列が  $R$  となるような  $n$  コの乱数  $Z_1, Z_2, \dots, Z_n$  を得る方法である。

$X_1, X_2, \dots$  は独立なので、任意の  $i, j \in \{1, 2, \dots, n\}$  について共分散  $\text{Cov}(X_i, X_j) = 0$ ,  $\text{Corr}(X_i, X_j) = 0$  である。これらを用いて新たな  $Z_1, Z_2, \dots$  を逐次構成していく方法を考えるが、天下りの以下のようにする：

$$\begin{cases} Z_1 = a_{11}X_1, \\ Z_2 = a_{21}X_1 + a_{22}X_2, \\ \vdots \\ Z_n = a_{n1}X_1 + a_{n2}X_2 + \cdots + a_{nn}X_n \end{cases} \quad (9.2)$$

これを行列で書くと

$$Z = \underbrace{\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ a_{n1} & a_{n2} & & \cdots & a_{nn} \end{pmatrix}}_{=:A} X = AX \quad (9.3)$$

として、下半三角行列  $A$  を用いた表現を得る。 $Z_i$  と  $Z_j$  について、

$$\begin{aligned}
 \text{Corr}(Z_i, Z_j) &= \frac{\text{Cov}(Z_i, Z_j)}{\sqrt{\text{Var}[Z_i]\text{Var}[Z_j]}} \\
 &= \frac{1}{\sigma^2} \text{Cov}(Z_i, Z_j) \\
 &= \frac{1}{\sigma^2} \text{Cov}\left(\sum_{k=1}^i a_{ik} X_k, \sum_{k=1}^j a_{jk} X_k\right) \\
 &= \frac{1}{\sigma^2} \left( \sum_{k=1}^{\min\{i,j\}} a_{ik} a_{jk} \right) \sigma^2 \\
 &= \sum_{k=1}^{\min\{i,j\}} a_{ik} a_{jk}
 \end{aligned}$$

となる。そのため、 $R$  は

$$R = \begin{pmatrix} a_{11}a_{11} & a_{21}a_{11} & \cdots & a_{n1}a_{11} \\ a_{21}a_{11} & a_{21}a_{21} + a_{22}a_{22} & & \\ \vdots & & \ddots & \vdots \\ a_{n1}a_{11} & a_{n1}a_{21} + a_{n2}a_{22} & \cdots & a_{n1}a_{n1} + a_{n2}a_{n2} + \cdots + a_{nn}a_{nn} \end{pmatrix} = AA^\top \quad (9.4)$$

という表現を持つことがわかり、 $R$  の Cholesky 分解を与えていることがわかる<sup>2</sup>。 $R$  について  $\rho_{ij}$  を用いた表現、 $a_{ij}$  を用いた表現を得たので、 $R$  の Cholesky 分解を得るアルゴリズムは  $a_{ij}$  について解けばよく、

$$a_{ij} = \begin{cases} \sqrt{\rho_{ii} - \sum_{k=1}^{i-1} a_{ik}^2}, & i = j \\ \frac{1}{a_{jj}} \left( a_{ij} - \sum_{k=1}^{j-1} a_{ik} a_{jk} \right), & i > j \end{cases} \quad (9.5)$$

で得られる<sup>3</sup>。簡単の為、 $n = 2$  の場合を考えると

$$R = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} = \begin{pmatrix} a_{11}a_{11} & a_{21}a_{11} \\ a_{21}a_{11} & a_{21}a_{21} + a_{22}a_{22} \end{pmatrix}$$

となり、 $\rho_{11} = \rho_{22} = 1$  なので

$$\begin{cases} a_{11}^2 &= 1, \\ a_{21}a_{11} &= \rho_{21}, \\ a_{21}^2 + a_{22}^2 &= 1 \end{cases}$$

となり、任意の  $i$  について  $a_{ii} > 0$  としておくと、

$$\begin{cases} a_{11} &= 1, \\ a_{21} &= \rho_{12}, \\ a_{22} &= \sqrt{1 - \rho_{12}^2} \end{cases} \quad (9.6)$$

<sup>2</sup> $R$  は相関行列のため、定義より実数値の正定値な行列であるから Cholesky 分解可能である。

<sup>3</sup>ちなみに、下三角行列の各成分を計算するにあたって、左上端より行ごとに計算するこのアルゴリズムを Cholesky–Banachiewicz 法という。他方、左上端より列ごとに計算するアルゴリズムを Cholesky–Crout 法という。

を得る。そのため、

$$Z_1 = X_1, Z_2 = \rho_{12}X_1 + \sqrt{1 - \rho_{12}^2}X_2$$

とすれば良いことがわかる。

- 演習 9.1.**
1.  $n \times n$  行列 ( $n$  は任意の自然数) をインプットした時に、Cholesky 分解を返す関数を作成せよ。
  2. インプットした行列が正方行列でない場合にエラーをハンドルできるように修正せよ。
  3. インプットした行列が対称行列でない場合にエラーをハンドルできるように修正せよ。
  4. 作成した関数を用いて、標準正規分布に従う独立した乱数 (個数は自由) に対し相関を与えよ。

## 9.2 モンテカルロ法

### 9.3 2項ツリー



## 第10章 時系列解析（発展）

### 10.1 非定常な時系列

時系列解析（基礎）では、定常な場合を中心に議論を進めてきたが、非定常なものに出会うことももちろんある。そこで、非定常の時系列を扱ういくつかの直感的なアイデアを示す。

#### 10.1.1 分散における非定常

時系列の平均レベルが決定的に変化し、時系列の分散が平均レベルに沿って変化する状況:

$$y_t = \mu_t + z_t, \text{Var}[y_t] = \text{Var}[z_t] = h^2(\mu_t)\sigma^2$$

を考える。ここで、 $\mu_t$  は時間に応じて定められた値を返す物である。この式は、 $y_t$  の分散が平均  $\mu_t$  に比例することを表している。このような現象を説明するべく、 $g(Y_t)$  の分散が一定となるような変換  $g$  を見つけたい。

Taylor 展開を用いれば、

$$g(y_t) \approx g(\mu_t) + (y_t - \mu_t)g'(\mu_t)$$

を得る。従って、

$$\text{Var}[g(y_t)] = \text{Var}[g(\mu_t) + (y_t - \mu_t)g'(\mu_t)] = (g'(\mu_t))^2 \text{Var}[y_t] = (g'(\mu_t))^2 h^2(\mu_t) \sigma^2$$

となる。 $g(y_t)$  の分散が定数となるためには、[1]  $h(\mu_t) = \mu_t$  であれば、 $g'(\mu_t) = 1/h(\mu_t) = 1/\mu_t$  すなわち  $g(\mu_t) = \ln(\mu_t)$  であれば良く、これは通常対数変換 [2]  $h(\mu_t) = \sqrt{\mu_t}$  であれば、 $g'(\mu_t) = 1/\sqrt{\mu_t}$  となる。これは、 $g(\mu_t) = 2\sqrt{\mu_t}$  を意味しており平方根過程である。

#### 10.1.2 heikinn

### 10.2 EGARCH モデルとボラティリティ変動モデル



## 第11章 金利モデルの基礎

### 11.0 基本的用語の復習

#### 1. 複利

- (a) 単利：1 円の  $T$  年後の価値  $= 1 + rT$
- (b) 1 年複利 (p.a.)：1 円の  $T$  年後の価値  $= (1 + r)^T$
- (c) 半年単利 (s.a.)：1 円の  $T$  年後の価値  $= (1 + \frac{1}{2}r)^{2T}$
- (d)  $t$  年複利：1 円の  $T$  年後の価値  $= (1 + rt)^{T/t}$
- (e) 連続複利：1 円の  $T$  年後の価値  $= \lim_{n \rightarrow \infty} \left(1 + \frac{r}{n}\right)^{nT} = e^{rT}$

#### 2. daycount fraction

- (a) 30/365  
1 か月を 30 日、1 年間を  $30 \times 12 = 360$  日と考えて計算する
- (b) ACT/ACT  
1 年間を 365 日（閏年は 366 日）として、当該期間を実日数で計算する
- (c) ACT/365  
1 年間を 365 日として、当該期間を実日数で計算する
- (d) ACT/360  
1 年間を 360 日として、当該期間を実日数で計算する

#### 3. ゼロレート・スワップレート

- (a) ゼロレート：当該期間  $T$  年について、期中の利払いはなく、期末にのみ利払いがあるときの金利
- (b) スワップレート：当該期間  $T$  年に  $r$  年ごとに等しい金利で利払いがある場合の金利

#### 4. スポットレート・フォワードレート

- (a) spot rate: 現時点から将来のある時点までの期間に対する金利
- (b) forward rate: 将来のある時点から、さらに先の別の時点までの期間に対応する金利

### 11.1 イールドカーブの推定

任意の期間  $t$  に対応するディスカウント・ファクター（或いはスポットゼロレート）がわかれば、どのような割引債も評価可能となるので、連続的な  $t$  に対するディスカウントファクター  $DF(t)$  を推定しておくことは重要である。

### 11.1.1 スプライン補間

時点  $t_0, t_1, \dots, t_n$  における関数  $f(x)$  の値、 $f(t_0), f(t_1), \dots, f(t_n)$  が既知で、 $f(t_i) = y_i$  と与えられているとき、 $t \in (t_i, t_{i+1}), i \in \{0, 1, 2, \dots, n-1\}$  をどのような値として補間するか？その補間手法についていくつか解説する。

#### 0 次スプライン（階段関数による補間）

0 次スプラインは、以下のように節点間の値を以下のように定める手法をいう：

$$f(t) := \begin{cases} y_i, & t \in [t_i, t_{i+1}), \\ y_n, & t = t_n. \end{cases} \quad (11.1)$$

#### 1 次スプライン（線形補間）

1 次スプラインは、所謂線形補間のことで、節点は標本点に一致する：

$$f(t) = \frac{y_{i-1}(t_i - t) + y_i(t - t_{i-1})}{t_i - t_{i-1}}, \quad t \in [t_{i-1}, t_i], i = 0, 1, 2, \dots, n-1. \quad (11.2)$$

#### 3 次スプライン

ここまでに掲げた 2 つの補間手法は、節点間の値を与えることは出来ているが、滑らかさが保証されないという問題がある。そこで、

- 節点は標本点に一致し
- 十分に滑らかで 2 回までの微分可能性を保証する

手法を考える。2 階微分可能性を担保すべく  $f$  は 3 次式で表現されるものとする。

$$M_i := f''(t_i), \quad i = 0, 1, 2, \dots, n$$

とおく。また、 $\Delta_i := t_i - t_{i-1}$  とおく。 $f''(t)$  は 1 次式であるから

$$f''(t) = M_{i-1} \frac{t_i - t}{\Delta_i} + M_i \frac{t - t_{i-1}}{\Delta_i}, \quad t \in [t_{i-1}, t_i] \quad (11.3)$$

となり、これを 2 回積分し、積分定数を  $f(t_{i-1}) = y_{i-1}k, f(t_i) = y_i$  により定めると

$$f(t) = M_{i-1} \left( \frac{(t_i - t)^3}{6\Delta_i} - \frac{t_i - t}{6} \Delta_i \right) + M_i \left( \frac{(t - t_{i-1})^3}{6\Delta_i} - \frac{t - t_{i-1}}{6} \Delta_i \right) + \frac{t_i - t}{\Delta_i} y_{i-1} + \frac{t - t_{i-1}}{\Delta_i} y_i \quad (11.4)$$

を得る。これを 1 回微分すれば

$$f'(t) = M_{i-1} \left( -\frac{(t_i - t)^2}{2\Delta_i} + \frac{\Delta_i}{6} \right) + M_i \left( \frac{(t - t_{i-1})^2}{2\Delta_i} - \frac{\Delta_i}{6} \right) + \frac{y_i - y_{i-1}}{\Delta_i} \quad (11.5)$$



を得る。さて、ここから

$$\begin{aligned} f'(t_{i,-}) &= \frac{\Delta_i}{6} M_{i-1} + \frac{\Delta_i}{3} M_i + \frac{y_i - y_{i-1}}{\Delta_i}, \\ f'(t_{i,+}) &= -\frac{\Delta_{i+1}}{3} M_i - \frac{\Delta_{i+1}}{6} M_{i+1} + \frac{y_{i+1} - y_i}{\Delta_{i+1}} \end{aligned}$$

となるが、 $f'$  が連続であることから  $f'(t_{i,-}) = f'(t_{i,+})$  でなくてはならず、 $\kappa_i := \frac{\Delta_{i+1}}{\Delta_i + \Delta_{i+1}}$ ,  $\tilde{\kappa}_i := 1 - \kappa_i$  とおくと

$$\tilde{\kappa}_i M_{i-1} + 2M_i + \kappa_i M_{i+1} = 6 \frac{1}{\Delta_i + \Delta_{i+1}} \left( \frac{y_{i+1} - y_i}{\Delta_{i+1}} - \frac{y_i - y_{i-1}}{\Delta_i} \right) =: \delta_i, \quad i = 1, 2, \dots, n-1$$

を得る。これだけでは、変数  $n+1$  コ、方程式が  $n-1$  本なので、一意に解を定められないが、追加的に

$$2M_0 + \kappa_0 M_1 = \delta_0, \quad \tilde{\kappa}_n M_{n-1} + 2M_n = \delta_n$$

という条件を課しても一般性は失わず、これを行列の形で書けば、以下のようになる。

$$\begin{pmatrix} 2 & \kappa_0 & 0 & \cdots & \cdots & 0 \\ \tilde{\kappa}_1 & 2 & \kappa_1 & 0 & \cdots & 0 \\ 0 & \tilde{\kappa}_2 & 2 & \kappa_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & 2 & \kappa_{n-1} & 0 \\ & & & & \tilde{\kappa}_{n-1} & 2 & \kappa_{n-1} \\ 0 & \cdots & & & \tilde{\kappa}_n & 2 \end{pmatrix} \begin{pmatrix} M_0 \\ M_1 \\ M_2 \\ \vdots \\ M_{n-2} \\ M_{n-1} \\ M_n \end{pmatrix} = \begin{pmatrix} \delta_0 \\ \delta_1 \\ \delta_2 \\ \vdots \\ \delta_{n-2} \\ \delta_{n-1} \\ \delta_n \end{pmatrix}. \quad (11.6)$$

3 次の自然スプライン ( $M_0 = 0, M_n = 0$  と置く場合) は、 $\kappa_0 = \delta_0 = 1 - \kappa_n = \delta_n = 0$  とすればよい。

## B-スプライン

B-スプライン補間は、Local Support を持つ関数からなる規定を用いて関数を表現する方法の一種で、(…日本語考える)

### 11.1.2 市場価格データを用いた最適化によるカーブ推定

#### McCulloh のアプローチ

割引関数  $DF(t)$  をいくつかの滑らかな関数  $f_1, f_2, \dots, f_n$  の 1 次結合としてモデル化する。 $f_1, f_2, \dots, f_n$  は  $f_i(0) = 0$  なる滑らかな関数、割引関数  $DF(t)$  がこれらの 1 次結合で

$$DF(t) = 1 + \sum_{i=1}^n a_i \cdot f_i(t)$$

と書けると仮定する。

のとき、各時刻  $t_i$  で金額  $C_i (i = 1, 2, \dots, N)$  が支払われる利付債の時点  $t$  での価格  $P$  は

$$\begin{aligned} P &= \sum_{i=1}^N C_i \text{DF}(t_i) \\ &= \sum_{i=1}^N C_i \left( 1 + \sum_{j=1}^n a_j f_j(t) \right) \end{aligned}$$

## 11.2 確率微分方程式を用いたスポットレートのモデル化

以下、瞬間的なスポットレート  $r(t)$  が観測確率  $\mathbf{P}$  に関して確率微分方程式でモデル化されるとする（スポットレートモデル）。

### 11.2.1

## 第12章 生保・年金数理入門

本章では、生命保険と年金数理に関する数理についての入門的な内容について概説する。生保数理に関する各参考文献を参照することを見越して、ここまでは異なる記法を使うことに留意されたい。

### 12.1 現価計算

利息の計算法には、単利・複利の計算の仕方があるが、以下、複利計算のみを用いることとする。

$A$  円を年利  $i$  で  $n$  年間預けたときの  $n$  年後の元利合計（＝元本＋利息）は  $(1+i)^n A$  円となる。逆に  $n$  年後に  $A$  円の価値を持つために現在いくら持っていればよいかを考え、その値を  $B$  円とすると  $A = (1+i)^n B$  の関係が成り立つので

$$B = \frac{1}{(1+i)^n} A$$

という関係が成り立つ。ここで、 $v := 1/(1+i)$  とすれば、 $B = v^n A$  と書くことが出来る。将来の価値を現在の価値に変換するときに使う  $v$  を**現価率**などと呼ぶ。以下、保険料や積立金などの計算の基礎となる利率を**予定利率**といい  $i$  で表す。また、 $i$  から定まる現価率を  $v \equiv v(i)$  で表す。

利息を元金に組入れることを**転化**といい、1年間の組入れ回数を**転化回数**と呼ぶ。転化回数  $k$  のとき、転化回数が1回のときより元利合計は大きくなると考えられる。年利  $\tilde{i}$ （これを**名称利率**という）で転化回数  $k$  のときの実質的な利率（**実利率**）はいくらかを考えると、1年後の元利合計は

$$\left(1 + \frac{\tilde{i}}{k}\right)^k A$$

となることから、このときの実利率を  $i$  とすると

$$1+i = \left(1 + \frac{\tilde{i}}{k}\right)^k$$

となる。逆に実利率  $i$  を固定して、転化回数が  $k$  であるときの名称利率を  $\tilde{i}(k)$  とすると

$$\tilde{i}(k) = k \left( (1+i)^{1/k} - 1 \right)$$

となる。

ここで、 $(1+i)^{1/k} = e^{\frac{1}{k} \log(1+i)}$  と書き換えたうえで Taylor 展開を用いれば

$$\tilde{i}(k) = \log(1+i) + \frac{1}{2!k} (\log(1+i))^2 + \frac{1}{3!k^2} (\log(1+i))^3 + \dots$$

となり、 $k \rightarrow \infty$  としたときの  $\tilde{i}(k)$  を  $\delta$  とすると

$$\delta = \log(1+i)$$

となる。この  $\delta$  を**利力**といい、利力を用いれば  $v^t = e^{-\delta t}$  と表される。

## 12.2 確定年金

## 第13章 実践演習

**演習 13.1.** InterestRatesData.xlsx の [01Spline] シートに与えられたデータをもとに、節点間を1次スプライン、2次スプライン、3次スプラインの3つの方法で補間し、グラフを描画せよ。



## 第III部

# 真面目に確率論と確率解析





最後に（ほぼおまけ）、真面目に数学をやっておく。



## 第14章 確率解析

本章では、確率解析について解説を行う。

確率解析にたどり着くためには、根底に測度論に基づく確率論が必要となり、その理解には Lebesgue 積分論が必要となる。そのため、本章は以下のように構成した。まず、[§14.1](#)において、測度論およびそれに基づく確率論について解説を行い<sup>1</sup>、そのうえで、確率過程とその性質について基本的な内容を [§14.2](#)で取り扱う。ここまでが準備段階にあたり、既知であればもちろんスキップしてかまわない内容である。その後、確率解析については [§14.3](#) および [§14.4](#)で説明を行う。

基本的には、主張のみを並べており適宜証明については、[\[1\]](#)、[\[7\]](#)、[\[8\]](#)、[\[2\]](#)、[\[9\]](#)などを参照されたい。

### 14.1 測度と確率の復習

#### 14.1.1 測度から積分まで

$S$  を空でない集合とする。

**定義 14.1** ( $\sigma$ -加法族,  $\sigma$ -algebra).  $S$  の部分集合族  $\mathcal{F}$  が

(F1)  $\emptyset \in \mathcal{F}$ ,

(F2)  $A \in \mathcal{F} \implies A^c \in \mathcal{F}$ ,

(F3)  $A_n \in \mathcal{F} (n = 1, 2, \dots) \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ ,

の3つの条件を満たすとき、 $\mathcal{F}$  は  $S$  上の  $\sigma$ -加法族 ( $\sigma$ -algebra) といい、組  $(S, \mathcal{F})$  を可測空間 (measurable space) という。

**定義 14.2** (測度, measure). 可測空間  $(S, \mathcal{F})$  に対し、写像  $f : \mathcal{F} \rightarrow [0, \infty]$  が

(1)  $\mu(\emptyset) = 0$ ,

(2)  $A_n \in \mathcal{F} (n = 1, 2, \dots), A_n \cap A_m = \emptyset (n \neq m) \implies f(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} f(A_n)$ ,

を満たすとき、 $f$  は  $(S, \mathcal{F})$  上の測度 (measure) といい、3つ組  $(S, \mathcal{F}, f)$  を測度空間 (measure space) という。測度  $f$  は  $f(S) < \infty$  になるとき、有限測度 (finite measure) といい、特に  $f(S) = 1$  になるとき確率測度 (probability measure) という。

**命題 14.1.**  $(S, \mathcal{F})$  を可測空間とする。この時、以下が成り立つ。

<sup>1</sup>必要となる測度論的確率論については [§14.1.2](#)、そしてそのために必要な Lebesgue 積分論については [§14.1.1](#) で解説した。

1.  $S \in \mathcal{F}$ ,
2.  $A_n \in \mathcal{F} (n = 1, 2, \dots) \implies \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$ .

ma-alg

**定義 14.3** (部分集合族によって生成される  $\sigma$ -加法族). 集合  $S$  の部分集合族  $\mathcal{G}$  に対し,

$$\sigma(\mathcal{G}) := \bigcap_{\mathcal{H} \in \mathcal{C}} \mathcal{H}, \quad \mathcal{C} := \{ \mathcal{H} : \mathcal{H} \text{ は } S \text{ の } \sigma\text{-加法族で, } \mathcal{G} \subset \mathcal{H} \text{ を満たす} \}$$

により定義される  $S$  上の  $\sigma$ -加法族  $\sigma(\mathcal{C})$  を,  $\mathcal{G}$  で生成される  $\sigma$ -加法族あるいは,  $\mathcal{G}$  を含む最小の  $\sigma$ -加法族という.

**Q** 定義 14.3 で定義される  $\sigma(\mathcal{G})$  が  $\sigma$ -加法族であることを確認せよ.

**定義 14.4** (ボレル  $\sigma$ -加法族, Borel  $\sigma$ -algebra). 位相空間  $S$  のすべての開集合の集まり  $\mathcal{O}(S)$  に対し,  $\mathcal{B}(S) := \sigma(\mathcal{O}(S))$  とおき,  $S$  上の Borel  $\sigma$ -加法族と呼ぶ. また,  $\mathcal{B}(S)$  の元を  $S$  の Borel 集合という.

**注意 19.**  $S$  の開集合や閉集合, およびそれらから可算個の和・積演算によって得られる集合は, すべて Borel 集合である.

**注意 20.** 重要な例として,  $S = \mathbb{R}^n$  の場合の Borel  $\sigma$ -加法族  $\mathcal{B}(\mathbb{R}^n)$  は重要となる. そのため, 以下, 特段の断りがない限り,  $\mathbb{R}^n$  を可測空間とみるとときには, それは  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  のこととする.

**注意 21.** 再度, 同様に,  $[a, b], (a, b], [a, b)$  などの区間  $I \subset \mathbb{R}$  上の  $\sigma$ -加法族も, 特段の断りがない限り  $\mathcal{B}(I)$  のこととする.

**定義 14.5** (ルベグ測度, Lebesgue measure). 実数の区間  $[a, b], -\infty < a \leq b < \infty$  に対し,  $([a, b], \mathcal{B}([a, b]))$  上の測度  $\lambda$  で,  $a \leq x \leq y \leq b$  なる  $x, y$  に対し,  $\lambda([x, y]) = y - x$  となるものが一意に存在<sup>2</sup>する. この  $\lambda$  を  $([a, b], \mathcal{B}([a, b]))$  上の Lebesgue 測度という.

測度について, これまでの定義から得られる性質をいくつか纏めておこう.

**命題 14.2** (有限加法性, 劣加法性).  $(S, \mathcal{F}, f)$  を測度空間,  $A_1, A_2, \dots \in \mathcal{F}$  とする.

- (1)  $A_1 \cap A_2 = \emptyset \implies f(A_1 \cup A_2) = f(A_1) + f(A_2)$ . (有限加法性)
- (2)  $A_1 \subset A_2 \implies f(A_2) = f(A_1) + f(A_2 \setminus A_1), f(A_1) \leq f(A_2)$
- (3)  $f(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} f(A_n)$ . (劣加法性)

**系 14.3.**  $(S, \mathcal{F}, f)$  を測度空間,  $A_1, A_2, \dots \in \mathcal{F}$  とする.

- (1)  $f(A_n) = 0 (n = 1, 2, \dots) \implies f(\bigcup_{n=1}^{\infty} A_n) = 0$ .
- (2)  $f(S) < \infty$  かつ  $f(A_n) = f(S) (n = 1, 2, \dots) \implies f(\bigcap_{n=1}^{\infty} A_n) = f(S)$ .

**定理 14.4** (単調収束定理, monotone convergence theorem).  $(S, \mathcal{F}, f)$  を測度空間,  $A_1, A_2, \dots \in \mathcal{F}$  とする.

<sup>2</sup> これを確認するには, かなり厳密にやる必要がある

$$(1) A_n \subset A_{n+1} (n = 1, 2, \dots) \implies f(A_n) \uparrow f(\bigcup_{n=1}^{\infty} A_n)$$

$$(2) f(A_1) < \infty \text{ かつ } A_n \supset A_{n+1} (n = 1, 2, \dots) \implies f(A_n) \downarrow f(\bigcap_{n=1}^{\infty} A_n).$$

**定義 14.6** (部分  $\sigma$ -加法族). 可測空間  $(S, \mathcal{F})$  に対して,  $\mathcal{G}$  が  $\mathcal{F}$  の部分  $\sigma$ -加法族であるとは,  $\mathcal{G}$  自体が  $S$  の  $\sigma$ -加法族であってかつ,  $\mathcal{G} \subset \mathcal{F}$  を満たすことをいう.

**系 14.5.**  $\mathcal{G}, \mathcal{H}$  をそれぞれ  $\mathcal{F}$  の部分  $\sigma$ -加法族とする. このとき,  $\mathcal{G} \cap \mathcal{H}$  もまた,  $\mathcal{F}$  の部分  $\sigma$ -加法族.

一方で, 和集合は一般に  $\sigma$ -加法族とならない. そこで, 以下の記号を導入しておく.

$$\mathcal{G} \vee \mathcal{H} := \sigma(\mathcal{G} \cup \mathcal{H}).$$

**定義 14.7** (可測, measurable). 2つの可測空間  $(S, \mathcal{F}), (S', \mathcal{G})$  に対し, 写像  $f: S \rightarrow S'$  が可測であるとは,  $F \in \mathcal{G} \implies f^{-1}(F) \in \mathcal{F}$  が成り立つことであり,  $f$  を  $(S, \mathcal{F})$  から  $(S', \mathcal{G})$  への可測写像であるという.

**命題 14.6.** 3つの可測空間  $(S, \mathcal{F}), (S', \mathcal{G}), (S'', \mathcal{H})$  と, 可測写像  $f: S \rightarrow S', g: S' \rightarrow S''$  について,  $g \circ f: S \rightarrow S''$  も可測.

**定義 14.8** (可測関数). 2つの可測空間  $(S, \mathcal{F})$  に対し, 関数  $f: S \rightarrow \mathbb{R}$  が可測であるとは,  $F \in \mathcal{B}(\mathbb{R}) \implies f^{-1}(F) \in \mathcal{F}$  が成り立つことをいう. 特に,  $\mathcal{F}$  を強調する場合には,  $f$  は  $\mathcal{F}$ -可測であるという.

**命題 14.7.**  $f: S \rightarrow \mathbb{R}$  が  $\mathcal{F}$ -可測であることと,  $y \in \mathbb{R} \implies \{x \in S \mid f(x) \leq y\} \in \mathcal{F}$  が成り立つことは同値.

**系 14.8.** 2つの  $\mathcal{F}$ -可測関数  $f, g: S \rightarrow \mathbb{R}$  と,  $a, b \in \mathbb{R}$  に対し,  $af + bg$  もまた  $\mathcal{F}$ -可測である.

「 $\mathcal{F}$ -可測な関数」という言葉を頻繁に用いることから, 以下のクラスを準備しておく.

$$\mathbf{m}\mathcal{F} := \{f \mid f \text{ は } \mathcal{F}\text{-可測関数}\}.$$

**系 14.9.** 可測空間  $(S, \mathcal{F})$  上の可測関数列  $f_1, f_2, \dots \in \mathbf{m}\mathcal{F}$  に対し, 次はいずれも可測関数:

$$(1) \sup_n f_n, \inf_n f_n.$$

$$(2) \limsup_{n \rightarrow \infty} f_n, \liminf_{n \rightarrow \infty} f_n.$$

また,  $E := \{s \in S \mid \lim_{n \rightarrow \infty} f_n(s) \in \mathbb{R} \text{ が存在する}\}$  とおくと,  $E \in \mathcal{F}$  が成り立つ.

**命題 14.10.**  $(S, \mathcal{F})$  を可測空間とし,  $f: S \rightarrow S'$  とする.  $\mathcal{G}$  を  $S'$  の部分集合族 ( $\sigma$ -加法族とは限らない!!) に対し,  $\mathcal{G}' := \sigma(\mathcal{G})$  とおく. このとき,

$$G' \in \mathcal{G}' \implies f^{-1}(G') \in \mathcal{F}$$

が成り立つならば,  $f$  は  $(S, \mathcal{F})$  から  $(S', \mathcal{G}')$  への可測写像である.

$A \subset S$  に対し,  $S$  上の関数  $\mathbf{1}_A$  を

$$\mathbf{1}_A(x) := \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

で定義する. この  $\mathbf{1}_A$  を  $A$  の指示関数 (indicator function) という.

**定義 14.9** (単関数, simple function).  $(S, \mathcal{F})$  を可測空間とする. 可測関数  $f$  が単関数であるとは, ある  $n \in \mathbb{N}$  と  $a_1, a_2, \dots, a_n \in \mathbb{R}$  と  $F_1, F_2, \dots, F_n \in \mathcal{F}$  があって,

$$f(x) = \sum_{k=1}^n a_k \mathbf{1}_{F_k}(x), x \in S$$

と書けることをいう.

$SF$  で単関数全体,  $SF^+$  で非負単関数全体の集合を表すものとする.

**定義 14.10** (殆どいたるところ, almost everywhere). 測度空間  $(S, \Sigma, \mu)$  に対し,  $x \in S$  に関する命題  $\mathcal{S}(x)$  が殆どいたるところで成り立つとは,

$$\mathcal{S}_F := \{x \in S \mid \mathcal{S}(x) \text{ が偽} \} \in \mathcal{F} \text{ かつ } \mu(\mathcal{S}_F) = 0$$

が成り立つことである. almost everywhere を省略して, a.e. と書き, 測度  $f$  を強調する場合には,  $\mu$ -a.e. と書く.

a.e. の使い方を例で紹介する. 例えば,  $S$  上の 2 つの可測関数  $f, g$  に対し,  $f = g$   $\mu$ -a.e. とは,  $\mu(\{x \in S \mid f(x) \neq g(x)\}) = 0$  が成り立つことである.

**定義 14.11** (積分 (単関数の場合)).  $(S, \Sigma, \mu)$  を測度空間,  $f : S \rightarrow [0, \infty)$  を非負単関数とする. この  $f$  に対し

$$\int_S f(x) \mu(dx) := \sum_{k=1}^n a_k \mu(E_k)$$

と置き, これを  $f$  の  $\mu$  に関する  $S$  上の積分という. 但し,  $0 \times \infty = 0$  とする.

**注意 22.** 積分の値は  $f$  の表示形式によらない.

**定義 14.12** (積分 (非負可測関数の場合)). 一般の非負可測関数  $f : S \rightarrow [0, \infty]$  に対して,

$$\int_S f d\mu = \sup \left\{ \int_S g d\mu \mid g \in SF^+, 0 \leq g(x) \leq f(x) (\forall x \in S) \right\}$$

によって積分を定義する.

**定義 14.13** (可積分, integrable). 可測関数  $f : S \rightarrow [-\infty, \infty]$  で

$$\int_S |f(x)| \mu(dx) < \infty$$

を満たすものを可積分という.

**定義 14.14** (積分). 可積分な  $f$  に対して, その  $\mu$  に関する  $S$  上の積分を

$$\int_S f(x) \mu(dx) := \int_S f^+(x) \mu(dx) - \int_S f^-(x) \mu(dx) \in \mathbb{R}$$

で定義する. 但し,  $f^+(x) := \max\{f(x), 0\}$ ,  $f^-(x) := -\min\{f(x), 0\}$ ,  $x \in S$  で定義する.

非負または単に可積分な関数  $f$  の  $\mu$  に関する積分を  $I(f)$  あるいは,  $\mu$  に関する積分であることを強調するように  $I^\mu(f)$  と書く:

$$I(f) = I^\mu(f) := \int_S f(s) \mu(ds).$$

ルベーグ積分に対しては, 次のように書くことにする:

$$\int_{[a,b]} f(x) \lambda(dx) = \int_a^b f(x) dx, \quad \int_{\mathbb{R}} f(x) \lambda(dx) = \int_{-\infty}^{\infty} f(x) dx.$$

積分の性質を見ておこう.

**命題 14.11.**  $[0, \infty]$  に値をとる可測関数  $f, g$  に対し, 次がそれぞれ成り立つ.

- (1)  $a, b \in [0, \infty)$  に対して,  $I(af + bg) = a \cdot I(f) + b \cdot I(g)$  (積分の線形性),
- (2)  $I(f) \geq 0$ ,
- (2)'  $f \geq g$  a.e.  $\implies I(f) \geq I(g)$ ,
- (3)  $\mu(f > 0) > 0 \implies I(f) > 0$ ,
- (4)  $I(f) = 0 \implies f = 0$  a.e.,
- (5)  $I(f) < \infty \implies f < \infty$  a.e. .

**注意 23.** D. Williams(1991) では,  $\{s \in S \mid f(s) > 0\}$  を省略して  $\{f > 0\}$  という記法導入している. また,  $\mu(\{f > 0\})$  を  $\mu(f > 0)$  と書いたりしている.

可積分関数についても性質を纏めておこう.

**命題 14.12.**  $f, g$  を可積分関数とする. 次が成り立つ.

- (1)  $a, b \in \mathbb{R}$  に対して,  $I(af + bg) = a \cdot I(f) + b \cdot I(g)$  (積分の線形性),
- (2)  $f \geq g$  a.e.  $\implies I(f) \geq I(g)$ ,
- (3)  $|I(f)| \leq I(|f|)$ .

極限操作, 積分の順序交換に関して重要な主張を示しておく.

**定理 14.13** (単調収束定理, monotone convergence theorem). 測度空間  $(S, \Sigma, \mu)$  状の可測関数  $f$  および  $f_n (n = 1, 2, \dots)$  について,

$$0 \leq f_n \uparrow f \text{ a.e. } I(f_n) \uparrow I(f).$$

**定理 14.14** (Fatou の補題, Fatou's Lemma). 測度空間  $(S, \Sigma, \mu)$  上の可測関数  $f$  および  $f_n (n = 1, 2, \dots)$  について,

$$\text{任意の } n \in \mathbb{N} \text{ について } f_n \geq 0 \text{ a.e.} \implies I\left(\liminf_{n \rightarrow \infty} f_n\right) \leq \liminf_{n \rightarrow \infty} I(f_n)$$

DOM

**定理 14.15** (優収束定理 或いは Lebesgue の収束定理, dominated convergence theorem). 測度空間  $(S, \Sigma, \mu)$  上の可測関数  $f$  および  $f_n (n = 1, 2, \dots)$  について,  $f_n \rightarrow f$  a.e. で任意の  $n \in \mathbb{N}$  について,  $|f_n| \leq g$  a.e. を満たす可積分関数  $g$  が存在するとき,  $I(f_n) \rightarrow I(f)$ .

**注意 24.** 定理 14.15 において,  $\mu(S) < \infty$  かつ  $g$  が正の定数の場合を有界収束定理 (bounded convergence theorem) という.

積分に関する不等式については, 以下の 2 つが有名である.

**定理 14.16** (Holder の不等式).  $f, g$  は測度空間  $(S, \Sigma, \mu)$  上の可測関数とする.  $p \in (1, \infty), (1/p) + (1/q) = 1$  とすると,

$$I(|fg|) \leq I(|f|^p)^{1/p} I(|g|^q)^{1/q}.$$

**注意 25.** 特に,  $p = q = 2$  の時を Schwarz の不等式という.

Minkowski

**定理 14.17** (Minkowski の不等式).  $f, g$  は測度空間  $(S, \Sigma, \mu)$  上の可測関数とする.  $p \in [1, \infty)$  とすると,

$$I(|f + g|^p)^{1/p} \leq I(|f|^p)^{1/p} + I(|g|^p)^{1/p}.$$

**定義 14.15** ( $\mathcal{L}^p$ -空間).  $(S, \Sigma, \mu)$  上の可測関数  $f$  で  $I(|f|^p) < \infty$  を満たすものの全体を  $\mathcal{L}^p(S, \mathcal{F}, \mu)$  と書く. 省略した記法として  $\mathcal{L}^p, \mathcal{L}^p(S), \mathcal{L}^p(\mu)$  といったものがある.

**注意 26.** もう少し, 厳密にルールを設けておくと,  $f, g \in \mathcal{L}^p(S, \Sigma, \mu)$  に対し,  $f = g$   $\mu$ -a.e. のとき,  $f = g$  と書いて  $f, g$  を同一視する.

**注意 27.** 定義よりあきらかに  $\mathcal{L}^1(S, \mathcal{F}, \mu)$  は可積分関数のクラスと一致する.

**定義 14.16** ( $p$  乗可積分,  $p$ -th power integrable).  $f \in \mathcal{L}^p(S, \Sigma, \mu)$  のとき,  $f$  は  $(S$  上)  $p$  乗可積分であるという.

**注意 28.** Minkowski の不等式 (定理 14.17) より,  $\mathcal{L}^p$  が線形空間をなしていることが分かる.

$f \in \mathcal{L}^p$  に対して,  $\|f\|_p := I(|f|^p)^{1/p}$  とおくと, Minkowski の不等式 (定理 14.17) より  $\|\cdot\|_p$  は  $\mathcal{L}^p$  上のノルムとなる.  $\|f\|_p$  を  $f$  の  $\mathcal{L}^p$ -ノルムという.

**定義 14.17** ( $\mathcal{L}^p$ -収束). 関数列  $f_n, n = 1, 2, \dots$  が  $f \in \mathcal{L}^p$  に  $\mathcal{L}^p$ -収束するとは,  $\|f_n - f\|_p \rightarrow 0$  が成り立つことをいう. このことを  $f_n \rightarrow f$  in  $\mathcal{L}^p$  や,  $f_n \xrightarrow{\mathcal{L}^p} f$  などと書く. これは,  $\mathcal{L}^p$  上の距離  $d(f, g) := \|f - g\|_p$  に関する収束に他ならない.

このノルムから定まる距離  $d(\cdot, \cdot)$  に関し  $\mathcal{L}^p$  は完備である.

**定理 14.18** (Riesz-Fischer).  $\mathcal{L}^p (1 \leq p < \infty)$  は完備距離空間.

完備なノルム空間を Banach 空間と呼ぶのであったことを思い出せば, この主張は  $\mathcal{L}^p$  はバナッハ空間であるということに他ならない.



**命題 14.19.**  $1 \leq p < \infty$  とする.  $f_n \xrightarrow{\mathcal{L}^p} f$  のとき, 次の主張が成り立つ:

- (1)  $\sup_n \|f_n\|_p < \infty$ ,
- (2)  $|f_n| \xrightarrow{\mathcal{L}^p} |f|$ ,
- (3)  $|f_n|^p \xrightarrow{\mathcal{L}^1} |f|^p$ .

続いて, 2次元以上の積分を扱うべく, 直積  $\sigma$ -加法族の概念を導入する.

**定義 14.18** (直積  $\sigma$ -加法族, product  $\sigma$ -algebra). 2つの可測空間  $(S_1, \Sigma_1), (S_2, \Sigma_2)$  に対し,  $S_1 \times S_2$  上の  $\sigma$ -加法族  $\Sigma_1 \otimes \Sigma_2$  を

$$\Sigma_1 \otimes \Sigma_2 := \sigma(\mathcal{F}), \quad \mathcal{F} := \{F_1 \times F_2 \mid F_1 \in \Sigma_1, F_2 \in \Sigma_2\}$$

で定義し, これを,  $\Sigma_1, \Sigma_2$  の直積  $\sigma$ -加法族という.

**命題 14.20.**  $\Sigma, \mathcal{F}, \mathcal{G}$  をそれぞれ  $S, U, V$  上の  $\sigma$ -加法族とする. このとき以下が成り立つ:

- (1)  $u: S \rightarrow U, v: S \rightarrow V$  は可測とする. すると,  $x \mapsto f(x) := (u(x), v(x))$  で定義される写像  $f: S \rightarrow U \times V$  は  $(S, \Sigma)$  から  $(U \times V, \mathcal{F} \otimes \mathcal{G})$  への可測写像である.
- (2)  $u: U \rightarrow S$  を可測とする. すると,  $(x, y) \mapsto f(x, y) := u(x)$  で定義される写像  $f: U \times V \rightarrow S$  は  $(U \times V, \mathcal{F} \otimes \mathcal{G})$  から  $(S, \Sigma)$  への可測写像である.
- (3)  $f: U \times V \rightarrow S$  は  $(U \times V, \mathcal{F} \otimes \mathcal{G})$  から  $(S, \Sigma)$  への可測写像とし,  $y_0 \in V$  とする. このとき,  $x \mapsto g(x) := f(x, y_0)$  で定義される写像  $g: U \rightarrow S$  は可測である.

**定義 14.19** ( $\sigma$ -有限). 測度空間  $(S, \Sigma, \mu)$  が  $\sigma$ -有限であるとは, 可測集合の列  $A_n \in \Sigma$  で  $S = \bigcup_n A_n$  かつ  $\mu(A_n) < \infty$  となるものが存在することである.

有限測度空間や  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$  は  $\sigma$ -有限な測度空間の例である.

**定義 14.20** (直積測度). 2つの  $\sigma$ -有限な測度空間  $(S_i, \Sigma_i, \mu_i), i = 1, 2$  と  $S_1 \times S_2$  上の直積  $\sigma$ -加法族  $\Sigma_1 \otimes \Sigma_2$  に対し, 可測空間  $(S_1 \times S_2, \Sigma_1 \otimes \Sigma_2)$  上の測度  $\mu$  で

$$F_i \in \Sigma_i, i = 1, 2 \implies \mu(F_1 \times F_2) = \mu_1(F_1) \cdot \mu_2(F_2)$$

なるものがただ一つ存在し, この  $\mu$  を  $\mu_1, \mu_2$  の直積測度といい,  $\mu = \mu_1 \otimes \mu_2$  で表す.

Fubini

**定理 14.21** (Fubini). 2つの  $\sigma$ -有限な測度空間  $(S_i, \Sigma_i, \mu_i), i = 1, 2$  の直積測度空間  $(S_1 \times S_2, \Sigma_1 \otimes \Sigma_2, \mu_1 \otimes \mu_2)$  上の可測関数  $f$  に対し, 次の主張が成り立つ:

- (1)  $f \geq 0$  a.e., ならば  $y \mapsto \int_{S_1} f(x, y) \mu_1(dx)$  と  $x \mapsto \int_{S_2} f(x, y) \mu_2(dy)$  はそれぞれ  $\Sigma_1, \Sigma_2$  について可測であり,  $+\infty$  の値をとる場合も含め

$$\begin{aligned} \int_{S_1 \times S_2} f(z) \mu_1 \otimes \mu_2(dx) &= \int_{S_2} \left( \int_{S_1} f(x, y) \mu_1(dx) \right) \mu_2(dy) \\ &= \int_{S_1} \left( \int_{S_2} f(x, y) \mu_2(dy) \right) \mu_1(dx). \end{aligned}$$

(2)  $f \in \mathcal{L}^1(\mu_1 \otimes \mu_2)$  ならば,  $y \mapsto \int_{S_1} f(x, y) \mu_1(dx)$  と  $x \mapsto \int_{S_2} f(x, y) \mu_2(dy)$  はそれぞれ  $\mathcal{L}^1(\mu_2), \mathcal{L}^1(\mu_1)$  の元であり, (1) と同様に

$$\begin{aligned} \int_{S_1 \times S_2} f(z) \mu_1 \otimes \mu_2(dx) &= \int_{S_2} \left( \int_{S_1} f(x, y) \mu_1(dx) \right) \mu_2(dy) \\ &= \int_{S_1} \left( \int_{S_2} f(x, y) \mu_2(dy) \right) \mu_1(dx) \end{aligned}$$

が成り立つ. この場合, これら 3 つの積分は同じ  $\mathbb{R}$  の値をとる.

**注意 29.** 上の定理における (1) の主張を Fubini-Tonelli の定理ともいう.

**注意 30.** (2) で必要な  $f$  の可積分性は,  $|f| \geq 0$  に注意して, (1) を用いて, 例えば

$$\int_{S_2} \left( \int_{S_1} |f(x, y)| \mu_1(dx) \right) \mu_2(dy) < \infty$$

を確かめればよい.

測度および積分についての準備を終えたので, ここからは測度論に基づいた確率論について, 基本的な用語や事実を纏める.

### 14.1.2 測度論的確率論

**定義 14.21** (確率空間). 測度空間  $(\Omega, \mathcal{F}, \mathbf{P})$  は,  $\mathbf{P}$  が確率測度であるとき確率空間といい, このとき  $\Omega$  を標本空間 (sample space),  $\omega \in \Omega$  を標本元 (sample element),  $F \in \mathcal{F}$  を事象 (event) と呼ぶ.

以下, 本稿では, 確率空間は常に完備であると仮定しておく.

**定義 14.22** (完備). 確率空間  $(\Omega, \mathcal{F}, \mathbf{P})$  が完備であるとは,

$$N \in \mathcal{F}, \mathbf{P}(N) = 0, A \subset N \implies A \in \mathcal{F}$$

が成り立つことをいう.

**定義 14.23** (確率変数). 確率空間  $(\Omega, \mathcal{F}, \mathbf{P})$  上の可測関数  $X: \Omega \rightarrow \mathbb{R}^n$  を  $\mathbb{R}^n$ -値確率変数 (random variable, R.V.) という. 特に  $n = 1$  のとき実確率変数と呼んだりもする.

**注意 31.** 確率変数が, 可測関数  $X: \Omega \rightarrow [-\infty, \infty]$  である場合もある.

確率空間  $(\Omega, \mathcal{F}, \mathbf{P})$  の場合, almost everywhere(a.e.) の代わりに, **ほとんど確実に (almost surely, a.s.)** という. 特に確率測度  $\mathbf{P}$  を強調する場合には,  $\mathbf{P} -$  (a.s.) と書く.  $\omega \in \Omega$  に関するある命題  $S(\omega)$  が a.s. に成り立つことは,  $E := \{\omega \in \Omega \mid S(\omega) \text{ は真} \} \in \mathcal{F}$  かつ  $\mathbf{P}(E) = 1$  が成り立つことと同値である.

**例 14.1.**  $X = Y, \mathbf{P} -$  (a.s.) とおは,  $X$  と  $Y$  が確率 1 で等しい  $\mathbf{P}(X = Y) = 1$  が成り立つことである. 事象  $\{\omega \in \Omega \mid X(\omega) = Y(\omega)\}$  を単に  $\{X = Y\}$  と書いたりする.

**定義 14.24.** 確率変数  $X$  が有界 (bounded) であるとは, ある  $K \in (0, \infty)$  が存在して,  $\mathbf{P}(|X| \leq K) = 1$  が成り立つことをいう.

**定義 14.25** (分布・密度).  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  上の確率測度を**分布 (distribution)** という。分布  $\mu$  に対して、

$$\mu(B) := \int_B f(x)dx, B \in \mathcal{B}(\mathbb{R}^n)$$

を満たす非負の可測関数  $f$  が存在するとき、 $f$  を  $\mu$  の**密度 (density)** という。

**注意 32.**  $\mathbb{R}^n$  上の非負可測関数  $f$  がある分布の密度になるための必要十分条件は、 $f$  が

$$\int_{\mathbb{R}^n} f(x)dx = 1$$

を満たすことである。分布  $\mu$  が密度  $f$  を持ち、可測関数  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  が非負または測度  $\mu$  に対し可積分であるとき、

$$\int_{\mathbb{R}^n} g(x)\mu(dx) = \int_{\mathbb{R}^n} g(x)f(x)dx$$

が成り立つ。

**定義 14.26** (正規分布).  $m \in \mathbb{R}, \sigma \in (0, \infty)$  に対し、密度

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), x \in \mathbb{R}$$

を持つ 1 次元分布を、平均  $m$  分散  $\sigma^2$  の正規分布 (normal distribution) あるいは、Gauss 分布 (Gaussian distribution) といい、 $N(m, \sigma^2)$  で表す。

**定義 14.27** (法則). 確率空間  $(\Omega, \mathcal{F}, \mathbf{P})$  上の  $\mathbb{R}^n$  値確率変数  $X$  に対し、 $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  上の確率測度  $\mathbf{P}_X$  を次で定義し、 $X$  の分布あるいは法則 (law) という；

$$\mathbf{P}_X(B) := \mathbf{P}(X^{-1}(B)), B \in \mathcal{B}(\mathbb{R}^n).$$

$X$  の分布  $\mathbf{P}_X$  がある分布  $\mu$  に一致するとき、 $X$  は分布  $\mu$  に従うという。また、 $\mathbf{P}_X$  が密度  $f$  を持つとき、 $f$  を  $X$  の密度という。

**例 14.2.**  $X$  が正規分布  $N(m, \sigma^2)$  に従うとは、次が成り立つこと；

$$\mathbf{P}(X \in B) = \int_B \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx, B \in \mathcal{B}(\mathbb{R}).$$

**定義 14.28.** 確率変数  $X$  が**可積分**であるとは、 $I^{\mathbf{P}}(|X|) < \infty$  が成り立つ<sup>3</sup>ことをいう。

**定義 14.29** (期待値). 非負または可積分な確率変数  $X$  の**期待値 (expectation)**  $\mathbf{E}[X]$  を、 $X$  の  $P$  に関する積分として次のように定義する：

$$\mathbf{E}[X] := \int_{\Omega} X(\omega) \mathbf{P}(d\omega).$$

確率測度  $\mathbf{P}$  を強調したいときには、 $\mathbf{E}[X]$  を  $\mathbf{E}^{\mathbf{P}}[X]$  などと書く。

<sup>3</sup>非負または可積分な  $f$  の  $\mu$  に関する積分を  $I(f)$  と書く：

$$I(f) = I^{\mu}(f) := \int_S f(s)\mu(ds).$$

**注意 33.** 期待値は積分に他ならないため、これまでに示した積分に関する種々の主張は期待値についても同様に成り立つ。

**系 14.22.**  $X$  が可積分であることと、 $\mathbf{E}[|X|] < \infty$  は同値。

$\mathbb{R}^n$  値確率変数  $X$  と可測関数  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  に対し、 $g(X)$  は非負または可積分であるとする。このとき、

$$\mathbf{E}[g(X)] = \int_{\mathbb{R}^n} g(x) \mathbf{P}_X(dx)$$

が成り立つ。また、 $X$  が密度  $f$  を持つとき、

$$\mathbf{E}[g(X)] = \int_{\mathbb{R}^n} g(x) f(x) dx$$

が成り立つ。 $|g(x)| \geq 0$  より、 $\mathbf{E}[|g(X)|] = \int_{\mathbb{R}^n} |g(x)| \mathbf{P}_X(dx)$  は常に成り立つため、密度  $f$  をもつ確率変数  $X$  に対して、 $\int_{\mathbb{R}^n} |g(x)| f(x) dx < \infty$  ならば  $g(x)$  は可積分である。

確率変数  $X$  は  $\mathbf{E}[X^2] < \infty$  (すなわち、2 乗可積分) とする。このとき、Schwarz の不等式から  $\mathbf{E}[|X|] \leq \mathbf{E}[X^2]^{1/2} < \infty$  が成り立つので、 $X$  は期待値  $\mathbf{E}[X] \in \mathbb{R}$  を持つ。このことに注意し、

$$V(x) = \mathbf{E}[(X - \mathbf{E}[X])^2] \in [0, \infty)$$

を  $X$  の**分散 (variance)** という。

**系 14.23.**  $V(X) = \mathbf{E}[X^2] - \mathbf{E}[X]^2$ 。

**定義 14.30** (概収束).  $\{X_n\}_{n \in \mathbb{N}}$  が  $X$  に概収束 (almost surely convergence) するとは、 $X_n \rightarrow X$  (a.s.) なることをいう。

**定義 14.31** (確率収束).  $\{X_n\}_{n \in \mathbb{N}}$  が  $X$  に確率収束 (convergence in probability) するとは、任意の  $\varepsilon$  に対して、 $\lim_{n \rightarrow \infty} \mathbf{P}(|X_n - X| > \varepsilon) = 0$  なることをいう。

**定義 14.32.** 確率変数  $X$  に対して、 $X$  で生成される  $\mathcal{F}$  の部分  $\sigma$ -加法族  $\sigma(X)$  を

$$\sigma(X) := \{X^{-1}(A) \mid A \in \mathcal{B}(\mathbb{R})\}$$

で定義する。確率変数  $X$  が  $\mathcal{F}$  の部分  $\sigma$ -加法族  $\mathcal{G}$  に対して  $\mathcal{G}$ -可測であるとは、 $\sigma(X) \subset \mathcal{G}$  となることである。

**定義 14.33.** 確率変数の属  $\{X_\lambda\}_{\lambda \in \Lambda}$  に対して、 $\{X_\lambda\}_{\lambda \in \Lambda}$  で生成される  $\mathcal{F}$  の  $\sigma$ -加法族  $\sigma(X_\lambda : \lambda \in \Lambda)$  を

$$\sigma(X_\lambda : \lambda \in \Lambda) := \sigma\left(\bigcup_{\lambda \in \Lambda} \sigma(X_\lambda)\right)$$

で定義する。

**注意 34.**  $\sharp(\Lambda) \geq 2$  のとき、 $\bigcup_{\lambda} \sigma(X_\lambda)$  は  $\sigma$ -加法族になるとは限らない。

**系 14.24.** すべての確率変数は  $\sigma(X)$ -可測。また、任意の  $\ell \in \Lambda$  に対して、 $X_\ell$  は  $\sigma(X_\lambda : \lambda \in \Lambda)$ -可測。

**定義 14.34** (独立).  $\mathcal{F}$  の  $n$  個の部分  $\sigma$ -加法族  $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n$  が**独立 (independent)** であるとは、すべての  $G_1 \in \mathcal{G}_1, G_2 \in \mathcal{G}_2, \dots, G_n \in \mathcal{G}_n$  に対して、

$$\mathbf{P}(G_1 \cap G_2 \cap \dots \cap G_n) = \mathbf{P}(G_1)\mathbf{P}(G_2) \cdots \mathbf{P}(G_n)$$

が成り立つことをいう。

**定義 14.35** (独立-その2).  $\mathcal{F}$  の部分  $\sigma$ -加法族の族  $\{\mathcal{G}_\lambda\}_{\lambda \in \Lambda}$  が**独立 (independent)** であるとは、任意の有限個の異なる  $\lambda(1), \lambda(2), \dots, \lambda(n) \in \Lambda$  に対して、 $\mathcal{G}_{\lambda(i)}, i = 1, 2, \dots, n$  が独立となることをいう。が成り立つことをいう。

**系 14.25.**  $n$  この確率変数  $X_1, X_2, \dots, X_n$  が独立であることは、 $\sigma(X_1), \sigma(X_2), \dots, \sigma(X_n)$  が独立であることであり、これは、

$$\mathbf{P}(X_1 \in B_1, X_2 \in B_2, \dots, X_n \in B_n) = \mathbf{P}(X_1 \in B_1)\mathbf{P}(X_2 \in B_2) \cdots \mathbf{P}(X_n \in B_n), \forall B_1, B_2, \dots, B_n \in \mathcal{B}(\mathbb{R})$$

が成り立つことと同値。

### 14.1.3 条件付き期待値

**定理 14.26.**  $(\Omega, \mathcal{F}, \mathbf{P})$  を確率空間、 $X$  を  $(\Omega, \mathcal{F}, \mathbf{P})$  上の確率変数、 $\mathcal{G}$  を  $\mathcal{F}$  の部分  $\sigma$ -加法族とする。

1.  $\mathbf{E}[|X|] < \infty$  を仮定する。このとき、次を満たす確率変数  $Y$  が存在する:

(A)  $Y$  は  $\mathcal{G}$ -可測で、 $\mathbf{E}[|Y|] < \infty$  を満たす。

(B) 任意の  $G \in \mathcal{G}$  に対し、 $\mathbf{E}[X\mathbf{1}_G] = \mathbf{E}[Y\mathbf{1}_G]$ 。

(C) (A),(B) を満たす別の確率変数  $Y'$  が存在するとき、 $\mathbf{P}(Y = Y') = 1$  が成り立つ。

2.  $X \geq 0$  (a.s.) を仮定する。このとき、次を満たす確率変数  $Y$  が存在する:

(A)  $Y$  は  $\mathcal{G}$ -可測で、 $Y \geq 0$  (a.s.) を満たす。

(B) 任意の  $G \in \mathcal{G}$  に対し、 $\mathbf{E}[X\mathbf{1}_G] = \mathbf{E}[Y\mathbf{1}_G]$ 。

(C) (A),(B) を満たす別の確率変数  $Y'$  が存在するとき、 $\mathbf{P}(Y = Y') = 1$  が成り立つ。

## 14.2 確率過程

### 14.2.1 準備

### 14.2.2 Brown 運動

### 14.2.3 停止時刻

### 14.2.4 Martingale

### 14.2.5 Doob の不等式

### 14.2.6 Martingale の空間

### 14.2.7 局所 Martingale

## 14.3 確率積分と確率解析

### 14.3.1 Crash Introduction

### 14.3.2 確率積分の定義・その性質

### 14.3.3 稠密性

### 14.3.4 $\mathcal{L}_2^{\text{loc}}$ に対する確率積分の構成

### 14.3.5 伊藤過程

### 14.3.6 伊藤の公式

### 14.3.7 Girsanov の定理

### 14.3.8 Martingale 表現定理

### 14.3.9 確率微分方程式

## 14.4 (発展) より一般の確率解析

## 第15章 連続時間のファイナンス入門

### 15.1 資産のモデル化と Black-Scholes





## 関連図書

- |             |   |
|-------------|---|
| Williams    | [1] David Williams, 1999, <i>Probability with Martingales</i> , Cambridge University Press.     |
| KS          | [2] Ioanis Karatzas & Steven Shreve, <i>Brownian Motion and Stochastic Calculus</i> , Springer. |
| Don:tsa     | [3] James D. Hamilton, 1994, <i>Time Series Analysis</i> , Princeton University Press.          |
| Don:tsa     | [4] 沖本 竜義, 2010, 経済・ファイナンスデータの計量時系列分析   |
| SMText      | [5] 日本アクチュアリー会, モデリング   |
| SMProb      | [6] 日本アクチュアリー会, 確率・統計・モデリング問題集  |
| Sato        | [7] 佐藤 担, 1999, 初めての確率論 -測度から確率へ-, 共立出版.  |
| FunakiP     | [8] 舟木 直久, 1999, 確率論, 朝倉書店.   |
| FunakiS     | [9] 舟木 直久, 1999, 確率微分方程式, 朝倉書店.   |
| Kijima      | [10] 木島 正明, 1999, 期間構造モデルと金利デリバティブ, 朝倉書店.   |
| Yatsuka     | [11] 白塚 重典, 2023, 金融政策, 慶應義塾大学出版会   |
| Yakami      | [12] 二神 孝一, 2012, 動学マクロ経済学, 日本評論社   |
| Hasumi      | [13] 蓮見 亮, 2020, 動学マクロ経済学へのいざない, 日本評論社  |
| Yukawa      | [14] 吉川 大介, 2022, データ駆動型ファイナンス , 共立出版   |
| Fischer:ToI | [15] Fischer, Irving, 1930, The Theory of Interest, The Macmillan company                       |