

# 線形確率システム

経済動学 2016Q1  
資料番号: 16ED07

mail@kenjisato.jp

2016年5月26日

本稿ではこれまで扱った線形システムに確率的な要因を導入しよう。本稿は確率過程に関する初級の知識を前提にしているが、Klein (2000) の手法を理解するのに確率論の高度な知識は必要ではない。読み進めながら定義などを確認していけばよいだろう。本稿では本文の後に簡単な補論を設けて、未定義語を補うようにしている。

$(\Omega, P, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0})$  をフィルター付き確率空間、入力項  $(u_t)$  を  $\{\mathcal{F}_t\}_{t \geq 0}$  に適合した確率過程とする。<sup>1</sup> システム変数に関する (期待の) 時間発展が方程式 (1) で表されているものとする。

$$E\mathbb{E}_t x_{t+1} = Ax_t + Bu_t, \quad t = 0, 1, \dots, \quad (1)$$

ただし  $\mathbb{E}_t x_{t+1} = \mathbb{E}[x_{t+1} | \mathcal{F}_t]$  は  $\mathcal{F}_t$  に関する  $x_{t+1}$  の条件付き期待値である。<sup>2</sup>  $\mathbb{E}_t x_t = x_t$ ,  $\mathbb{E}_t u_t = u_t$  が成り立つことに注意しておこう。したがって, (1) は

$$E\mathbb{E}_t x_{t+1} = A\mathbb{E}_t x_t + B\mathbb{E}_t u_t$$

と書けば、決定論との関係がより明確になるだろう。ただし、条件付き期待値はそれ自身確率変数なので対象の複雑性が大幅に増していることは注意しておこう。それでもなお決定論の解法が必要な仕事をしてくれる。

以下では、決定論的な方法をなぞりながら仮定すべき条件を探していこう。まずは  $(E, A)$  の QZ 分解を行う。したがって、次の条件を仮定する。

仮定 1.  $(E, A)$  はレギュラーである。

$Q, Z$  をユニタリ行列,  $Q^*EZ = S$  および  $Q^*AZ = T$  を上三角行列であるとする。変数変換  $x = Zy$  の部分行列分解を

$$\begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} = \begin{bmatrix} Z_{1s} & Z_{1u} \\ Z_{2s} & Z_{2u} \end{bmatrix} \begin{bmatrix} y_t^s \\ y_t^u \end{bmatrix}$$

とする。  $x^1 \in \mathbb{R}^{n_1}$ ,  $x^2 \in \mathbb{R}^{n_2}$ ,  $y^s \in \mathbb{R}^{n_s}$ ,  $y^u \in \mathbb{R}^{n_u}$  とする。次の条件を仮定しよう。

<sup>1</sup>適合性の定義は省略しているが、実用上は  $\mathbb{E}_t u_t = u_t$ ,  $t = 0, 1, \dots$  が成り立つことを認識していればよい。

<sup>2</sup>任意の確率変数について  $\mathbb{E}_t \mathbb{E}_{t+k} z = \mathbb{E}_t z$ ,  $k \geq 0$ , が成り立つことと、 $\mathbb{E}_t$  が確率変数を確率変数に写す線形作用素であることを認識しておけばよい。

仮定 2.  $n_u \neq 0$ .

方程式の適当な並び替えによって  $x^1$  は先決変数,  $x^2$  は非先決変数であるように選んでいる. ただし, Klein (2000) の意味で先決変数とは

$$x_{t+1}^1 = \mathbb{E}_t x_{t+1}^1 + \xi_{t+1}$$

なる, 外生的な ( $\mathcal{F}_t$ )-適合確率過程 ( $\xi_t$ ) が存在する変数のこと. 定義より  $\mathbb{E}_t \xi_{t+1} = 0$  であることに注意せよ. 先決変数は  $t$  期に獲得できる情報によって ( $t+1$ ) 期の値を予測不可能な誤差 ( $\xi_{t+1}$ ) を除いて決定できる変数である. 0 期時点の先決変数の期待値は ( $-1$ ) 期には既に決定されているはずなので,  $x_0^1$  は初期値として与えられている値に外生的な攪乱を加えたものと考えてよい. これは決定論の場合に仮定したことと同等である.<sup>3</sup>

次の仮定の役割は今や明らかであろう.

仮定 3.  $n_1 = n_s$ ,  $\det Z_{1s} \neq 0$ .

さて, 決定論の場合と同様に QZ 分解の一般化固有値の並び替えによって, ブロック行列表現

$$\begin{bmatrix} S_{ss} & S_{su} \\ 0 & S_{uu} \end{bmatrix} \mathbb{E}_t \begin{bmatrix} y_{t+1}^s \\ y_{t+1}^u \end{bmatrix} = \begin{bmatrix} T_{ss} & T_{su} \\ 0 & T_{uu} \end{bmatrix} \begin{bmatrix} y_t^s \\ y_t^u \end{bmatrix} + \begin{bmatrix} C_s \\ C_u \end{bmatrix} u_t \quad (2)$$

が次のような性質をもつとする. すなわち,  $(S_{ss}, T_{ss})$  は安定 (絶対値が 1 のものを含む) な一般化固有値に対応するブロックであり,  $(S_{uu}, T_{uu})$  は不安定 (無限大固有値を含む) 一般化固有値に対応するブロックである. レギュラー性の仮定により  $T_{uu}$  は正則行列であるから, バックワードサブシステム

$$y_t^u = T_{uu}^{-1} S_{uu} \mathbb{E}_t y_{t+1}^u - T_{uu}^{-1} C_u u_t \quad (3)$$

が定義できる. これを forward-looking に解いて  $y_t^u$  を決定する.

$$\begin{aligned} y_t^u &= T_{uu}^{-1} S_{uu} \mathbb{E}_t y_{t+1}^u - T_{uu}^{-1} C_u u_t \\ &= T_{uu}^{-1} S_{uu} \mathbb{E}_t [T_{uu}^{-1} S_{uu} \mathbb{E}_{t+1} y_{t+2}^u - T_{uu}^{-1} C_u u_{t+1}] - T_{uu}^{-1} C_u u_t \\ &= (T_{uu}^{-1} S_{uu})^2 \mathbb{E}_t y_{t+2}^u - (T_{uu}^{-1} S_{uu}) T_{uu}^{-1} C_u \mathbb{E}_t u_{t+1} - T_{uu}^{-1} C_u u_t \\ &= \dots \\ &= (T_{uu}^{-1} S_{uu})^\tau \mathbb{E}_t y_{t+\tau}^u - \sum_{k=0}^{\tau-1} (T_{uu}^{-1} S_{uu})^k T_{uu}^{-1} C_u \mathbb{E}_t u_{t+k}. \end{aligned} \quad (4)$$

$T_{uu}^{-1} S_{uu}$  は安定行列であることに注意しよう. 式 (4) の第 2 項が収束性するために次を仮定しよう. 一番簡単な十分条件は  $u$  が有界なサポートを持つことである.

仮定 4. 確率過程 ( $u_t$ ) に対する期待は安定である. すなわち,

$$\sup_t \|\mathbb{E} u_t\| < \infty.$$

<sup>3</sup> 決定論とは異なり,  $x_0^1$  が外生的な確率変数であることは許容される. あえて冗長な表現をすれば  $x_0^1 = \mathbb{E}_0[x_0^1] + \xi_0$  となる.

さらに, (4) の第1項が収束することを保証するために, 解空間を制限する.  $\lim_{\tau \rightarrow \infty} (T_{uu}^{-1} S_{uu})^\tau = 0$  であるから, いつものように指数的に発散する解を排除してやればよい. 以上より,

$$y_t^u = - \sum_{k=0}^{\infty} (T_{uu}^{-1} S_{uu})^k T_{uu}^{-1} C_u \mathbb{E}_t u_{t+k} \quad (5)$$

を得る.

非先決変数を決定するための線形方程式

$$\begin{bmatrix} Z_{1s} & 0 \\ Z_{2s} & -I \end{bmatrix} \begin{bmatrix} y_t^s \\ x_t^2 \end{bmatrix} = \begin{bmatrix} I & -Z_{1u} \\ 0 & -Z_{2u} \end{bmatrix} \begin{bmatrix} x_t^1 \\ y_t^u \end{bmatrix}$$

から,

$$y_t^s = Z_{1s}^{-1} (x_t^1 - Z_{1u} y_t^u).$$

さらに, (2) の上段の式から,

$$S_{ss} \mathbb{E}_t y_{t+1}^s + S_{su} \mathbb{E}_t y_{t+1}^u = T_{ss} y_t^s + T_{su} y_t^u + C_s u_t$$

$$\begin{aligned} S_{ss} \mathbb{E}_t Z_{1s}^{-1} (x_{t+1}^1 - Z_{1u} y_{t+1}^u) + S_{su} \mathbb{E}_t y_{t+1}^u &= T_{ss} Z_{1s}^{-1} (x_t^1 - Z_{1u} y_t^u) + T_{su} y_t^u + C_s u_t \\ S_{ss} Z_{1s}^{-1} \mathbb{E}_t x_{t+1}^1 + (S_{su} - S_{ss} Z_{1s}^{-1} Z_{1u}) \mathbb{E}_t y_{t+1}^u &= T_{ss} Z_{1s}^{-1} x_t^1 + (T_{su} - T_{ss} Z_{1s}^{-1} Z_{1u}) y_t^u + C_s u_t \end{aligned}$$

$$\begin{aligned} S_{ss} Z_{1s}^{-1} \mathbb{E}_t x_{t+1}^1 &= T_{ss} Z_{1s}^{-1} x_t^1 + (S_{ss} Z_{1s}^{-1} Z_{1u} - S_{su}) \mathbb{E}_t y_{t+1}^u \\ &\quad + (T_{su} - T_{ss} Z_{1s}^{-1} Z_{1u}) y_t^u + C_s u_t \\ \mathbb{E}_t x_{t+1}^1 &= Z_{1s} S_{ss}^{-1} T_{ss} Z_{1s}^{-1} x_t^1 + (Z_{1u} - Z_{1s} S_{ss}^{-1} S_{su}) \mathbb{E}_t y_{t+1}^u \\ &\quad + Z_{1s} S_{ss}^{-1} (T_{su} - T_{ss} Z_{1s}^{-1} Z_{1u}) y_t^u + Z_{1s} S_{ss}^{-1} C_s u_t. \end{aligned}$$

式 (3) より,

$$Z_{1s} S_{ss}^{-1} (T_{su} - T_{ss} Z_{1s}^{-1} Z_{1u}) y_t^u = Z_{1s} S_{ss}^{-1} (T_{su} - T_{ss} Z_{1s}^{-1} Z_{1u}) (T_{uu}^{-1} S_{uu} \mathbb{E}_t y_{t+1}^u - T_{uu}^{-1} C_u u_t)$$

だから,

$$\mathbb{E}_t x_{t+1}^1 = \Omega_x x_t^1 + \Omega_u u_t + \Omega_y \mathbb{E}_t y_{t+1}^u$$

$$\begin{aligned} \Omega_x &= Z_{1s} S_{ss}^{-1} T_{ss} Z_{1s}^{-1} \\ \Omega_u &= Z_{1s} S_{ss}^{-1} [(T_{ss} Z_{1s}^{-1} Z_{1u} - T_{su}) T_{uu}^{-1} C_u + C_s] \\ \Omega_y &= Z_{1u} - Z_{1s} S_{ss}^{-1} S_{su} + Z_{1s} S_{ss}^{-1} (T_{su} - T_{ss} Z_{1s}^{-1} Z_{1u}) T_{uu}^{-1} S_{uu}. \end{aligned}$$

ここで, 先決変数の定義を使うと,

$$x_{t+1}^1 = \Omega_x x_t^1 + \Omega_u u_t + \Omega_y \mathbb{E}_t y_{t+1}^u + \xi_{t+1} \quad (6)$$

を得る. 非先決変数については,

$$\begin{aligned} x_t^2 &= Z_{2s}y_t^s + Z_{2u}y_t^u \\ &= Z_{2s}Z_{1s}^{-1}(x_t^1 - Z_{1u}y_t^u) + Z_{2u}y_t^u \\ &= Z_{2s}Z_{1s}^{-1}x_t^1 + (Z_{2u} - Z_{2s}Z_{1s}^{-1}Z_{1u})y_t^u \end{aligned}$$

より,

$$x_t^2 = \Psi_x x_t^1 + \Psi_y y_t^u, \quad (7)$$

$$\begin{aligned} \Psi_x &= Z_{2s}Z_{1s}^{-1} \\ \Psi_y &= Z_{2u} - Z_{2s}Z_{1s}^{-1}Z_{1u} \end{aligned}$$

とできる.

### $u$ の特定化

ここで,  $u$  が AR 過程

$$u_{t+1} = \Phi u_t + \varepsilon_{t+1}$$

であると仮定しよう. ただし,  $(\varepsilon_t)$  は独立同分布をもつ確率過程であり,  $\mathbb{E}_t \varepsilon_{t+1} = 0$ ,  $\mathbb{E}_t \varepsilon_{t+1} \varepsilon_{t+1}^\top = \Sigma$ .<sup>4</sup> 行列  $\Phi$  の固有値はすべて絶対値が 1 より小さいとする. 式 (5) と  $\mathbb{E}_t u_{t+1} = \Phi u_t$  より,<sup>5</sup>

$$y_t^u = - \sum_{k=0}^{\tau-1} (T_{uu}^{-1} S_{uu})^k T_{uu}^{-1} C_u \Phi^k u_t.$$

$M$  をシルベスタ方程式

$$M - T_{uu}^{-1} S_{uu} M \Phi = -T_{uu}^{-1} C_u$$

の解とすれば

$$y_t^u = M u_t$$

と書ける. (6) と (7) から

$$\begin{aligned} x_{t+1}^1 &= \Omega_x x_t^1 + \Omega_u u_t + \Omega_y \mathbb{E}_t y_{t+1}^u + \xi_{t+1} \\ &= \Omega_x x_t^1 + (\Omega_u + \Omega_y M \Phi) u_t + \xi_{t+1} \end{aligned}$$

および

$$x_t^2 = \Psi_x x_t^1 + \Psi_y M u_t$$

を得る. これで解の特徴づけが完了した.

<sup>4</sup>自己共分散に対する仮定は使われているか?

<sup>5</sup> $\mathbb{E}_t u_{t+2} = \mathbb{E}_t \mathbb{E}_{t+1} u_{t+2} = \mathbb{E}_t \Phi u_{t+1} = \Phi^2 u_t$  に注意せよ.

## A. 確率過程について

ランダムネスというのは、ごく荒く言って、発生する現象に複数の可能性があるということである。明日は雨が降るかもしれないし、雨は降らないかもしれない。天気予報は「明日雨になる確率」を教えてくれる。<sup>6</sup> ランダムな現象の素朴な数学的表現は、確率変数と呼ばれる2つの可能な値を持つ「変数」（ここでは  $X$  を名付けよう）のそれぞれの値に、足して1になる数字  $p_1, p_2$  を割り当てて、

$$P(X = \text{fine}) = p_1, \quad P(X = \text{rainy}) = p_2$$

のように書くことである。 $\{p_1, p_2\}$  を確率分布という。測度論的な確率論では、 $X$  を関数として定式化して、定義域  $\{\omega_1, \omega_2\}$  の方に確率を割り振る。もちろん、まったく同じことであるがこうすることによって関数解析の枠組みで確率論を扱えるので大変都合が良い。この補論では、マクロ経済学で頻出するフィルトレーションとか条件付き期待値といった用語がなんとなくイメージできるようになる程度に確率論の簡単な紹介を行なってみたいと思う。

### A.1. 確率変数

ランダムな現象を表現するために関数を使うということは、関数  $X: \omega \mapsto X(\omega)$  の取りうる値（値域）の方に「生起するかもしれないと認識している値」を対応させるということは想像がつかだろう。例えば株価のように非負の値を取る確率変数であれば

$$X: \bullet \rightarrow \mathbb{R}_+$$

という関数を考えるのである。定義域の方は何が入るだろうか？実は取りうる値を区別するのに十分な数の要素を持っていれば何でもよい。慣例に従って  $\Omega$  という記号を使うと

$$X: \Omega \rightarrow \mathbb{R}_+$$

と書ける。 $\Omega$  を状態空間、 $\Omega$  の元を根元事象、 $\Omega$  の部分集合を事象という。 $\Omega$  が意味するものは何か？という問いには特に意味がないことを強調しておこう。<sup>7</sup>  $\Omega$  がどのように分割されるか、 $\Omega$  の分割のそれぞれにどのように確率を割り振るかが大事なのである。例えば、 $X$  が2値関数であれば、 $\Omega_1 = \{\omega_1, \omega_2\}$  という2点集合を取って、 $\Omega_1 = \{\omega_1\} \cup \{\omega_2\}$  と分割し、

$$X_1(\omega) = \begin{cases} 0, & \omega \in \{\omega_1\} \\ 1, & \omega \in \{\omega_2\} \end{cases}$$

<sup>6</sup> 「天気予報が外れた」という表現を「今日の降水確率は30%ではなく50%だ」の意味で使う人はおらず、「降水確率20%で予報は曇りだと言っていたのに、実際には雨が降った」といったケースに使う。確率という言葉が浸透すればするほど、確率概念が表現するドライな実体とはかけ離れた理解が広まっていくのかもしれない。

<sup>7</sup> 神様（あるいは nature）が  $\Omega$  のうちの1つを選んでいますが、我々にはどれを選んだか分からない、といったイメージで  $\Omega$  が説明されることもあるが、そのようなイメージは理解の助けになるかもしれないし、ならないかもしれない。もっと悪いことに、神様（あるいは nature）とは何かという問題に発展してしまう。個人的には「確率変数は関数である。関数には定義域が必要だから形式的に  $\Omega$  としよう」で十分だと思う。

としてもいいし,  $\Omega_2 = [0, 1]$  として,  $\Omega_2 = [0, 1/2) \cup [1/2, 1]$  と分割し,

$$X_2(\omega) = \begin{cases} 0, & \omega \in [0, 1/2) \\ 1, & \omega \in [1/2, 1] \end{cases}$$

とするのでもよい. さらに, この分割にさえ自由度がある. 例えば,  $\Omega_3 = [0, 1]$  として,  $\Omega_3 = [0, 1/3) \cup [1/3, 1]$  と分割し,

$$X_3(\omega) = \begin{cases} 0, & \omega \in [0, 1/3) \\ 1, & \omega \in [1/3, 1] \end{cases}$$

としても問題は起こらない. 3つの例において, 関数  $X_1, X_2, X_3$  はそれぞれ異なる関数であるにも関わらず, 同じ確率的現象の表現と解釈するためには, 各事象に対して確率を割り振るやり方を同じにしてやる必要がある. つまり,

$$\begin{aligned} X_1^{-1}(\{0\}) &= \{\omega_1\}, & X_1^{-1}(\{1\}) &= \{\omega_2\} \\ X_2^{-1}(\{0\}) &= [0, 1/2), & X_2^{-1}(\{1\}) &= [1/2, 1] \\ X_3^{-1}(\{0\}) &= [0, 1/3), & X_3^{-1}(\{1\}) &= [1/3, 1] \end{aligned}$$

に注意して,

$$\begin{aligned} P_1(\{\omega_1\}) &= P_2([0, 1/2)) = P_3([0, 1/3)) = p_1 \\ P_1(\{\omega_2\}) &= P_2([1/2, 1]) = P_3([1/3, 1]) = p_2 \end{aligned}$$

と「確率」 $P_1, P_2, P_3$  を定義すれば, これらが同じ現象を表現していると言える. ただし,  $p_1 + p_2 = 1, p_1, p_2 \geq 0$ .

連続値を取るような場合には, 事象としてどのようなものを考えればよいだろうか. 例を挙げて考えてみよう. 「株価が5000円以上6000円未満である」という事象は

$$\{\omega \in \Omega \mid 5000 \leq X(\omega) < 6000\}$$

と書ける. あるいは同じことを

$$X^{-1}([5000, 6000)) = X^{-1}([0, 5000))^c \cap X^{-1}([0, 6000))$$

とも書ける.  $X^{-1}(A)$  は  $A$  の  $X$  による逆像,  $A^c$  は  $A$  の補集合である. 非負の実数値を取るような確率現象を扱う上で, 「事象全体の集合」として考えるべきは,  $\{\omega \in \Omega \mid X(\omega) \leq a\}$ ,  $\{\omega \in \Omega \mid X(\omega) < a\}$ ,  $a \in \mathbb{R}_+ \cup \{+\infty\}$  やその補集合, 和集合, 積集合などを含むものとするのがよい. さらに技術的な理由から, そのような部分集合の可算無限個の和集合も含んでいるとするのである.

以上の要請を汲んだ上で, 確率論の教科書を確認してほしい. そこには次のような定義が書いていると思う:  $\Omega$  の部分集合族  $\mathcal{F}$  が  $\sigma$ -集合族であるとは, (i)  $\Omega \in \mathcal{F}$ , (ii)  $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$ , (iii)  $A_1, A_2, \dots, A_n, \dots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$  の3つが成り立つことをいう.  $\Omega$  とその部分集合族  $\mathcal{F} = (\Omega, \mathcal{F})$  のペアを可測空間という. の出来上がりである. さらに,  $\{\omega \mid X(\omega) < a\}$  の形式の部分集合やそれらの補集合, 可算無限個の和集合, 積集合が  $\mathcal{F}$  に含まれているとき,  $X$  は可測関数という.

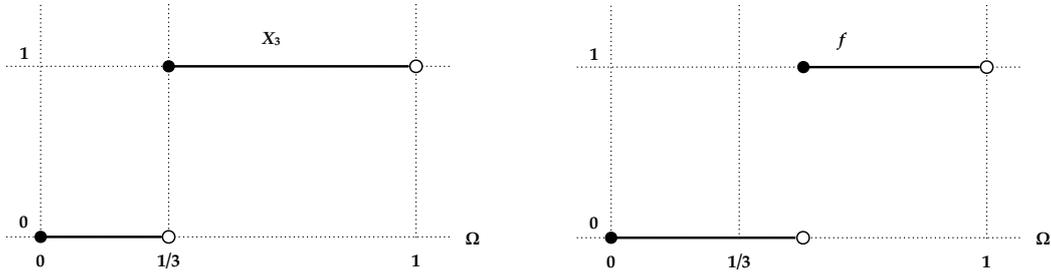


図 1: 分割  $[0, 1] = [0, 1/3) \cup [1/3, 1]$  に対する可測関数  $X_3$  と非可測関数  $f$

確率は確率測度と呼ばれる写像  $P : \mathcal{F} \rightarrow [0, 1]$  によって定式化される。確率測度は次の性質を満たす。(i)  $P(\Omega) = 1$ , (ii)  $P(A^c) = 1 - P(A)$ , (iii)  $A_1, A_2, \dots$  が  $A_i \cap A_j = \emptyset$ ,  $i \neq j$  ならば  $P(\cup_i A_i) = \sum_i P(A_i)$ . 3つ組  $(\Omega, \mathcal{F}, P)$  を確率空間という。

以上でランダムネスの構成要素が全部出揃った。適当に状態空間  $\Omega$  を設定し、状態空間を事象を表す部分集合に分割し ( $\mathcal{F}$  はその一般化・抽象化である), さらに事象ごとの生起確率を割り振っていく ( $P$  がその役割を果たしてくれる)。確率変数  $X$  は、 $\Omega$  を定義域にもつ関数であって、

$$\{\omega \in \Omega \mid X(\omega) < a\} \in \mathcal{F}$$

が成り立つものである。  $\mathcal{F}$  の元になるということは、「 $X < a$  である確率」がきちんとして定義されるように作っているということである。ところで、可測性は  $\mathcal{F}$  の選び方に依存するので、明示的に  $\mathcal{F}$ -可測と言ったりもする。

$\Omega$  が有限個の部分集合の族に分割される場合を考えると可測関数の直感的な理解が得られやすい。  $\Omega = A_1 \cup A_2 \cup \dots \cup A_n$ ,  $A_i \cap A_j = \emptyset$ ,  $i \neq j$  としよう。この場合、分割  $\Delta = \{A_1, A_2, \dots, A_n\}$  が事象の集合族を定める。すなわち、

$$\mathcal{F}_\Delta = \{\cup_{k=1}^r A_{n_k} \mid \{n_1, \dots, n_r\} \subset \{1, \dots, n\}, r = 0, 1, \dots, n\}$$

とすれば、 $\mathcal{F}_\Delta$  は  $\sigma$ -集合族となる。  $\mathcal{F}_\Delta$  は  $\Delta$  の元をすべて含む最小の  $\sigma$ -集合族であって  $\mathcal{F} \supset \mathcal{F}_\Delta$  なる  $\sigma$ -集合族であれば何を考えても以下の議論は上手くいく。2値関数の場合に可測関数と非可測関数を図示してみよう。  $\Omega_3 = [0, 1]$ ,  $\Delta_3 = \{[0, 1/3), [1/3, 1]\}$  とすれば、 $\mathcal{F}_{\Delta_3} = \{\emptyset, [0, 1/3), [1/3, 1], \Omega_3\}$  である。図1の  $X_3$  は分割  $\Delta_3$  の各集合上で定数値を取る関数である。一方で、 $f$  は  $[1/3, 1]$  上で定数関数でない。  $f$  が可測性の定義  $\{\omega \in \Omega \mid f(\omega) < a\} \in \mathcal{F}_{\Delta_3}$  を満たさないことを確認してほしい。有限個の事象のみを扱うケースでは、可測関数すなわち確率変数は、分割を構成する各集合上で定数値を取る関数に外ならない。2値関数、2事象への分割を考えながら、より一般のケースでも同様のことを確認できる。

なお、 $\Omega$  上で定数値をとる関数はもちろん可測関数である。不確実性がない現象の表現に使う。

## A.2. 期待値

期待値は積分で定義できる.  $(\Omega, \mathcal{F}, P)$  を確率空間,  $X$  を  $\mathbb{R}$ -値確率変数とする.  $X$  の期待値を

$$\mathbb{E}X = \int_{\Omega} X(\omega) dP(\omega) \quad (8)$$

と定義する. ここでの積分はいわゆるルベーグ積分である. ルベーグ積分論を未修であれば, 次のように理解すればよいだろう. 先ほど述べた可測関数の定義より,

$$\{\omega \in \Omega \mid x \leq X(\omega) < x + dx\} \in \mathcal{F}$$

が成り立つ.  $\mathcal{F}$  は  $P$  の定義域なので,

$$P(\{\omega \in \Omega \mid x \leq X(\omega) < x + dx\})$$

という確率に対応している. これを  $dP(\omega)$  と思えば,

$$\mathbb{E}X = \int_{x \in \mathbb{R}} x P(\{\omega \in \Omega \mid x \leq X(\omega) < x + dx\})$$

と書いて, 馴染みのある定式化とかなり近い形式になる. 累積分布関数を  $F$  とすれば,

$$P(\{\omega \in \Omega \mid x \leq X(\omega) < x + dx\}) = F(x + dx) - F(x)$$

だから,

$$\mathbb{E}X = \int_{x \in \mathbb{R}} x [F(x + dx) - F(x)]$$

はスティルチェス積分による表現であるし, 確率密度関数が存在する場合には,

$$P(\{\omega \in \Omega \mid x \leq X(\omega) < x + dx\}) = p(x)dx$$

と書き換えることができるので,

$$\mathbb{E}X = \int_{x \in \mathbb{R}} xp(x)dx$$

と, よりいっそう馴染み深い表現になる.

この操作で一つ見えてくることは,  $(\Omega, \mathcal{F})$  に備えた測度  $P$  が  $\mathbb{R}$  上の測度を定めるということである.  $A \subset \mathbb{R}$  が普通の (例えば区間  $(a, b)$  など) 部分集合であれば

$$X^{-1}(A) \in \mathcal{F}$$

が成り立つので,  $P(X^{-1}(A))$  が定義されている. これを  $\mathbb{R}$  の部分集合  $A$  に対して確率  $P(X^{-1}(A))$  を対応させる集合関数  $P \circ X^{-1}$  と見れば,  $P \circ X^{-1}$  は確率測度の性質をもっている.<sup>8</sup>  $X$  から誘導された確率測度を  $X$  の分布という.

<sup>8</sup> $\mathbb{R}$  の部分集合族からなる  $\sigma$ -集合族を考えなければ「測度」といっても意味がない. 通常, 开区間をすべて含む最小の  $\sigma$ -集合族を備えているものとする. これを ボレル集合族という.

簡単な例を見てみよう.

$$\begin{aligned} \mathbb{E}X_3 &= 0 \cdot P([0, 1/3)) + 1 \cdot P([1/3, 1]) \\ &= 0 \cdot p_1 + 1 \cdot p_2 \\ &= p_2 \end{aligned}$$

である. 分布は

$$\begin{aligned} P \circ X^{-1}(\{0\}) &= p_1 \\ P \circ X^{-1}(\{1\}) &= p_2 \end{aligned}$$

によって定まる.

### A.3. 条件付き期待値

$\sigma$ -集合族, あるいはそれを生成する分割が不確実な現象に対する情報精度の上限という意味を持っていることを確認しておこう.  $(\Omega, \mathcal{F}, P)$  を確率空間,  $X, Y$  を実数値確率変数とする. 説明の都合上,  $X$  が先に判明して,  $Y$  は後で判明するという時間差があるとしておく.

根元事象  $\omega \in \Omega$  を特定できれば,  $X(\omega)$  にも  $Y(\omega)$  にも不確実性は存在しない. したがって, 不確実性はどの  $\omega$  に対応する値がでてくるか分からないということに起因している. また, 先に明らかになる確率変数  $X$  の観測によって得られる情報は, あくまで  $X(\omega) = a$  とか,  $a < X(\omega) < b$  とか,  $X(\omega) \in A$  という生起した値に関する情報だけであるから  $\omega \in X^{-1}(\{a\})$  や  $\omega \in X^{-1}((a, b))$ ,  $\omega \in X^{-1}(A)$  といった情報を得ることができるが,  $\omega$  の完全な特定はできない. もし,  $Y$  が  $X^{-1}(A)$  の上で定数でなければ,  $Y$  の値を正確に予見することはできない. ただし,  $\{Y(\omega) \mid \omega \in X^{-1}(A)\}$  のいずれかの値が生起するという事だけでは分かる.

2つの重要なメッセージがある. 第一に,  $\mathcal{F}$  が細かく細分化されていけばいるほど,  $\mathcal{F}$ -可測関数の値から得られる情報は多くなる. すなわち観測値を用いて, より小さな集合の中に  $\omega$  を特定できる. 第二に, 時間を通じて  $\omega$  に関する情報が徐々に明らかになっていくということである.

具体的な例を通して見ていこう.  $\Omega = [0, 1]$  が  $\Delta = \{[0, \frac{1}{4}), [\frac{1}{4}, \frac{1}{2}), [\frac{1}{2}, \frac{3}{4}), [\frac{3}{4}, 1]\}$  と分割されているとする. 簡単のため分割を構成する各集合の確率は等しく  $\frac{1}{4}$  であるとしよう. 確率変数を

$$\begin{aligned} X(\omega) &= \begin{cases} 0 & \text{if } \omega \in [0, \frac{1}{2}) \\ 1 & \text{if } \omega \in [\frac{1}{2}, 1] \end{cases} \\ Y(\omega) &= \begin{cases} 0 & \text{if } \omega \in [0, \frac{1}{4}) \\ 1 & \text{if } \omega \in [\frac{1}{4}, \frac{1}{2}) \\ 2 & \text{if } \omega \in [\frac{1}{2}, \frac{3}{4}) \\ 3 & \text{if } \omega \in [\frac{3}{4}, 1] \end{cases} \end{aligned}$$

と特定化する.  $X, Y$  が  $\mathcal{F}_\Delta$ -可測関数であることを確認してほしい.

$X$  が実現する前には,  $\omega \in \Omega$  という情報しかないので,  $\Delta_0 = \{\Omega\}$  で表される情報集合を持っているとってよいだろう.  $X$  が実現した後は,  $\omega \in [0, \frac{1}{2})$  あるいは  $\omega \in [\frac{1}{2}, 1]$  の

いずれかであるので、 $X$  によって  $\Delta_1 = \{[0, \frac{1}{2}), [\frac{1}{2}, 1]\}$  という  $\Delta_0$  より細かいが  $\Delta$  より粗い分割で表される情報を獲得できる。さらに、 $Y$  が明らかになると、 $\Delta_2 = \Delta$  を獲得できる。それでは、 $X$  が授けてくれた情報を使って、 $Y$  について何が言えるようになるか、を考えてほしい。

まずは初級の確率論の議論をなぞってみよう。 $X$  の実現値で条件付けると、 $Y$  がそれぞれの値をとる確率が変わる。例えば、 $X = 0$  という条件の下で、 $Y = 0$  となる確率は

$$P(Y = 0|X = 0) = \frac{P(Y = 0)}{P(Y = 0) + P(Y = 1)} = \frac{1}{2}.$$

また、

$$P(Y = 2|X = 0) = 0$$

などが分かる。次に、 $X$  の実現値で条件付けると、 $Y$  に関する期待値も変化する。例えば、

$$\begin{aligned} \mathbb{E}[Y | X = 0] &= 0 \cdot P(Y = 0|X = 0) + 1 \cdot P(Y = 1|X = 0) + 2 \cdot P(Y = 2|X = 0) + 3 \cdot P(Y = 3|X = 0) \\ &= \frac{1}{2}. \end{aligned}$$

また、

$$\mathbb{E}[Y | X = 1] = \frac{5}{2}.$$

したがって、条件付き期待値は実現値  $x$  の関数として

$$\mathbb{E}[Y | X = x] = \begin{cases} \frac{1}{2} & \text{if } x = 0 \\ \frac{5}{2} & \text{if } x = 1 \end{cases}$$

などと書かれるのである。

$\mathbb{E}[Y|X = x]$  と書いてはいるが、実現値  $x$  に依存していないことに違和感を感じないだろうか。もちろん、 $x$  に依存するような条件付き期待値を作ろうと思えば作れるし、そのような例が重要であるのは事実だが、 $x$  の関数として条件付き期待値を特徴付けると本質を見誤ることになる。では何に依存していると認識するべきか？それは、特定の実現値を生成する  $\omega$  のなす集合である。すなわち条件付き期待値は  $\omega$  に依存していて、 $\Delta_1$  のどの元に  $\omega$  が属するかによって値が変わるものと捉えるべきである。したがって、次のように書く方がより本質に近い。

$$\mathbb{E}[Y|\Delta_1](\omega) = \begin{cases} \frac{1}{2} & \text{if } \omega \in [0, \frac{1}{2}) \\ \frac{5}{2} & \text{if } \omega \in [\frac{1}{2}, 1]. \end{cases}$$

この定式化は、 $\mathbb{E}[Y|\Delta_1]$  は確率変数であり、 $\mathcal{F}_{\Delta_1}$ -可測であるということを示唆している。有限分割が得られないケースでは、 $\mathbb{E}[Y|\mathcal{F}_{\Delta_1}]$  と書く。ここで、

$$\mathcal{F}_{\Delta_1} \subset \mathcal{F}_{\Delta}$$

に注意してほしい。粗い分割で得られた情報で「予測」をしているので ( $Y$  は  $\mathcal{F}_{\Delta_1}$ -可測ではない)、 $\omega$  への依存性が残ってしまうのである。

条件付き期待値の性質 一般に, 条件付き期待値は次の性質をもつ:  $(\Omega, \mathcal{F}, P)$  を確率空間  $Z$  を確率変数,  $\mathcal{G}$  を  $\mathcal{G} \subset \mathcal{F}$  なる  $\sigma$ -集合族とすると  $\mathbb{E}[Z|\mathcal{G}]$  は  $\mathcal{G}$ -可測な確率変数になる. さらに,  $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$  なる  $\sigma$ -集合族を取れば,

$$\mathbb{E}[\mathbb{E}[Z|\mathcal{G}|\mathcal{H}] = \mathbb{E}[Z|\mathcal{H}]$$

が成り立つ (正式には「ほとんど確実に」というのを付けなければならない).

これを分割の言葉で確認しておこう. 前半の部分はすでに確認できているので, 後半を確認する. 分割  $\Delta_1, \Delta_2$  に対して,  $\Delta_2$  が  $\Delta_1$  より細分化されているとは, 任意の  $A \in \Delta_2$  についてある  $B \in \Delta_1$  があって  $A \subset B$  が成り立つこととし,  $\Delta_1|\Delta_2$  と書こう.  $\Delta_1|\Delta_2$  ならば  $\mathcal{F}_{\Delta_1} \subset \mathcal{F}_{\Delta_2}$  が成り立つ. したがって,

$$\mathbb{E}[\mathbb{E}[Z|\Delta_2|\Delta_1] = \mathbb{E}[Z|\Delta_1]$$

を確かめよう.

$\Delta$  が確率空間を構成する分割 (最も細かい分割) であるとするれば,  $\Delta_2|\Delta$  がなりたつ. 任意の  $A \in \Delta_2$  について,  $A = \cup_i A_i$ ,  $A_i \in \Delta$ , なる互いに素な分割がある. 上の計算を一般化すると, 条件付き期待値の定義は  $P(A) \neq 0$ ,  $\omega \in A \in \Delta_2$  のとき,

$$\mathbb{E}[Z|\Delta_2](\omega) = \sum_i \frac{Z(A_i)P(A_i)}{P(A)}$$

とできる.  $P(A) = 0$  のときは,  $\mathbb{E}[Z|\Delta_2](\omega) = 0$  である. ただし,  $Z(A_i)$  は  $\omega \in A_i$  のときの  $Z(\omega)$  の値, 以下では  $\mathbb{E}[Z|\Delta_2](A)$  も同様の意味で使う. 可測関数は区分的な定数関数であったことに注意しよう.

$B \in \Delta_1$  に対して,  $\Delta_2$  の元による互いに素な分割  $B = \cup_j B_j$  が存在する. 各  $B_j \in \Delta_2$  には,  $\Delta$  の元による互いに素な分割  $B_j = \cup_i B_{ij}$  が存在するので,  $\omega \in B$  に対して

$$\begin{aligned} \mathbb{E}[\mathbb{E}[Z|\Delta_2|\Delta_1](B) &= \sum_j \frac{\mathbb{E}[Z|\Delta_2](B_j)P(B_j)}{P(B)} \\ &= \sum_j \frac{\sum_i \left[ \frac{Z(B_{ij})P(B_{ij})}{P(B_j)} \right] P(B_j)}{P(B)} \\ &= \sum_j \sum_i \frac{Z(B_{ij})P(B_{ij})}{P(B)}. \end{aligned}$$

$B = \cup_{i,j} B_{ij}$  は,  $B \in \Delta_1$  の  $\Delta$  の元による互いに素な分割になっているので,

$$\mathbb{E}[\mathbb{E}[Z|\Delta_2|\Delta_1](B) = \mathbb{E}[Z|\Delta_1](B)$$

が成り立つ.

#### A.4. 確率過程

前節では確率変数が情報を明らかにしていく様子を具体例をもって描写してみた。どの程度成功しているかは分からないが、我々はもっと先に進まなければならない。ここまでの話を逆転させてみよう。すなわち、しだいに細分化されていく分割の列

$$\{\Omega\} = \Delta_0 | \Delta_1 | \Delta_2 | \dots$$

あるいは、 $\sigma$ -集合族の増大列

$$\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \dots$$

が予め与えられていて、確率変数の列  $X_1, X_2, \dots$  がこれらの分割、 $\sigma$ -集合族に関して可測になっているという状況を考えるのである。  $\sigma$ -集合族の列  $\mathbb{F} = (\mathcal{F}_n)$  をフィルトレーションといい、 $X_n$  が  $\mathcal{F}_n$ -可測であるような確率変数の列を  $\mathbb{F}$ -適合確率過程という。

具体的なイメージとしては、コイントスの繰り返しを考えるとよい。ひとまず繰り返しの回数を3回としてみよう。  $\Omega = [0, 1)$  の分割列を

$$\begin{aligned}\Delta_0 &= \{\Omega\} \\ \Delta_1 &= \left\{ \left[0, \frac{1}{2}\right), \left[\frac{1}{2}, 1\right) \right\} \\ \Delta_2 &= \left\{ \left[0, \frac{1}{4}\right), \left[\frac{1}{4}, \frac{1}{2}\right), \left[\frac{1}{2}, \frac{3}{4}\right), \left[\frac{3}{4}, 1\right) \right\} \\ \Delta_3 &= \left\{ \left[0, \frac{1}{8}\right), \left[\frac{1}{8}, \frac{1}{4}\right), \dots, \left[\frac{3}{4}, \frac{7}{8}\right), \left[\frac{7}{8}, 1\right) \right\}\end{aligned}$$

と構成する。1回目が表であるという事象 ( $x_1 = 0$ ) を  $\omega \in \left[\frac{0}{2}, \frac{1}{2}\right)$ 、裏であるという事象 ( $x_1 = 1$ ) を  $\omega \in \left[\frac{1}{2}, 1\right)$  に対応させる。さらに1回目が  $x_1$  で2回目が表 ( $x_2 = 0$ ) あるいは裏 ( $x_2 = 1$ ) であるという事象を、 $\omega \in \left[\frac{x_1}{2} + \frac{x_2}{2^2}, \frac{x_1}{2} + \frac{x_2+1}{2^2}\right)$  に対応させる。1回目が  $x_1$  で、2回目が  $x_2$  で3回目が  $x_3$  である ( $x_3 = 0$  or  $1$ ) という事象を  $\omega \in \left[\frac{x_1}{2} + \frac{x_2}{2^2} + \frac{x_3}{2^3}, \frac{x_1}{2} + \frac{x_2}{2^2} + \frac{x_3+1}{2^3}\right)$  に対応させると、コイントスモデルの分割ができあがる。この分割列に適合した確率過程を作ることは難しくない。各ステップの分割にあわせて区分的定数関数を作ればよい。

適合過程を考察する理由は明白であろう。確率過程  $X_n(\omega)$  が  $n$  回目のコイントス後の所持金を表している場合を考えてみよ。

#### 参考文献

KLEIN, P. (2000): "Using the generalized Schur form to solve a multivariate linear rational expectations model," *Journal of Economic Dynamics & Control*, 24, 1405–1423.