

ある条件下における確率変数の期待値と その数値計算法

森 治憲, 新居 玄武

1. 問題の提起と定式化

企業が、自社製品のマーケット・シェアを、自社及び競合他社が投下した広告費に応じて予測するという問題を考えてみよう。ただし、過去のデータから、各製品のマーケット・シェアの分布と広告費との相関係数は推定されているものとする。一つの方法として考えられるのは、広告費を説明変数として離散選択モデルを適用することであるが、ここでは、シェアと広告費との関係を直接表現する相関係数を用いて予測する方法を考える。このシェア予測は、シェアの合計が1、さらに全製品の売上合計と広告費の相関係数を固定するという制約下での予測問題となる。そして、平均二乗誤差の意味で最適な予測値が条件付期待値ということから、所与の広告費の下で最適な予測値は、上記二つの条件を課したときの条件付期待値になる。

この予測問題は次のように定式化される。製品の個数を T 個 ($T \geq 4$) とし¹、各製品のシェアを T 次元確率変数ベクトル $X' = (X_1, \dots, X_T)$ で表す。対象とする企業のシェアは第1成分 X_1 とする。また、 X は同時密度関数 $f(x)$ を持つことを仮定する。

全成分が1の T 次元列ベクトルを e とすれば、シェアの合計が1という制約は、

$$e'X = 1$$

と表される。これを条件 I と呼ぶことにする。 T 個の製品のシェアと対応する広告費の相関係数を $r \in (-1, 1)$, $r \neq 0$ とする。相関係数は平行移動と定数倍に関して不変だから、各製品の広告費 $\mu' = (\mu_1, \dots, \mu_T)$ は、

$$\sum_{i=1}^T \mu_i = 0, \quad \sum_{i=1}^T \mu_i^2 = 1$$

と基準化されているものとする。 $ee' = E$ とすれば、シェアと広告費との相関係数が r という制約は、

$$\mu'X = r \sqrt{X' \left(I - \frac{1}{T} E \right) X}$$

と表すことができる。これを条件 II とする。

¹条件を満たす X を求めるという意味では $T \geq 2$ で十分であるが、解法を示すうえで最も一般的な状況は $T \geq 4$ の場合である。

本稿では、条件 I と II を与えたときの条件付期待値、

$$E\left(X_t \mid e'X=1, \mu'X=r, \sqrt{X'(I-\frac{1}{T}E)X}\right)$$

の導出について考察していく。

2. 数値計算のための公式

前節で定義した条件付期待値を解析的に求めることは一般に困難である。 X の確率分布が T 変量の一様分布に基づいて、条件を満たす乱数を発生させることができなければ、通常のモンテカルロ法による数値計算を適用することもできない。ここでは、実際に乱数を発生させることのできる確率分布を用いた数値計算の方法を説明する。

この節では、前述した確率モデルを一般化して議論していく。 R^T における T 次元开区間 I^T の部分集合から成る Borel 集合族を B とし、 P_t を Borel 空間 (I^T, B) で定義された確率測度とする。そして、 X を、Borel 空間 (I^T, B) 上の値をとる確率空間 (I^T, B, P_t) で定義された絶対連続型の確率変数ベクトルとする。 X の同時密度関数を $f(x)$ とし、

$$x \in I^T \Rightarrow f(x) > 0$$

を仮定する。ここで条件を、 m 個 ($m < T$) の Borel 可測関数 $C_j(x); j=1, \dots, m$ で定義された陰関数に一般化する。 C を関数ベクトルとして、条件を $C(x)=0$ と表すことにする。条件を満たす点 $x \in I^T$ の Borel 可測集合 $\Gamma = \{x \mid C(x)=0\} \cap I^T$ の部分集合から成る Borel 集合族を $B_\Gamma = \{A \mid A=B \cap \Gamma, B \in B\}$ と定義して、各 X_t の条件付期待値を $E_t(X_t \mid B_\Gamma)$ と書くことにする。

ここで、 X の各成分 X_t は独立に分布し、开区間 I^T 上の T 次元一様分布に従っている場合を考えよう。確率測度を P_0 とする。 X は Borel 空間 (I^T, B) 上の値をとるから、

$$E_t(X_t \mid B_\Gamma) = \int_\Gamma x_t dP_0(x \mid B_\Gamma); t=1, \dots, T$$

を満たす正則条件付確率 $P_\mu(\cdot \mid B_\Gamma)$ が存在する。正則条件付確率は確率測度だから、この正則条件付確率 $P_\mu(\cdot \mid B_\Gamma)$ を確率測度 U と表すことにする。

このとき、次の定理が成立する。

定理 2-1

$$E_t(X_t \mid B_\Gamma) = C \int_\Gamma x_t f(x) dU(x); t=1, \dots, T$$

ただし、 C は基準化定数である。

²ベクトル $x'=(x_1 \dots x_T)$ が与えられたとき、座標 (x_1, \dots, x_T) を持つ点を表すのに、同じ記号“ x ”を使うことにする。

証明は補足で与えられる。ここでは、確率測度 U が具体的に構成されていないため、条件付期待値を解析的に求めることはできない。しかし、この定理から数値計算のための公式が導かれる。

条件 $C(X) = \theta$ を満たす T 次元乱数を、 X の確率分布あるいは T 次元一様分布に基づいて発生できるとは限らない。そこで、新たに X を、実際に乱数を発生できる確率分布に従う絶対連続型の確率変数ベクトルとしよう。確率測度は P_g とする。同時密度関数は $g(x)$ とし、ここでも、

$$x \in I^T \Rightarrow g(x) > 0$$

を仮定する。条件 $C(X) = \theta$ を課したときの正則条件付確率を $P_g(\cdot | B_r)$ とすれば、次の定理が成立する。

定理 2-2

確率分布 $P_g(\cdot | B_r)$ に従う確率変数ベクトル $X'_n = (X_{1n}, \dots, X_{Tn})$; $n=1, 2, \dots$ は n に関して独立であるものとする。各 t について、

$$\frac{\sum_{n=1}^N X_{tn} \frac{f(X_n)}{g(X_n)}}{\sum_{n=1}^N \frac{f(X_n)}{g(X_n)}} \xrightarrow{p} E_t(X_t | B_r)$$

が成立する。

証明は補足で与えられる。以下の議論では、同時密度関数 g を元の同時密度関数 f と区別して加重関数³と呼ぶことにする。

条件を満たす T 次元乱数 X_n を発生させることができれば、この公式を用いて条件付期待値 $E_t(X_t | B_r)$ の一致推定量を求めることができる。この数値計算において定数は約分されるので、同時密度関数 f の基準化定数と、乱数 X_n を発生させることができれば、加重関数 g の基準化定数は不要である。

3. 加重関数の導出

第一節で提起した問題に議論を戻そう。ただし、確率モデルを次のように拡張して考えていくことにする。まず、 X の定義域を T 次元開区間、

$$I^T = (m_1, M_1) \times \dots \times (m_T, M_T)$$

³モンテカルロ法における加重サンプリングから名前をとった。条件付期待値の導出方法は加重サンプリングと同じ考え方である。しかし、加重サンプリングは推定量の分散を減少させるために用いられる。

とする。シェアの合計が1という条件Iは、 T 次元列ベクトル w の各成分を重みとした加重平均が u という条件、

$$w'x = u$$

に一般化する。この条件を改めて条件Iとする。ただし、 $\sum w_i = 1, w_i \geq 0$ は当然だが、基準化された広告費 μ に関して、

$$\mu = aw + be$$

を満たす実数 a, b は存在しないことを仮定する。

条件Iを満たす点 x の集合 Γ_1 と条件IIを満たす点 x の集合 Γ_2 により、すべての条件を満たす点 x の集合を $\Gamma = \Gamma_1 \cap \Gamma_2 \cap I^T$ と表すことにする。ここで、点 $u = (u, \dots, u)$ を始点とし、点 $x \in R^T$ を通る直線を、

$$l(x) = \{z \mid z = a(x - ue) + ue, a > 0\}$$

と定義すれば、

$$\Gamma_1 \cap \Gamma_2 = \bigcup_{x \in \Gamma_1 \cap \Gamma_2} l(x)$$

であることが容易に示される。次の関係は明らかである。

$$\Gamma_1 \cap \Gamma_2 = \bigcup_{x \in R^T} l(x)$$

この包含関係から、まず、点 u を始点とする直線 $l \in \Gamma_1 \cap \Gamma_2$ を選び、次に、直線 l 上の点 $x \in \Gamma$ を選ぶという方法で T 次元乱数 $x \in \Gamma$ を求められることが分かる。そこで、本稿では数値計算を容易にするため、

- i) 点 u を始点とする線分 $L \subset I^T$ には等確率が定義される。
- ii) その線分 L 上の点 x には等確率が定義される。

を満たす加重関数 $g(x)$ を提案する。

開区間 I^T の境界を ∂I^T とし、点 u が始点で点 $x \in I^T$ を通る直線 $l(x)$ の境界 ∂I^T までの部分を線分 $L(x)$ とする。この線分の長さ $|L(x)|$ により、加重関数を、

$$g(x) = \frac{C}{|L(x)|}, x \in I^T - \{u\}$$

と定義する。前節で述べたように、基準化定数 C を求める必要はない。この加重関数が前出した二つの要請を満たすことは補足で示される。

線分 $L(x)$ は直線 $l(x)$ の境界 ∂I^T までの部分だから、 $l(x)$ が ∂I^T と交わるときの係数 a が分かれば、線分 $L(x)$ の長さ $|L(x)|$ を求めることができる。このときの係数 a の値は；

- i) $x_i > u$ となる x_i の中で $z_i = M_i$ となる係数 a の最小値
- ii) $x_i < u$ となる x_i の中で $z_i = m_i$ となる係数 a の最小値

のうち小さい方である。この値を $x \in I^T - \{u\}$ の関数として $A(x) > 0$ とすれば、

$$\begin{aligned}
 T^{(+)} &= \{t | x_t > u; t=1, \dots, T\} \\
 T^{(-)} &= \{t | x_t < u; t=1, \dots, T\} \\
 A^{(+)} &= \min \left\{ A_t \left| A_t = \frac{M_t - u}{x_t - u}, t \in T^{(+)}(x) \right. \right\} \\
 A^{(-)} &= \min \left\{ A_t \left| A_t = \frac{M_t - u}{x_t - u}, t \in T^{(-)}(x) \right. \right\} \\
 A(x) &= \min \{A^{(+)}(x), A^{(-)}(x)\}
 \end{aligned}$$

である。直線 $l(x)$ と境界 ∂I^T との交点 z の座標は、

$$z_t = A(x)(x_t - u) + u; t = 1, \dots, T$$

だから、点 u と点 z を結ぶ線分 $L(x)$ の長さは、

$$|L(x)| = A(x) \sqrt{\sum (x_t - u)^2}$$

となる。

4. 乱数発生の方法

T 次元乱数を発生させる方法は、

- i) 点 u を始点とする線分 $LC \Gamma$ を等確率で選ぶ。
- ii) 線分 L 上の点 x を等確率で選ぶ。

という二つのプロセスに分けて考えることができる。結論から先に述べると、この論文では、 T 変量正規分布 $N(ue, I)$ を用いて線分 $LC \Gamma$ を選ぶ方法を提案する。このプロセスにおける確率モデルは、正規乱数を乱数 X と区別して Y とすれば、

$$\begin{aligned}
 Y &\sim N(ue, I) \\
 \text{s.t. } w'Y &= u, \mu'Y = r \sqrt{Y'(I - \frac{1}{T}E)Y}
 \end{aligned}$$

である。

正規分布 $N(ue, I)$ を用いるのは、点 u を始点とするすべての直線に等しい確率が定義されるからである。したがって、乱数 y から直線 $l(y)$ を求めたということは、点 u を通るすべての直線 $l \subset \Gamma_1 \cap \Gamma_2$ の中から直線 $l(y)$ を等確率で選んだことと同値である。当然、直線 $l(y)$ から線分 $L(y) \subset \Gamma$ を求めれば、線分 $L(y)$ も点 u を通るすべての線分 $LC \Gamma$ の中から等確率で選ばれたことになる。もちろん、点 $y' \in l(y)$ を選んでも、線分 $L(y) = L(y')$ が選ばれるという結果に変わりはない。ここから、線分 $LC \Gamma$ を求めるのに必要なのは、乱数 y の値というより、直線 $l(y)$ の方向比、

$$y_1 - u : y_2 - u : \dots : y_T - u$$

であることがわかる¹。

¹このように、すべてを「ランダム」に乱数 y を求める必要はない。7-2 参照。

Y の乱数を発生させる具体的な方法は次節以降で説明するので、ここでは、乱数 $Y=y$ が得られたものとして、点 y から線分 $L(y)$ 上の点 x を等確率で選ぶ二番目の方法を説明する。前節で定義した関数 $A(y)$ を使うと、線分 $L(y)$ は、

$$L(y) = \{x \mid x = a(y - ue) + ue, 0 < a < A(y)\}$$

と表される⁵。したがって、係数 $A=a$ を一様分布 $A \sim U(0, A(y))$ に従う乱数として求めれば、線分 $L(y)$ 上の点 x を等確率で選ぶことができる。一様乱数 $U \sim U(0, 1)$ を、

$$A = A(y) U$$

と逆変換すれば、係数 a は乱数として求められる。最終的に、すべての条件を満たす T 次元乱数 $x \in \Gamma$ は、

$$x = a(y - ue) + ue$$

となる。

ところで、線分 $L \subset \Gamma$ を求める最初の方法では、確率モデルは正規分布 $N(ue, I)$ だから加重関数 g を必要としない。線分 L 上の点 x を選ぶ二番目の方法でも、確率モデルは一様分布 $U(0, A(y))$ だから加重関数 g は用いられない。ここから、乱数 $x \in \Gamma$ を発生させるのに、加重関数の基準化定数 C は不要であることが分かる。

なお、正規乱数 $Y \in \Gamma_1 \cap \Gamma_2$ を発生させる方法は、

- i) Step 1 ~ Step 5 : 確率モデルの変換
- ii) Step 6 ~ Step 8 : 乱数の発生
- iii) Step 9 : 変換に伴う調整

という構造になっている。したがって、乱数を発生させるアルゴリズムは、

- 「条件を満たす乱数の発生 : Step 6 → Step 8」
- ⇒ 「元の確率モデルの乱数へ逆変換 : Step 5 → Step 1」
- ⇒ 「変換に伴う調整 : Step 9」

となる。詳細は次節以降で説明する。

5. 加重平均に関する条件 I の消去 : Step 1 ~ Step 3

5-1 Step 1

条件 II の両辺を二乗して、 Y の 2 次形式で表された条件に拡張する。Step 8 まで、この条件が用いられる。行列 R を、

$$R = \mu\mu' - r^2 \left(I - \frac{1}{T} E \right)$$

⁵ $y \in I^T$ とは限らないが、線分に $L(y)$ に関する前節の議論はそのまま成立する

と置けば, 確率モデルは次のようになる。

$$Y \sim N(ue, I) \\ \text{s.t. } w'Y = u, Y'RY = 0$$

5-2 Step 2

$Z = Y - ue$ と変数変換すれば, $w'e = 1$ だから, 加重平均に関する条件 I は,

$$w'Y = u \Leftrightarrow w'Z = w'(Y - ue) = 0$$

となる。 Y の各成分から定数 u を引いても相関係数は変わらないから, 二次形式に関する条件は,

$$Y'RY = 0 \Leftrightarrow Z'RZ = 0$$

である。 $Z \sim N(0, I)$ だから, Z を改めて Y と置けば, この Y に関する確率モデルは次のようになる。

$$Y \sim N(0, I) \\ \text{s.t. } w'Y = 0, Y'RY = 0$$

5-3 Step 3

条件 I の下では確率 1 で線型関係 $w'Y = 0$ が成立するため, Y の条件付同時密度関数は存在しない。まず, 密度関数を用いない多変量正規分布の定義を最初に述べておく⁶。

定義

p 次元確率変数 Y は, $Z \sim N(0, I_m)$ により $Y = \mu + GZ$ と表現できるならば, p 変量正規分布 $N(\mu, \Sigma)$ に従うという。ただし, $\Sigma = GG'$ であり, G は階数が m の $(p \times m)$ 型行列である。

このように多変量正規分布を定義すると, 条件 I を課したときの Y の条件付分布を求めることができる。また, 行列 A の列ベクトルが張るベクトル空間を $V(A)$ とし, その直交補空間を $V^\perp(A)$ と書くことにする。

定理 5-1

条件 $w'Y = 0, \|w\| = 1$ を与えたときの, $Y \sim N(0, I_T)$ の条件付分布は,

$$Y | w'Y = 0 \sim N(0, I - ww')$$

である。ただし, $\text{rank}(I - ww') = T - 1$ である。

(証明) $G: T \times (T - 1)$ を $V^\perp(w)$ の正規直交基底を列ベクトルとした行列とする。 Y を直交行

⁶Rao(1973, 8 a. 2)参照。

列 $P=(G \ w)$ により,

$$Y=PZ=(G \ w) \begin{pmatrix} Z_1 \\ Z_T \end{pmatrix}$$

と直交変換する。 $Z \sim N(0, I_T)$ だから、 $Z_T = w' Y = 0$ の条件としたときの、 Z_1 の条件付分布は $N(0, I_{T-1})$ となる。

ところで、条件 $w' Y = 0$ の下では、確率 1 で $Y \in V^\perp(w)$ であり、確率 1 で、

$$Y = GZ_1$$

が成立する。したがって、多変量正規分布の定義から、

$$Y | w' Y = 0 \sim N(0, GG')$$

であることが分かる。行列 G は $V^\perp(w)$ の正規直交基底を列ベクトルとした行列だから一意ではない。しかし、

$$\begin{aligned} PP' &= (G \ w) \begin{pmatrix} G' \\ w' \end{pmatrix} = GG' + ww' = I \\ \Rightarrow GG' &= I - ww' \end{aligned}$$

であるから、行列 $GG' = I - ww'$ は行列 G の選び方に依存しない。その階数が $(T-1)$ であることは明らか。以上により、

$$Y | w' Y = 0 \sim N(0, I - ww')$$

が示された。(証明終)

定理 5-1 を適用するため、 $w / \|w\|$ を新しく w と書くことにする。この段階で、確率モデルは条件 I が消去された構造となる。

$$\begin{aligned} Y &\sim N(0, I - ww') \\ \text{s.t. } Y'RY &= 0 \end{aligned}$$

6. 二次形式の標準化：Step4～Step5

6-1 Step4

直交行列 $P=(G \ w)$ を用いて、 $Y \sim N(0, I - ww')$ を $Y=PZ$ と直交変換すると、 Z の同時分布は、

$$\begin{aligned} E(Z) &= P'E(Y) = 0 \\ \text{Var}(Z) &= \begin{pmatrix} G' \\ w' \end{pmatrix} GG' (G \ w) = \begin{pmatrix} I_{T-1} \\ 0 \end{pmatrix} (I_{T-1} \ 0) = \begin{pmatrix} I_{T-1} & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

の T 変量正規分布となる。各 Z_i は独立であり、 Z_T は確率 1 でゼロだから、以下の議論では $Z \sim N(0, I_{T-1})$ だけ考えれば十分である。そこで、新しく $Y=GZ$ と置けば、確率モデルは次のようになる。

$$\begin{aligned} Z &\sim N(0, I_{T-1}) \\ \text{s.t. } Z'G'RGZ &= 0 \end{aligned}$$

6-2 Step 5

$\alpha = G'\mu, \beta = G'\tau, \tau = e/\sqrt{T}$ と置く。行列 $G'RG$ の固有多項式は、

$$\begin{aligned} |xI - G'RG| &= |xI - G'\mu\mu'G + r^2G'(I - \frac{1}{T}E)G| \\ &= |(x+r^2)I - (\alpha\alpha' + r^2\beta\beta')| \end{aligned}$$

だから、その固有値は $\alpha\alpha' + r^2\beta\beta'$ の固有値から r^2 を引いた値である。まず、二つの補題を述べておく。

補題 6-1

$$\alpha = G' \mu \neq 0$$

(証明) 行列 G の列ベクトルは $V^\perp(w)$ の正規直交基底だから、

$$\alpha = 0 \Leftrightarrow \mu \in V(w)$$

である。すなわち、 $\alpha = 0$ と $\mu = aw$ を満たす実数 a が存在することは同値である。しかし、 $\mu = aw + be$ を満たす実数 a, b は存在しないから、 $\alpha \neq 0$ である。(証明終)

補題 6-2

$$\text{rank}(\alpha\alpha' + r^2\beta\beta') = 2$$

(証明) $\text{rank}(\alpha\alpha' + r^2\beta\beta') = 2$ となる必要十分条件は、

$$V(\alpha\alpha') \cap V(\beta\beta') = V(\alpha) \cap V(\beta) = \{0\}$$

が成立することである⁷。補題 6-1 より $\alpha \neq 0$ だから、 $V(\alpha) \cap V(\beta) = \{0\}$ とは $\alpha = d\beta$ を満たす実数 d が存在しないことと同値である⁸。その否定は、

$$\alpha = d\beta \Leftrightarrow G'\mu = dG'\tau \Leftrightarrow G'(\mu - be) = 0, \quad b = d/\sqrt{T}$$

である。ここで、 $V^\perp(G) = V(w)$ だから、 $\mu - be \in V^\perp(G)$ ということは、 $\mu = aw + be$ となる実数 a, b が存在することと同値である。すなわち、 $\alpha = d\beta$ を満たす d が存在しないことと、 $\mu = aw + be$ となる a, b が存在しないことは同値である。(証明終)

階数 2 の行列 $\alpha\alpha' + r^2\beta\beta'$ は非負値定符号行列だから、この行列は重複度 $(T-3)$ のゼロと二つの正の値を固有値として持つ。正の固有値を $\lambda_2 < \lambda_1$ すれば、 $G'RG$ の固有値は重複度 $(T-3)$ の $-r^2$ と $\lambda_1 - r^2$, $\lambda_2 - r^2$ となる。対応する固有ベクトルは $\alpha\alpha' + r^2\beta\beta'$ の固有ベクトルと等しいから、 $\alpha\alpha' + r^2\beta\beta'$ を対角化する $(T-1)$ 次直交行列は、 $G'RG$ を対角化することができる。こ

⁷Rao(1973, p.31)。

⁸ $\alpha = \beta = 0$ の場合、任意の実数 d は $\alpha = d\beta$ を満たす。

の直交行列を Q として, $Z \sim N(0, I_{T-1})$ を $Z=QX$ と直交変換すれば, 二次形式に関する条件は,

$$X'Q'G'RGQX=(\lambda_1-r^2)X_1^2+(\lambda_2-r^2)X_2^2-r^2(X_3^2+\cdots+X_{T-1}^2)=0$$

と標準化される。ここで, $X \sim N(0, I_{T-1})$ だから, 各 X_i^2 は独立にカイ 2 乗分布 $\chi^2(1)$ に従う。定理 6-1 で示すように, 制約式の係数について,

$$\lambda_1 - r^2 > 0, \lambda_2 - r^2 < 0$$

が成立する。そこで, $A = \frac{\lambda_1}{r^2} - 1, B = 1 - \frac{\lambda_2}{r^2}$ と置けば, 確率モデルは,

$$\begin{aligned} X_i^2, X_i^2 &\sim \chi^2(1); \quad i=1, \dots, T-1 \\ \text{s.t. } AX &= B X_2^2 + X_3^2 + \cdots + X_{T-1}^2, \quad A > B > 0 \end{aligned}$$

と表すことができる。

定理 6-1

$$0 < \lambda_2 < r^2 < \lambda_1$$

(証明) 固有値と主座小行列式との関係から,

$$\begin{aligned} \lambda_1 + \lambda_2 &= \text{tr}(\alpha\alpha' + r^2\beta\beta') \\ \lambda_1\lambda_2 &= \sum_{s < t} |(\alpha\alpha' + r^2\beta\beta')_{st}| \end{aligned}$$

が成立する。ただし,

$$\begin{aligned} (\alpha\alpha' + r^2\beta\beta')_{st} &= \begin{pmatrix} \alpha_s^2 + r^2\beta_s^2 & \alpha_s\alpha_t + r^2\beta_s\beta_t \\ \alpha_t\alpha_s + r^2\beta_t\beta_s & \alpha_t^2 + r^2\beta_t^2 \end{pmatrix} \\ |(\alpha\alpha' + r^2\beta\beta')_{st}| &= r^2(\alpha_s^2\beta_t^2 + \alpha_t^2\beta_s^2 - \alpha_s\alpha_t\beta_s\beta_t) \end{aligned}$$

である。 $\lambda_1\lambda_2$ の総和は $s < t$ となる s と t について定義されているが, 形式的に $s=t$ の場合も考えれば,

$$s=t \Rightarrow |(\alpha\alpha' + r^2\beta\beta')_{st}| = 0$$

だから, $\lambda_1\lambda_2$ は次のように表すことができる。

$$\begin{aligned} \lambda_1\lambda_2 &= \sum_{s < t} |(\alpha\alpha' + r^2\beta\beta')_{st}| = \frac{1}{2} \sum_{s=1}^{T-1} \sum_{t=1}^{T-1} |(\alpha\alpha' + r^2\beta\beta')_{st}| \\ &= r^2 \{ \|\alpha\|^2 \|\beta\|^2 - (\alpha'\beta)^2 \} \end{aligned}$$

この式と,

$$\lambda_1 + \lambda_2 = \text{tr}(\alpha\alpha' + r^2\beta\beta') = \|\alpha\|^2 + r^2\|\beta\|^2$$

を, λ_2 について整理すると,

$$\lambda_2^2 - (\|\alpha\|^2 + r^2\|\beta\|^2)\lambda_2 + r^2\{\|\alpha\|^2\|\beta\|^2 - (\alpha'\beta)^2\} = 0$$

となる。この式で λ_2 を x として, 左辺を $f(x)$ と表すことにする。

$0 < \lambda_2 < \lambda_1$ であり, Cauchy-Schwarz の不等式より $f(0) \geq 0$ だから, $f(r^2) < 0$ を示すことができる。 $0 < \lambda_2 < r^2 < \lambda_1$ であることが分かる。 $f(r^2)$ を r^2 の関数,

$$f^*(r^2) = f(r^2) = (1 - \|\beta\|^2)r^4 + \|\alpha\|^2\|\beta\|^2 - (\alpha'\beta)^2 - \|\alpha\|^2r^2$$

と考えると, $\|\beta\|^2 < 1$ より r^4 の係数は正であり, $f^*(0) = 0$ だから, $f^*(1) = 0$ が成立すれば, $r^2 \in (0, 1)$ に対して,

$$f^*(r^2) < 0 \Rightarrow f'(r^2) < 0$$

と結論することができる。そこで,

$$\begin{aligned} 1 - \|\alpha\|^2 &= \mu'\mu - \mu'GG\mu = \mu'ww'\mu \quad \because \mu'\mu = 1 \\ 1 - \|\beta\|^2 &= \tau'ww'\tau \quad \because \tau'\tau = 1 \\ (\alpha'\beta)^2 &= \mu'ww' \times \tau'ww'\tau \quad \because \mu'\tau = 0 \end{aligned}$$

を $f^*(1) = (1 - \|\alpha\|^2)(1 - \|\beta\|^2) - (\alpha'\beta)^2$ に代入すれば,

$$f^*(1) = \mu'ww'\mu \times \tau'ww'\tau - \mu'ww'\mu \times \tau'ww'\tau = 0$$

であることが分かる。

以上から, 行列 $\alpha\alpha' + r^2\beta\beta'$ の正の固有値 $\lambda_2 < \lambda_1$ について,

$$0 < \lambda_2 < r^2 < \lambda_1$$

であることが示された。(証明終)

7. 具体的な乱数発生方法: Step 6 ~ Step 9

7-1 Step 6

以下で示すように, すべての X_i^2 を乱数として求める必要はなく, $(T-2)$ 個だけ求めれば十分である。ここでは恒等的に,

$$X_i^2 \equiv 1$$

と置くことにする。

この乱数発生のプロセスで必要なのは, 条件を満たす正規乱数 $Y=y$ ではなく, そこから得られる直線 $l(y)$ の方向比,

$$y_1 - u : y_2 - u : \cdots : y_T - u$$

であった。Step 2 で $Y - ue$ を新しく Y と置いたことと, Step 4 と Step 5 の変換が線形写像であることから, ここで必要なのは, 乱数 x の成分比,

$$x_1 : x_2 : \cdots : x_{T-1}$$

ということになる。このように, X_i の一つを固定しても問題はない。以後の計算が簡単になるのは, $|X_i| = 1$ と置くことである。絶対値が必要なのは, 条件 II,

$$\mu'Y = r \sqrt{Y'(I - \frac{1}{T}E)Y}$$

⁹ 第 8 節 (Step 9) 参照。

を満たす乱数 y を得るためには、乱数 x_i^2 から $\pm\sqrt{x_i^2}$ を決めなければならないが、必ずしも $x_1 = 1$ が成立するとは限らないからである⁹。

$|X_1| = 1$ とは、Step 6 の変換で $X_1^2 \equiv 1$ と置くことに該当する。以上が、 $X_1^2 \equiv 1$ とする理由である。

7-2 Step 7

新しく確率変数を、

$$\begin{aligned} V &= X_2^2 \sim \chi^2(1) \\ W &= X_3^2 + \dots + X_{T-1}^2 \sim \chi^2(k), \quad k=T-3 \end{aligned}$$

と定義すれば、条件は $A=BV+W$ と書くことができる。 V と W は独立に分布しているから、その同時密度関数は、

$$f_{V,W}(v,w) \propto e^{-\frac{1}{2}(v+w)} v^{-\frac{1}{2}} w^{\frac{k}{2}-1}$$

となる。条件 $A=BV+W$ を与えたときの $V \in (0, A/B)$ の条件付密度関数 h は、定義式の分子、

$$f_{V,W}(v, A-Bv) \propto e^{-\frac{1}{2}(1-B)v} v^{-\frac{1}{2}} (A-Bv)^{\frac{k}{2}-1}$$

に比例する。基準化定数を C とおけば、条件付密度関数 h は、

$$\begin{aligned} h(v|A=BV+W) &= Ce^{-\frac{1}{2}(1-B)v} v^{-\frac{1}{2}} (A-Bv)^{\frac{k}{2}-1} \\ &= Ce^{-\frac{1}{2}(1-B)v} v^{-\frac{1}{2}} (A-Bv)^{\frac{T-5}{2}} \end{aligned}$$

となる。この式の指数関数部分は明らかに有界であり、残りの部分は开区間 $(0, A/B)$ 上のベータ分布の密度関数に比例しているから、条件付密度関数 h は定義域 $(0, A/B)$ で確かに積分可能である。

この条件付密度関数から分布関数 $H(v|A=BV+W)$ を求めれば、一様乱数を逆変換することで、乱数 $^{10}V=v$ を求めることができる。乱数 V が決まれば、乱数 $W=A-BV$ も一意に定められる。まとめると、Step 7 では、

$$\begin{aligned} V &= X_2^2 = v \\ W &= X_3^2 + \dots + X_{T-1}^2 = A-Bv \end{aligned}$$

が乱数として求められる。

7-3 Step 8

条件 $X_3^2 + \dots + X_{T-1}^2 = A-Bv$ を満たす $X_t^2 \sim \chi^2(1); t=3, \dots, T-1$ を求めるには、所与の定数 $Z > 0$ に関する条件、

$$Z = X + Y, \quad X \sim \chi^2(1), \quad Y \sim \chi^2(k)$$

¹⁰ $X_1^2 \equiv 1$ だから厳密には乱数ではないが、便宜上「乱数」と呼ぶことにする。

を満たす乱数 X, Y を選ぶ方法を説明すれば十分である。

X と Y の同時密度関数を f とする。まず、条件 $Z = X + Y$ を与えたときの X の条件付分布における条件付密度関数の分子は、

$$f_{X,Y}(x, Z-x) \propto e^{-\frac{x}{2}} e^{-\frac{Z-x}{2}} x^{-\frac{1}{2}} (Z-x)^{\frac{k}{2}-1} \propto x^{-\frac{1}{2}} (Z-x)^{\frac{k}{2}-1}$$

だから、区間 $(0, Z)$ で定義された母数 $\frac{1}{2}, \frac{k}{2}$ のベータ分布となる。そこで、 $\theta = X/Z$ の条件付密度関数を $g(\theta | Z = X + Y)$ とすれば、

$$g(\theta | Z = X + Y) \propto \theta^{-\frac{1}{2}} (1-\theta)^{\frac{k}{2}-1}$$

となり、その条件付分布は通常のベータ分布 $B(\theta | 1/2, k/2)$ となる。

一様乱数から乱数 $\theta = X/Z \Leftrightarrow X = Z\theta$ を求めれば、ここから乱数 $Y = Z - X$ も求めることができる。この手続きを繰り返すことで、条件を満たす乱数 X_1^2, \dots, X_{T-1}^2 を求めればよい。

8. 相関係数に関する調整: Step 9

Step 6 ~ Step 8 で求めた各 X_i^2 の平方根 $X_i = \pm \sqrt{X_i^2}$ をどのように選んでも、逆変換して求めた T 次元乱数 Y は条件 $Y'RY = 0$ を満たす。しかし、内積 $\mu'Y$ の符号から、 Y が相関係数 r に関する条件 II,

$$\mu'Y = r \sqrt{Y'(I - \frac{1}{T}E)Y}$$

を満たすとは限らない。この節では、各 X_i^2 から条件を満たす乱数 Y に変換する方法を説明する。

Step 6 ~ Step 8 では $X \sim N(0, I_{T-1})$ だから、条件 $X_i^2 = v_i$ の下では、

$$P(X_i = \sqrt{v_i} | X_i^2 = v_i) = P(X_i = -\sqrt{v_i} | X_i^2 = v_i) = \frac{1}{2}; i = 1, \dots, T-1$$

が成立する。したがって、各 i について、確率 $1/2$ で $x_i = \pm \sqrt{v_i}$ を決めればよい。これらの値を $\mathbf{x}^{(+)} = (x_1, \dots, x_{T-1})$ とすれば、 $\mathbf{X}^{(+)} \sim N(0, I_{T-1})$ であり、条件、

$$A\mathbf{X}_1^2 = B\mathbf{X}_2^2 + \mathbf{X}_3^2 + \dots + \mathbf{X}_{T-1}^2$$

を満たす。 $\mathbf{x}^{(-)} = -\mathbf{x}^{(+)}$ と定義すれば、 $\mathbf{x}^{(-)}$ は $\mathbf{x}^{(+)}$ と一対一に対応しているから、 $\mathbf{x}^{(-)}$ もまた条件を満たす正規乱数と考えることができる。そこで、 $\mathbf{x}^{(+)}, \mathbf{x}^{(-)}$ から Step 2 の T 次元乱数 $\mathbf{y}^{(+)}, \mathbf{y}^{(-)}$ を求めれば、これらの間にも、

$$\mathbf{y}^{(-)} = -\mathbf{y}^{(+)}$$

が成立する。そして、いずれか一方が相関係数に関する条件 II を満たす。

定義から明らかなように、 $\mathbf{y}^{(+)}, \mathbf{y}^{(-)}$ は乱数として無差別である。したがって、条件 II を満たす方を、Step 2 の乱数 \mathbf{y} とすればよい。

補足

定理 2-1 の証明

$Y=C(X)$ とし、その定義域 $C(I^T)$ の部分集合から成る Borel 集合族を \hat{B} とする。任意の集合 $A \in \hat{B}$ に対して、二つの確率測度を、

$$\begin{aligned}\hat{P}_u(A) &= P_u(x|y=C(x) \in A) \\ \hat{P}_f(A) &= P_f(x|y=C(x) \in A)\end{aligned}$$

と定義する。 T 次元 Lebesgue 測度を μ とすれば、条件付期待値の定義から、

$$\begin{aligned}\int_A E_f(X_i|Y=y) d\hat{P}_f(y) &= \int_{\{x|y \in A\}} x_i dP_f(x) \\ &= \int_{\{x|y \in A\}} x_i f(x) |I^T| \frac{1}{|I^T|} d\mu(x) \\ &= |I^T| \int_A E_u(X_i, f(x) | Y=y) d\hat{P}_u(y)\end{aligned}$$

が成立する。ここで、

$$\hat{P}_u(A) = 0 \Rightarrow P_u(x|y \in A) = \frac{1}{|I^T|} \mu(\{x|y \in A\}) = 0 \Leftrightarrow \mu(\{x|y \in A\}) = 0$$

だから、 $\hat{P}_u(A) = 0$ ならば、

$$\hat{P}_f(A) = \int_{\{x|y \in A\}} dP_f(x) = \int_{\{x|y \in A\}} f(x) d\mu(x) = 0$$

である。このように \hat{P}_f は \hat{P}_u に関して絶対連続だから、Radon-Nikodym の定理から、

$$\hat{P}_f(A) = \int_A h(y) d\hat{P}_u(y)$$

となる密度関数 h が存在する。そこで、 $d\hat{P}_f(y) = h(y) d\hat{P}_u(y)$ を条件付期待値の定義式に代入すれば、

$$\int_A E_f(X_i|Y=y) h(y) d\hat{P}_u(y) = |I^T| \int_A E_u(X_i, f(X) | Y=y) d\hat{P}_u(y)$$

となるから、

$$E_f(X_i|Y=y) h(y) = |I^T| \int_A E_u(X_i, f(X) | Y=y) \quad a.s.$$

であることが分かる。

\hat{P}_u に関する \hat{P}_f の密度関数 h は、仮定より $f(x) > 0$ だから、 m 次元 Lebesgue 測度に関してほとんどいたる所 $h(y) > 0$ である。そこで、 $C = |I^T|/h(0)$ とすれば、 $Y=0$ という条件は条件 $C(X) = 0$ のことだから、

$$E_f(X_i|B_f) = C E_u(X_i, f(X)|B_f)$$

が成立する。右辺の条件付期待値を定義する確率測度は U だから、

$$E_u(X_i, f(X)|B_f) = \int_{\Gamma} x_i f(x) dU(x)$$

と書くことができる。以上の結果は各 X_i について成立する。したがって、

$$E_f(X|B_f) = C \int_{\Gamma} x, f(x) dU(x); t=1, \dots, T$$

が成立する。

定理 2-2 の証明

ある条件下における確率変数の期待値とその数値計算法 (森, 新居)

定理 2-1 から, 基準化定数を C' とすれば, 任意の集合 $A \in B_T$ に対して,

$$P_g(A|B_T) = C' \int_A g(x) dU(x)$$

が成立する。したがって, 条件付期待値 $E_i(X_i|B_T)$ は,

$$\begin{aligned} E_i(X_i|B_T) &= C' \int_T x_i f(x) dU(x) \\ &= \frac{C}{C'} \int_T x_i \frac{f(X)}{g(X)} C' g(x) dU(x) = \frac{C}{C'} \int_T x_i \frac{f(X)}{g(X)} dP_g(X|B_T) \end{aligned}$$

と表すことができる。ここから,

$$E_i(X_i|B_T) = \frac{C}{C'} E_g \left(X_i \frac{f(X)}{g(X)} \middle| B_T \right)$$

であることが分かる。 $X_n; n=1, 2, \dots$ を n に関して独立な, 確率分布 $P_g(\cdot|B_T)$ に従う確率変数ベクトル列とすれば, 大数の法則から,

$$\frac{1}{N} \sum_{n=1}^N X_n \frac{f(X_n)}{g(X_n)} \xrightarrow{p} \frac{C'}{C} E_i(X_i|B_T)$$

が成立する。

右辺の基準化定数の比も同様にして求めることができる。まず, 次の関係式,

$$1 = E_i(1|B_T) = C' \int_T x_i f(x) dU(x) = \frac{C}{C'} \int_T \frac{f(X)}{g(X)} dP_g(x|B_T)$$

が成立する。したがって, 同じ確率変数ベクトル列 $X_n; n=1, 2, \dots$ によって,

$$\frac{1}{N} \sum_{n=1}^N X_n \frac{f(X_n)}{g(X_n)} \xrightarrow{p} \frac{C'}{C}$$

と表すことができる。二つの式の比を取れば, 確率収束の性質から, その値が条件付期待値 $E_i(X_i|B_T)$ に確率収束することが分かる。すなわち,

$$\frac{\sum_{n=1}^N X_n \frac{f(X_n)}{g(X_n)}}{\sum_{n=1}^N \frac{f(X_n)}{g(X_n)}} \xrightarrow{p} E_i(X_i|B_T)$$

となる。この式は各 X_i について成立する。

加重関数が二つの要請を満たすことの証明

点 u を始点とする線分 LCI' 上の点 x に等確率が定義されていることは, $g(x)$ の定義から明らかである。

次に, 点 u を始点とする線分 LCI' に等確率が定義されることを示す。 R^T の直線に定義される 1 次元 Lebesgue 測度を ν とし, R^T の $(T-1)$ 次元 Lebesgue 測度を S と表すことにする。そこで, $S(A_1) = S(A_2)$ を満たす集合 $A_1, A_2 \subset \partial I'$ をとる。このとき, 点 u が始点で終点が A_1 に含まれる

線分 $L \subset I'$ の集合 $B_1 = \{L \mid L=L(z), z \in A_1\}$ と、終点が A_2 に含まれる線分の集合 B_2 について、

$$P(B_1) = P(B_2)$$

が成立すれば、すべての線分に等確率が定義されているということができる。加重関数の定義から、 $x \in L(z)$ ならば、 $|L(x)| = |L(z)|$ であることに注意すると、

$$\begin{aligned} P(B_1) &= \int_{\{L \mid L=L(z), z \in A_1\}} \frac{C}{|L(X)|} d\mu(x) = \int_{A_1} \int_{L(z)} \frac{C}{|L(X)|} dv(x) dS(z) \\ &= CS(A_1) = CS(A_2) \\ &= P(B_2) \end{aligned}$$

となる。このように、最初の要請も満たされることが示された。

参考文献

- [1] 小山昭雄, 経済数学教室 2, 岩波書店, (1994)
- [2] 森正武, 室田一雄, 杉原正顕, 数値計算の基礎 (岩波講座応用数学), 岩波書店, (1993)
- [3] Breiman, L., *Probability*, Society for Industrial and Applied Mathematics, (1992)
- [4] Rao, C.R., *Linear Statistical Inference and Its Applications*, 2nd ed, John Wiley, (1973)