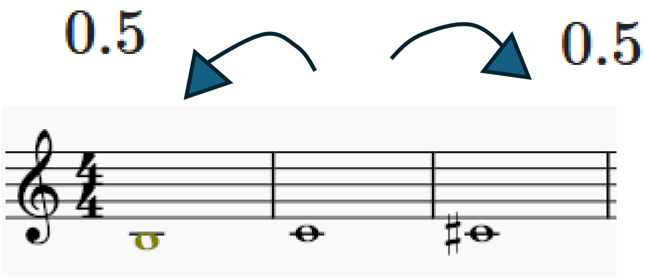
作曲の数理(ほぼマルコフ過程の話)

N.Y

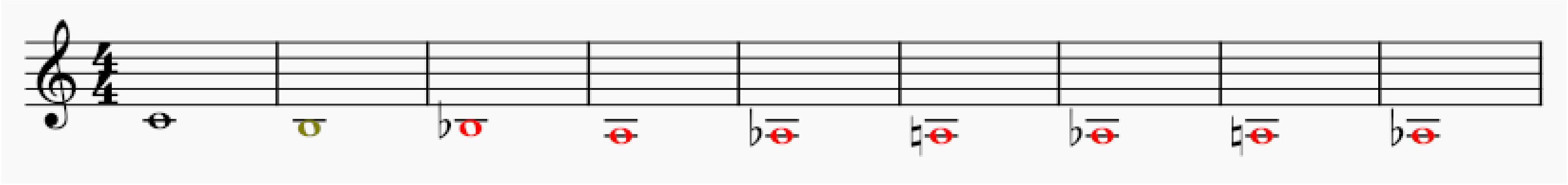
Iannis Xenakis(ヤニス・クセナキス)[1922~2001]という作曲家がいる.
彼は現代音楽というくくりに入る音楽を作曲しており,
特に数学を利用し先入観を排した音響の開拓に取り組んでいた.

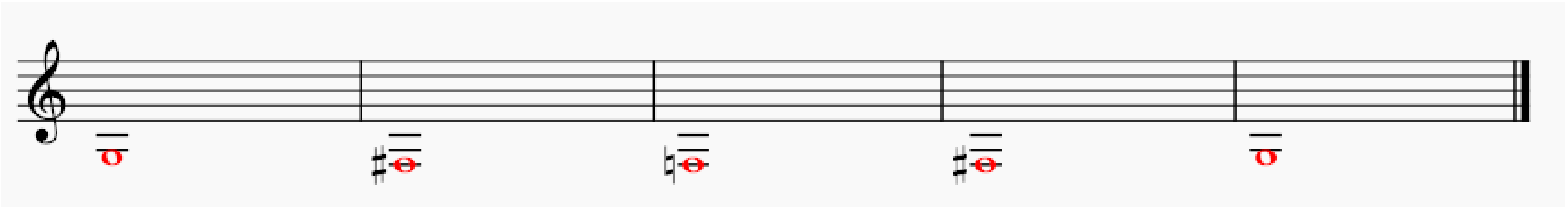
彼が利用した数学の中でマルコフ過程がある.

例えばドの音から始まり隣り合う半音に確率0.5で遷移するランダムウォークを考える.



独立なコイントスの結果, -1 -1 -1 -1 +1 -1 +1 -1 -1 -1 -1 +1 +1 となった場合,以下の音列が得られる.





要はランダムネスを利用して音を配列し,作曲をするということを
試みていたわけであるが, ランダムネスを利用して先入観を排する,
だけであると味気ないのでもう少し数学に踏み込み,マルコフ過程の
性質や応用について話す.

確率論について少しだけ

定義

Ωを集合とする.Ωの部分集合族$\varnothing \neq \mathcal{F} \subset 2^{\Omega}$が 以下の条件を満たすとき$\mathcal{F}$は完全加法族であるという

・$\varnothing \in \mathcal{F}$
・$A \in \mathcal{F}$ならば補集合A^Cも$\mathcal{F}$に属する
・$A_1, A_2, \ldots \in \mathcal{F}$ならば$\cup_{n=1}^{\infty} A_n \in \mathcal{F}$

定義

Ωを集合とし,$\mathcal{F}$がΩ上の完全加法族とする。写像$\mu : \mathcal{F} \to \mathbb{R}$が 以下の条件を満たすとき$\mu$は確率測度であるという.
・任意の$A \in \mathcal{F}$について$\mu(A) \geq 0$
・$\mu(\Omega) = 1$
・$A_1, A_2, \ldots \in \mathcal{F}$が互いに素なら$\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$

定義

集合Ω,完全加法族$\mathcal{F}$,確率測度μの組$(\Omega, \mathcal{F}, \mu)$を確率空間という.

<u>定義</u>

$(\Omega, \mathcal{F}, \mu)$を確率空間とし,$B \in \mathcal{F}$とする.
事象Bが起こる条件の下での条件付確率$\mu(\,\cdot\,|B)$を
$\mu(A|B) = \mu(A \cap B)/\mu(B)$
で定める.

<u>定義</u>

$(\Omega, \mathcal{F}, \mu)$を確率空間とする.$(\Omega', \mathcal{F}')$を別の集合と完全加法族の組とする.
関数$X : \Omega \to \Omega'$が確率変数であるとは,
任意の$B \in \mathcal{F}'$に対して$X^{-1}(B) \in \mathcal{F}$が成り立つことをいう.

例 (コイントス)

$\Omega = [0,1]$

$\mathcal{F}:\{[0,\frac{1}{2}),[\frac{1}{2},1]\}$を含む最小の完全加法族

μ :区間の長さを返す測度

とする.

$X:\Omega \to \{0,1\}$を

$$X(\omega)\begin{cases} 0 \ (0 \leq \omega < \dfrac{1}{2}) \\ 1 \ (\dfrac{1}{2} \leq \omega \leq 1) \end{cases}$$

で定めるとこれはコイントスをモデル化したものである.

Remark

確率変数を考えるとき,Ωが何者かはあまり気にしなくてよい.

例えば$\Omega = \{0,1\}, \mathcal{F} = 2^{\{0,1\}}, \mu(\{0\}) = \mu(\{1\}) = \frac{1}{2}$

$X(0) = 0, X(1) = 1$

とすればこれもコイントスをモデル化したものになっている.

定義

$X : (\Omega, \mathcal{F}, \mu) \to (\Omega', \mathcal{F}')$をな確率変数とするとき,

$B \in \mathcal{F}'$に対して$\mu(X^{-1}(B))$を$\mathbb{P}(B)$とかく.

Remark

$\mathcal{F}' = 2^{\Omega'}$である場合,$j \in \Omega'$に対して$\mu(X^{-1}(j))$を$\mathbb{P}(X = j)$とかく.

定義

$(\Omega, 2^{\Omega}, \mu)$を可算な確率空間とし,$X:\Omega \to \mathbb{R}$を確率変数とするとき,$X$の期待値を
$\mathbb{E}[X] = \sum_{\omega\in\Omega} X(\omega)\mu(\omega)$
とかく.

$A \in 2^{\Omega}$とする.
$\mathbb{E}[X, A] = \sum_{\omega\in A} X(\omega)\mu(\omega)$
で定める.

$A \in 2^{\Omega}$とする.事象Aのもとでの条件付期待値を
$\mathbb{E}[X|A] = \sum_{\omega\in\Omega} X(\omega)\mu(\omega|A)$
で定める.

マルコフ過程の定常分布とその性質

状態集合$\mathbb{S}$は有限か可算無限とする.(可算無限の場合は$\mathbb{S} = \mathbb{N}$と思ってもほぼよい)

定義

状態集合$\mathbb{S}$に値を取る(離散時間)マルコフ過程とは,

確率変数の列$X_0, X_1, \ldots$であって,

$\mathbb{P}(X_{n+1} = j | X_0 = i_0, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_n = i_n)$

を満たすものをいう.

Remark

これからは値域が高々可算な集合$\mathbb{S}$であるような確率変数の可算な族$X_0, X_1, \ldots$

を考えるので,Ωも可算と思ってよい.

実際,$\Omega = \mathbb{S} \times \mathbb{S} \times \cdots$,$\mathcal{F} = 2^{\Omega}$,

$\mu((i_0, i_1 \ldots)) = \mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots)$

$X_n((i_0, i_1 \ldots)) = i_n$

と考えればよい.

以後,$\mathbb{P}(X_{n+1}=j|X_n=i)$が時刻$n$に寄らない場合のみを考える.

$\mathbb{P}(X_{n+1}=j|X_n=i)$を$p_{ij}$とかくと,行列の計算のようにして確率の計算を行うことができる.

<u>定義</u>

・$p_{ij} \geq 0$
・各$i \in \mathbb{S}$について$\sum_{j\in\mathbb{S}} p_{ij} = 1$
を満たすような$\{p_{ij}\}_{i,j\in\mathbb{S}}$を確率遷移行列と呼び$P$とかく.

<u>定義</u>

$\mu_i \geq 0. \sum_{i\in\mathbb{S}} \mu_i = 1$を満たす横ベクトル$(\mu_1, \mu_2, \ldots)$
を確率ベクトルという.

初期分布が確率ベクトルμに従う,すなわち$\mathbb{P}(X_0 = i) = \mu_i$であり,$\mathbb{P}(X_{n+1} = j | X_n = i) = p_{ij}$であるようなマルコフ過程を考えてみよう.

例えば$\mathbb{P}(X_1 = j) = \sum_{i \in \mathbb{S}} \mathbb{P}(X_0 = i)\mathbb{P}(X_1 = j | X_0 = i) = \sum_{i \in \mathbb{S}} \mu_i p_{ij}$なので

確率ベクトル$(\mathbb{P}(X_1 = 1), \mathbb{P}(X_1 = 2), \ldots)$は$\mu P$と書ける.

同様に確率ベクトル$(\mathbb{P}(X_n = 1), \mathbb{P}(X_n = 2), \ldots)$は$\mu P^n$と書ける.

$$(\mu_1, \mu_2, \ldots) \begin{pmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots \end{pmatrix}$$

行列計算のアナロジーがきく！！

$\rho \in \mathbb{R}^{\mathbb{S}}$にたいして$||\rho||_1 = \sum_{j\in\mathbb{S}} |\rho_j|$で定める.

行列やベクトルの成分の書き方として$(P)_{ij}$みたいな書き方をすることがある

定常分布の存在条件のうち重要なものを述べる.

<u>定理</u>

確率遷移行列Pが$\exists j_0 \in \mathbb{S}, \exists\varepsilon > 0, \forall i \in \mathbb{S}(P)_{ij_0} \geq \varepsilon$を満たすならば

$\pi P = \pi$を満たす確率ベクトル(定常分布)がただ一つ存在して

任意の確率ベクトルμに対して

$||\mu P^n - \pi||_1 \leq (1-\varepsilon)^n ||\mu - \pi||_1 \leq 2(1-\varepsilon)^n$

を満たす.

(証明)

$\rho \in \mathbb{R}^{\mathbb{S}}, ||\rho||_1 < \infty$なら

$\sum_{j\in\mathbb{S}}(\rho P)_j = \sum_{i\in\mathbb{S}}(\rho)_i$

である.

実際,

$$(\rho_1, \rho_2, \ldots) \begin{pmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots \end{pmatrix}$$
$$= (\rho_1 p_{11} + \rho_2 p_{21} + \cdots)$$
$$+(\rho_1 p_{12} + \rho_2 p_{22} + \cdots) + \cdots$$
を考えればよい.

さらに,$\sum_{i \in \mathbb{S}} \rho_i = 0 \Rightarrow ||\rho P^n||_1 \geq (1-\varepsilon)^n ||\rho||_1$であることが帰納法でわかる.

$$|(\rho P)_j| = |\sum_i (\rho)_i (P)_{ij}|$$
$$= |\sum_i (\rho)_i ((P)_{ij} - \varepsilon \delta_{jj_0})|$$
$$\leq \sum_i |(\rho)_i| ((P)_{ij} - \varepsilon \delta_{jj_0})$$
だから

$$||\rho P||_1 \leq \sum_{j \in \mathbb{S}} (\sum_{i \in \mathbb{S}} |\rho_i| (P_{ij} - \varepsilon \delta_{jj_0}))$$
$$= \sum_{i \in \mathbb{S}} |\rho_i| (\sum_{j \in \mathbb{S}} (P_{ij} - \varepsilon \delta_{jj_0}))$$
$$= (1-\varepsilon) ||\rho||_1$$

μ:確率ベクトルとし,$\mu_n = \mu P^n$とする.
$1 \le m < n$ について,$\mu_n = \mu_{n-m} P^m$であり
$\sum_{i \in \mathbb{S}} ((\mu_{n-m})_i - \mu_i) = 1 - 1 = 0$だから

$$||\mu_n - \mu_m||_1 \le (1-\varepsilon)^m ||\mu_{n-m} - \mu||_1 \le 2(1-\varepsilon)^m$$

$\{\mu_n\}_1^\infty$はCauchy列.∃!π$||\mu_n - \pi||_1 \to 0$

$||\mu_n||_1 = 1$(定数列)だからπも確率ベクトルである.

$$\pi = \lim_{n\to\infty} \mu P^{n+1} = \lim_{n\to\infty} (\mu P^n) P = \pi P$$

しかも,$\pi_{j_0} = \sum_{i \in \mathbb{S}} (\pi)_i (P)_{i j_0} \ge \varepsilon \sum_{i \in \mathbb{S}} \pi_i = \varepsilon$

ν:確率ベクトルとするとき,

$$\begin{aligned}
&||\nu P^m - \pi||_1 \\
&= ||(\nu - \pi) P^m||_1 \\
&\le (1-\varepsilon)^m ||\nu - \pi||_1 \\
&\le 2(1-\varepsilon)^m
\end{aligned}$$

なので初期分布にかかわらず定常分布πに収束　□

状態集合が有限であれば定常分布がちょうど一つになる確率繊維行列
は簡単に作れる.

系

$\mathbb{S}$が有限であるとき,確率遷移行列Pが$(P)_{ij} > 0\ (i, j \in \mathbb{S})$をみたすならば,
定常分布がちょうど一つ存在する.

定常分布がちょうど一つになる条件はもっと緩めることができる.

命題

$\exists M, \exists \varepsilon > 0, \exists j_0 \in \mathbb{S}, \forall i \in \mathbb{S} (P^M)_{ij_0} \geq \varepsilon$をみたすならば
ただひとつの定常分布πが存在して,任意の初期分布μに対して
$||\mu P^n - \pi||_1 \leq 2(1-\varepsilon)^{\lfloor n/M \rfloor}$

定義

$A_n = \frac{1}{n}\sum_{m=0}^{n-1} P^m$とおく.
するとこれもまた確率遷移行列になっている.

定義

$$T_j = 1_j(X_0) + 1_j(X_1) + \cdots$$

$$T_j^{(n)} = 1_j(X_0) + 1_j(X_1) + \cdots + 1_j(X_{n-1})$$

$$\overline{T_j^{(n)}} = \frac{1}{n}(1_j(X_0) + 1_j(X_1) + \cdots + 1_j(X_{n-1}))$$

$(A_n)_{ij} = \frac{1}{n}\sum_{m=0}^{n-1} \mathbb{P}(X_m = j | X_0 = i)$
$= \mathbb{E}[\frac{1}{n}\sum_{m=0}^{n-1} 1_j(X_m) | X_0 = i]$
である.

命題

$\exists M, \exists \varepsilon > 0, \exists j_0 \in \mathbb{S}, \forall i \in \mathbb{S} (A_M)_{ij_0} \geq \varepsilon$をみたすならば
ただひとつの(Pについての)定常分布πが存在して,任意の初期分布μに対して
$||\mu A_n - \pi||_1 \leq \frac{M-1}{n\varepsilon}$

Remark

① $\exists j_0 \in \mathbb{S}, \exists \varepsilon > 0, \forall i \in \mathbb{S} (P)_{ij_0} \geq \varepsilon$

② $\exists M, \exists \varepsilon > 0, \exists j_0 \in \mathbb{S}, \forall i \in \mathbb{S} (P^M)_{ij_0} \geq \varepsilon$

③ $\exists M, \exists \varepsilon > 0, \exists j_0 \in \mathbb{S}, \forall i \in \mathbb{S} (A_M)_{ij_0} \geq \varepsilon$

はこの順でゆるくなっている.

そもそも状態集合が有限の場合は定常分布は存在はする.

じっさい,$\mathbb{R}^N$の確率ベクトルの集合は$\mathbb{R}^N$のコンパクト集合であるから,
点列$\{\mu A_M\}_{M\geq 1}$は収束部分列$\{\mu A_{M_k}\}_{k\geq 1}$をもつ.
この収束先をπとおくと

$$\pi P = \lim_{k\to\infty} \frac{1}{M_k} \sum_{m=0}^{M_k-1} \mu P^{m+1} = \lim_{k\to\infty} \left(\frac{\mu P^{M_k} - \mu}{M_k} + \mu A_{M_k} \right) = \pi$$

有限状態の場合で定常分布がuniqueになる必要十分条件としては以下が知られている.

命題

有限状態の場合,0が$I - P$の固有値であってなおかつ固有多項式の重根でないこと
が、定常分布が一意である事の必要十分条件である.

(有限状態であれば)

与えられた分布$\boldsymbol{\pi}$を定常分布に持つような遷移行列Pを求めるには線形計画

Minimize $\quad p_{11}$

Subject to
$$\begin{cases} \pi P = \pi \\ p_{ij} \geq 0 \ (i, j \in \mathbb{S}) \\ \sum_{j=1}^{N} p_{ij} = 1 \ (i \in \mathbb{S}) \end{cases}$$

を解けばよい.

定常分布の特徴づけにかかわる定理を紹介する.

<u>命題</u>

$\rho_j = \inf\{n \geq 1 : X_n = j\}$とする.

$\exists M, \exists \varepsilon > 0, \exists j_0 \in \mathbb{S}, \forall i \in \mathbb{S} (A_M)_{ij_0} \geq \varepsilon$をみたすならば

定常分布πについて

$(\pi)_j = \frac{1}{\mathbb{E}[\rho_j | X_0 = j]}$

(直感的アイデア)

$\mathbb{E}[T_j^{(n)} | X_0 = j] = (A_n)_{jj} \to \pi_j$なので

$X_0 = j \Rightarrow \overline{T}_j^{(\rho_j^{(m)})} = \frac{1}{\rho_j^{(m)}} \sum_{l=0}^{\rho_j^{(m)}-1} 1_j(X_l) = \frac{m}{\rho_j^{(m)}}$

$\pi_j = \lim_{m\to\infty} \mathbb{E}[\overline{T}_j^{(\rho_j^{(m)})} | X_0 = j] = \frac{1}{\mathbb{E}[\rho_j | X_0 = j]}$

ただし

$\rho_j^{(0)} = 0$

$\rho_j^{(1)} = \inf\{n \geq 1 : X_n = j\}$

$\rho_j^{m} = \inf\{n > \rho_j^{(m-1)} : X_n = j\}$

Q.


講演の準備期間は短いってわかってたのに・・・
どうしてもっと前もってちゃんと準備しようと思わなかったんだろう・・・

A.


俺は泣いた
スライドは書かないと出来上がらないんだよ
寝ている間に進捗を生やしてくれる妖精なんて御伽話なんだよ

参考文献

【Strook 2014】Daniel W. Stroock “An Introduction to Markov Processes” 2nd edition Springer 2014
【Xenakis 2001】 Iannis Xenakis “Formalized Music: Thought and Mathematics in Composition” Pendragon Pr 2001
【山田 2020】山田 鐘人 (原著), アベ ツカサ (イラスト), “葬送のフリーレン”,小学館サービス 2020
【吾峠2014】吾峠呼世晴 “鬼滅の刃 ”, 集英社 2014