

ジャズ・アドリブにおける音高と音長の列に対する時系列分析

駒木 泰
札幌大学経済学部

ジャズ・アドリブの単音の音高と音長の並びを時系列データとみなして時系列分析を適用し、演奏の特徴を見出した。Charlie Parker の Confirmation のテーマ、Charlie Parker と Clifford Brown のアドリブを対象に、自己相関、ガウス型と整数型の AR モデルの推定、VAR モデルの推定による Granger 因果性の検定を行った。ラグの長さ、パラメータの推定値について、いくつかの演奏の特徴が見出された。

The time series analysis for the sequences of pitch and duration of ad-lib playing in Jazz

Toru Komaki
Economics Department, Sapporo University

The time series analysis is applied to the sequences of pitch and duration of ad-lib playing in Jazz. Autocorrelation, Gaussian AR, Integer AR, and VAR models are estimated. Granger causality is tested by estimated VAR model. The applied data are the theme of Confirmation, and the ad-lib playing in the piece by Charlie Parker and Clifford Brown. The character of playing is found in the length of lag and estimated parameters of the models.

1. はじめに

本論では、音楽データの解析のため、単音の音高と音長の並びを時系列データとみなして時系列分析を適用し、演奏の特徴を見出す。

経済や物理現象で観測される時系列データは、一般的には等間隔に存在する。一方、音楽データは時間の順序で並んではいないものの、同時点に複数の音が観測される和音があり、観測時点としての発音時間が等間隔ではなく、さらにデータの無い区間としての休符が存在する。しかし、ジャズのアドリブは、1つの管楽器の演奏に限れば単音のデータが得られ、8分音符が羅列する部分では等間隔なデータに近くなる。また、音の数が多いため、統計的な分析に適している。

演奏の特徴を把握するため、音高と音長を単純に並べたデータに対して、その列に潜む規則

性を見出す。自己相関によりランダム性の程度を、また AR(p) (Autoregressive of pth order) モデルにより、過去からの影響を明示的に捉える。音高と音長とで音符のベクトルを構成するものと考え、VAR(p) (Vector Autoregressive of pth order) モデルを推定する。それにより音高と音長の関係を Granger 因果性の検定 (Hamilton(1994)、Lutkepohi and Kratzig(2004)) により明らかにする。

Charlie Parker (以下、Parker) の Confirmation はテーマでありながら、アドリブ的なフレーズが多いため、これをベンチマークのデータとする。アドリブは、同じ Parker と、同曲の Clifford Brown (以下、Brown) の演奏をとりあげる。Parker と Brown とでは、アルトサクソとトランペットという楽器の違いがある。しかし、ここでは演奏の特徴としてみなすことにより、楽

器の違いとしては識別しない。

2. 時系列モデルとデータ

楽器には音域があるので、音高のとり値には上限と下限がある。また音楽として成り立つには、音長も同様に限界がある。そこで両者とも定常データとみなすことができる。

ガウス型の AR(p)モデルは、音高を x_t ($t=1, \dots, T$) とすると、

$$x_t = \mu + \sum_{i=1}^p \beta_i x_{t-i} + \varepsilon_t$$

となる。ただし、 μ 、 β_i ($0 \leq \beta_i < 1$) はパラメータ、 ε_t は平均ゼロ分散 σ^2 のガウス分布に従う誤差項である。 ε_t にラグを課した MA (Moving Average) モデルも定式化出来るが、次の整数型モデルと対応させるため、AR(p)モデルを仮定した。

整数型データの INAR (Integer Valued Autoregressive) モデルは、

$$x_t = \alpha \circ x_{t-1} + \eta_t$$

となる (Al-Osh and Alzaid(1987))。ここで、 x_t は正の整数、 α ($0 \leq \alpha < 1$) はパラメータ、 η_t は平均 (分散) のパラメータが λ のランダムなポアソン分布に従う確率変数である。 \circ は binomial thinning operator といわれ、

$$\alpha \circ x_{t-1} = \sum_{i=1}^{x_{t-1}} v_i, \quad \begin{cases} \Pr(v_i = 1) = \alpha \\ \Pr(v_i = 0) = 1 - \alpha \end{cases}$$

となる。 η_t は常に正の値をとるので、誤差項というよりも、 $\alpha \circ x_{t-1}$ での減少分を補う役割をする。右辺に x_t のラグ変数を付け加えると、AP(p) のような一般的なモデルにはなるが、最尤法で推定する場合に収束しにくくなり、推定は困難となる。

他の整数型データの AR モデルとして、DAR モデル (Jacobs and Lewis (1983))、

$$x_t = u_t x_{t-1} + (1 - u_t) \omega_t, \quad \begin{cases} \Pr(u_t = 1) = \rho \\ \Pr(u_t = 0) = 1 - \rho \end{cases}$$

を用いる。 ρ ($0 \leq \rho \leq 1$) はパラメータ、 ω_t はパラメータ θ のランダムなポアソン分布に従う確率変数である。 ρ が大きいほど、 x_t に同じ値が連続して表れる。しかし、AP(p) のように、右辺に x_t のラグ変数を並べることはできない。

VAR(p)モデルは、AP(p)を連立させたものである。例えば、VAR(1)は t 番目の音長を y_t とすると、以下のようになる。

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

Granger 因果性は、以下を F 検定することで確認できる。

$\beta_{12} \neq 0$: 過去の音長が現在の音高へ影響

$\beta_{21} \neq 0$: 過去の音高が現在の音長へ影響

また、 ε_{1t} と ε_{2t} の相関行列により、音高と音長の同時点での関係も χ^2 検定ができる。

INAR、DAR はともにカウントデータ (比例尺度) を対象にしている。音高・音長は間隔尺度であるが、ここでは整数を扱うモデルとして援用した。

音程は MIDI コードの値とし、音長は全音符を 48 とし他の音長をすべて整数に変換した。休符は取り除いて、音と音を直接並べた。また、1 コーラスを 1 つのデータセットとした。テーマは 1 コーラスであり、Parker は 2 コーラスあるのでそれぞれ Parker1~Parker2、Brown は 4 コーラスで Brown1~Brown4 とした。計 5 データセットを作り、それぞれに分析を適用した。

なお、本論で用いた一連の分析方法は Hamilton(1994)、Lutkepohi and Kratzig(2004)、Cameron and Trivedi(1998) に詳しい。*計算は JMulTi (Lutkepohi and Kratzig(2004)) と Gauss で行った。

図1 音高の自己相関

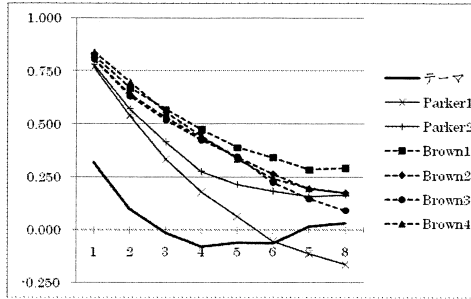


図2 音長の自己相関

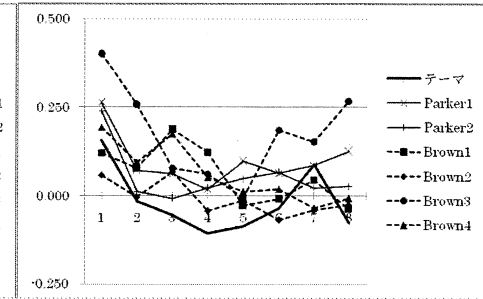


表1 音高の推定結果

コーラス	テーマ	Parker		Brown			
		Parker 1	Parker 2	Brown 1	Brown 2	Brown 3	Brown 4
T	179	247	233	206	217	202	200
				AR(p)			
p	1	1	2	1	3	1	1
β_1	0.319*	0.767*	0.861*	0.827*	0.807*	0.829*	0.873*
	(0.071)	(0.041)	(0.066)	(0.039)	(0.069)	(0.041)	(0.037)
β_2			-0.104		-0.063		
			(0.066)		(0.089)		
β_3					0.070		
					(0.070)		
μ	67.6*	65.0*	65.0*	72.4*	72.5*	77.9*	72.6*
	(0.334)	(0.853)	(0.870)	(1.22)	(1.24)	(1.31)	(1.54)
σ^2	9.61	10.1	10.7	9.55	12.0	10.6	8.09
Q_8	2.16	13.3*	10.4	8.73	2.33	6.14	6.67
				INAR			
α	0.887*	0.912*	0.907*	0.928*	0.912*	0.930*	0.942*
	(0.013)	(0.009)	(0.009)	(0.008)	(0.009)	(0.008)	(0.006)
λ	7.65*	5.71*	6.079*	5.15*	6.39*	5.44*	4.20*
	(0.881)	(0.560)	(0.615)	(0.554)	(0.676)	(0.605)	(0.456)

注) ()内はヘシアン行列から求めた標準誤差。*は5%で有意。 Q_8 はラグ8のLjung-Box統計量。

表2 音長の推定結果

コーラス	テーマ	Parker		Brown			
		Parker 1	Parker 2	Brown 1	Brown 2	Brown 3	Brown 4
T	179	247	233	206	217	202	200
				AR(p)			
p	1	1	2	1	3	1	3
β_1	0.226*	0.262*	0.261*	0.121	0.086	0.402*	0.180*
	(0.089)	(0.062)	(0.066)	(0.069)	(0.086)	(0.065)	(0.072)
β_2			-0.041		-0.040		0.114
			(0.066)		(0.087)		(0.085)
β_3					0.110		0.225*
					(0.087)		(0.085)
μ	6.92*	5.32*	5.52*	6.07*	5.98*	6.40*	6.84*
	(0.372)	(0.281)	(0.282)	(0.196)	(0.168)	(0.432)	(0.588)
σ^2	14.8	10.6	11.3	6.18	4.37	13.6	15.9
Q_8	6.67	5.60	1.15	10.1	2.46	33.1*	0.617
				INAR			
α	0.142*	0.246*	0.242*	0.184*	0.660*	0.321*	0.135*
	(0.052)	(0.049)	(0.050)	(0.089)	(0.072)	(0.052)	(0.045)
λ	5.92*	4.01*	4.12*	4.95*	2.09*	4.31*	5.78*
	(0.390)	(0.285)	(0.298)	(0.555)	(0.426)	(0.353)	(0.338)
				DAR			
ρ	0.659*	0.749*	0.725*	0.758*	0.811*	0.768*	0.767*
	(0.038)	(0.029)	(0.030)	(0.032)	(0.029)	(0.032)	(0.032)
θ	9.07*	6.97*	6.67*	7.13*	7.03*	8.32*	9.32*
	(0.396)	(0.344)	(0.327)	(0.375)	(0.413)	(0.436)	(0.464)

注) ()内はヘシアン行列から求めた標準誤差。*は5%で有意。 Q_8 はラグ8のLjung-Box統計量。

表3 グレンジャーの因果性

コーラス	テーマ	Parker		Brown			
		1	2	1	2	3	4
p	2	1	2	1	1	1	1
過去音高⇒現在音長	7.493*	0.332	4.32*	8.71*	0.115	0.023	7.03*
過去音長⇒現在音高	0.591	0.173	3.47*	2.51	24.0*	28.8*	4.03*
現在音高⇄現在音長	0.003	0.177	0.465	2.17	7.51*	2.14	0.693

注) F検定(過去と現在)と χ^2 検定(現在同士)の結果。*は5%で有意。

3. 分析結果

(1) 自己相関

図1に音高の8期差までの自己相関のパターンを示した。テーマは概ね0.3以下の値をとり、ランダム性が高いことを示す。BrownはParkerよりも値が大きく、4コーラス通じてランダム性のパターンは安定的である。

音長については図2に示した。テーマは±0.2以内であり小さな値をとり、Brown3を除いた他のアドリブはどれも同様なパターンを示した。

(2) AR(p)、INAR、DARモデル

AR(p)は、ラグの長さ p を AIC、BIC、HQ の各情報量基準で決定した後、最尤法で推定した。INARとDARも最尤法で推定した。INARは左辺のラグが長くなると収束が難しく、せいぜい1期のラグでしか収束しなかった。また、音高についてのDARは、すべてのデータセットで収束しなかった。

音高についての結果を表1に掲げた。情報量基準によりラグが $p > 1$ と選ばれても、 $p > 1$ のラグ変数は有意ではない。現在の音高が1個前の音高で決まることを示す。これは、マルコフ遷移確率を用いた作曲において、マルコフ性を1階と仮定できることの裏付けとなる。

誤差項に系列相関を含むか否かの検定を8期のLung-Box統計量によって行った。Parker1のみが有意となり、MAモデルの追加を示唆している。

AR(p)の β_i の推定値をみると、テーマのみが小さく、自己相関が小さいことと一致している。ParkerとBrownとではあまり相違はない。 σ^2 の推定値はParkerではどのコーラスでも同様であるが、Brownではコーラス毎に異なる。Brownは音高の流れからの逸脱の仕方がコーラスで異なる。

整数型のINARでは、テーマはアドリブに比べると、 α の推定値が小さい代わりに、 λ の推定値が大きい。ParkerとBrownともに同様の結果を示している。

表2に音長の結果を掲げた。テーマとParker

についてのAR(p)のラグの長さは、音高の結果と同様に、 $p > 1$ のラグ変数は有意ではない。Brownについては、コーラス毎にラグ変数の有意性が異なり、必ずしも1個前の音長から決まるとは限らない場合がある。

また、Brown1、Brown2、Brown4を除いて、 p の値とラグ変数の有意性が音高の結果と一致している。音高と音長とでラグ構造が似通っていることを示す。Lung-Box統計量はBrown3のみが有意となった。

AR(p)の β_i の推定値は、テーマとParkerとが近い。Parkerがどのコーラスでも安定している反面、Brownではコーラス毎に異なる。これは σ^2 、およびINARの α と λ の推定結果でもみられる。リズムの選択に関しては、Brownのアドリブの方がコーラス毎で多様である。

DARでは、テーマ、Parker、Brownの順に ρ の推定値が大きくなっており同じ音長が続く確率がその順に高くなることを示す。一方、 θ の推定値はテーマが大きく、変化に富んだ音長を生み出すことを示す。Brownは、Parkerに比べると θ の推定値が大きく、Parkerよりもランダムな音長を生み出している。すなわち、Brownの方が、音長の選択は多様である一方で、同じ音長を続けて演奏する確率が高いといえる。

(3) VAR(p)モデル

音高と音長を1つのベクトルとするVAR(p)では、ラグ p の長さをAIC、BIC、HQの各情報量基準で決定した後、OLSで推定した。表3に p の値とグレンジャーの因果性の検定結果を挙げた。ここでも、ラグの長さが1となる場合が多く、1階のマルコフ遷移確率の妥当性を示す。

グレンジャー因果性では、テーマは過去の音高が現在の音長に影響していることがわかる。Parkerは2コーラス目において音高と音長がお互いに影響しあうフィードバックの関係がみられる。Brownは方向性がコーラス毎に異なり、最後のコーラスでフィードバックの関係がみられる。誤差項に関する同時の因果性は、有意になったのはBrownの2コーラス目のみであ

る。Brownの方が、音高と音長を関連させながら演奏していることがわかる。

4. おわりに

単音のジャズのテーマとアドリブを対象に、音高と音長の列を時系列データとみなして、時系列分析を適用した。以下に要約する。

1) Confirmationのテーマがアドリブ的と言われる点：テーマの音高はアドリブに比べてランダムである。作曲者のParkerのアドリブについても共通点は見いだせなかった。

2) アドリブ演奏について：Parkerはスタジオ録音であるためか、2コーラスのアドリブには相違はあまりなかった。一方、Brownはライブ演奏で4コーラスもある。Brownの方が多様に音高と音長を選択し、しかも同じ音長を続けて演奏しがちである。また、Brownは音高と音長の関連性がコーラス毎に異なっていた。

3) 作曲アルゴリズムで用いられるマルコフ遷移確率の次数を1階とする点：AR(p)で $p=1$ のラグ変数が有意であり、VAR(p)のラグが概ね $p=1$ であった。また、音高と音長とでラグ構造が似る結果も得られた。

以上の結果を念頭に置きながら、実際の演奏を再度聴き直してみた。分析結果の解釈を直接聞き取るとは難しいが、新たな視点を与えられた上での発見がいくつか生まれる。音楽データの解析の目的は、新たな視点から音楽を聴く機会を与えることにもあると思う。

なお、パラメータの推定値を用いてのシミュレーションにより、アドリブを再生することも可能である。しかし、そこには統計学的な評価よりも、人間の耳による評価が必要となる。それよりも、フレーズパターンを時系列データとして分析できる手法を開発した方が効率的である。それらは今後の課題とする。

参考文献

[1] Al-Osh, M.A., and A.A. Alzaid (1987): "First-order integer-valued autoregressive

(INAR(1)) process," *Journal of Time Series Analysis*, 8, 261-275.

[2] Cameron, A.C., and P.K. Trivedi (1998): *Regression Analysis of Count Data*. Cambridge University Press.

[3] Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press.

[4] Jacobs, P.A., and P.A.W. Lewis (1983): "Stationary Discrete Autoregressive-Moving Average Time Series generated by Mixtures," *Journal of Time Series Analysis*, 4, 19-36.

[5] Lutkepohi, H., and M. Kratzig (2004): *Applied Time Series Econometrics*. Cambridge University Press.