Iダイバージェンスに基づく統計的音響信号処理の枠組に向けて

亀岡弘和 （NTT コミュニケーション科学基礎研究所）

NTT

1. 序論

動機

- 定常ガウス過程を仮定した信号モデルの最尤推定
  ＝ 板倉斎藤距離規準のスペクトル密度推定 [1]（詳しくは2章で）

→パラメトリックなスペクトルモデルを仮定する場合
板倉斎藤距離を最小化する最適化アルゴリズムが必要

しかし，板倉斎藤距離規準での数値最適化は厄介！(なことが多い)
非凸，勾配計算が数値的に不安定，小さい雑音の影響が大，等

- 最適化に関し好ましい性質の多い Iダイバージェンス [2] に着目
  - 凸性，相性の良い強力なパラメトリックモデルの存在，最適化におけるアルゴリズムの振る舞いの良さに関する実証報告あり
  - 近年，パワースペクトルモデルのフィッティング規準として注目され始めているがパワー領域で閉じているため用途は限定的

背後の信号の確率過程が明らかになれば Iダイバージェンスの好ましい性質を活かした新たな信号処理の枠組が実現可能に!?

本研究の目的

- Iダイバージェンス規準のスペクトル密度推定に帰着する信号の確率過程（以下の表中の「？」）を明らかにしたい

時間領域	周波数領域	スペクトル距離
定常ガウス過程→	複素正規分布→	板倉斎藤距離
？ ←	？ ←	Iダイバージェンス

2. 板倉斎藤距離 [1]

定常ガウス過程

- 離散時間信号のベクトル $\boldsymbol{x}=(x_1,\dots,x_I)^{\mathrm{T}}\in\mathbb{R}^I$ がガウス過程に従う→ $\boldsymbol{x}\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{\Sigma})$
- $\boldsymbol{x}$ のフーリエ変換 $\boldsymbol{z}=\boldsymbol{F}\boldsymbol{x}$ （$\boldsymbol{F}$ は離散フーリエ変換行列）もまた正規分布する→ $\boldsymbol{z}\sim\mathcal{N}_{\mathbb{C}}(\boldsymbol{0},\boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{F}^{\mathrm{H}})$（$\mathcal{N}_{\mathbb{C}}$ は複素正規分布）
- $\boldsymbol{x}$ が定常過程でさらに巡回性を仮定すると $\boldsymbol{\Sigma}$ は非負定値巡回行列

→ $\boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{F}^{\mathrm{H}}=\mathrm{diag}(\lambda_1,\dots,\lambda_I)$ $\lambda_i\geq 0:\boldsymbol{\Sigma}$ の固有値

→ $z_i\sim\mathcal{N}_{\mathbb{C}}(0,\lambda_i)$ $(i=1,\dots,I)$ つまり λ_i は周波数ビンi のパワースペクトル密度 $\mathbb{E}[|z_i|^2]$ に対応

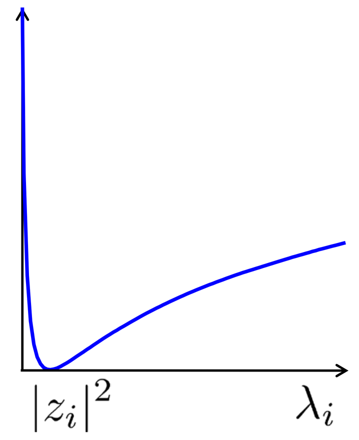
最尤スペクトル推定

- $z_i\sim\mathcal{N}_{\mathbb{C}}(0,\lambda_i)$ のとき最尤の λ_i は？

対数尤度関数 $L_{\mathcal{N}}(\lambda_i)=-\log\pi\lambda-|z_i|^2/\lambda_i$

$$\operatorname*{argmax}_{\lambda_i} L_{\mathcal{N}}(\lambda_i)=|z_i|^2\longrightarrow L_{\mathcal{N}}(|z_i|^2)\geq L_{\mathcal{N}}(\lambda_i)$$

$$\therefore L_{\mathcal{N}}(|z_i|^2)-L_{\mathcal{N}}(\lambda_i)=\underbrace{\frac{|z_i|^2}{\lambda_i}-\log\frac{|z_i|^2}{\lambda_i}-1}_{|z_i|^2\text{ と }\lambda_i\text{ の板倉斎藤距離}}\geq 0$$



$\lambda_1,\dots,\lambda_I$ の最尤推定 ＝ 標本スペクトル $|z_1|^2,\dots,|z_I|^2$ と
パワースペクトル密度 $\lambda_1,\dots,\lambda_I$ の板倉斎藤距離最小化
（パラメトリックなパワースペクトル密度モデルを仮定した場合は板倉斎藤距離規準の最適化問題になる）

3. 確率密度関数の提案

定義 (Complex Poisson Distribution)

$\mathcal{D}=\{z\in\mathbb{C}\mid |z|^p\in\mathbb{N}\}$ をサポートとする複素数 z の確率密度関数

$$f_z(z;\lambda,p)=\frac{pe^{-\lambda}}{2\pi}\frac{|z|^{p-2}\lambda^{|z|^p}}{(|z|^p)!}$$

を複素ポアソン分布（仮称）と呼ぶ。ただし, $\lambda>0$，$p>0$ である。z が複素ポアソン分布に従うことを

$$z\sim\mathrm{cPois}(\lambda,p)$$

と表記する。

Property 1.

f_z の $\mathcal{D}$ 上の積分は1。 $\int_{\mathcal{D}}f_z(z;\lambda,p)\mathrm{d}z=1$

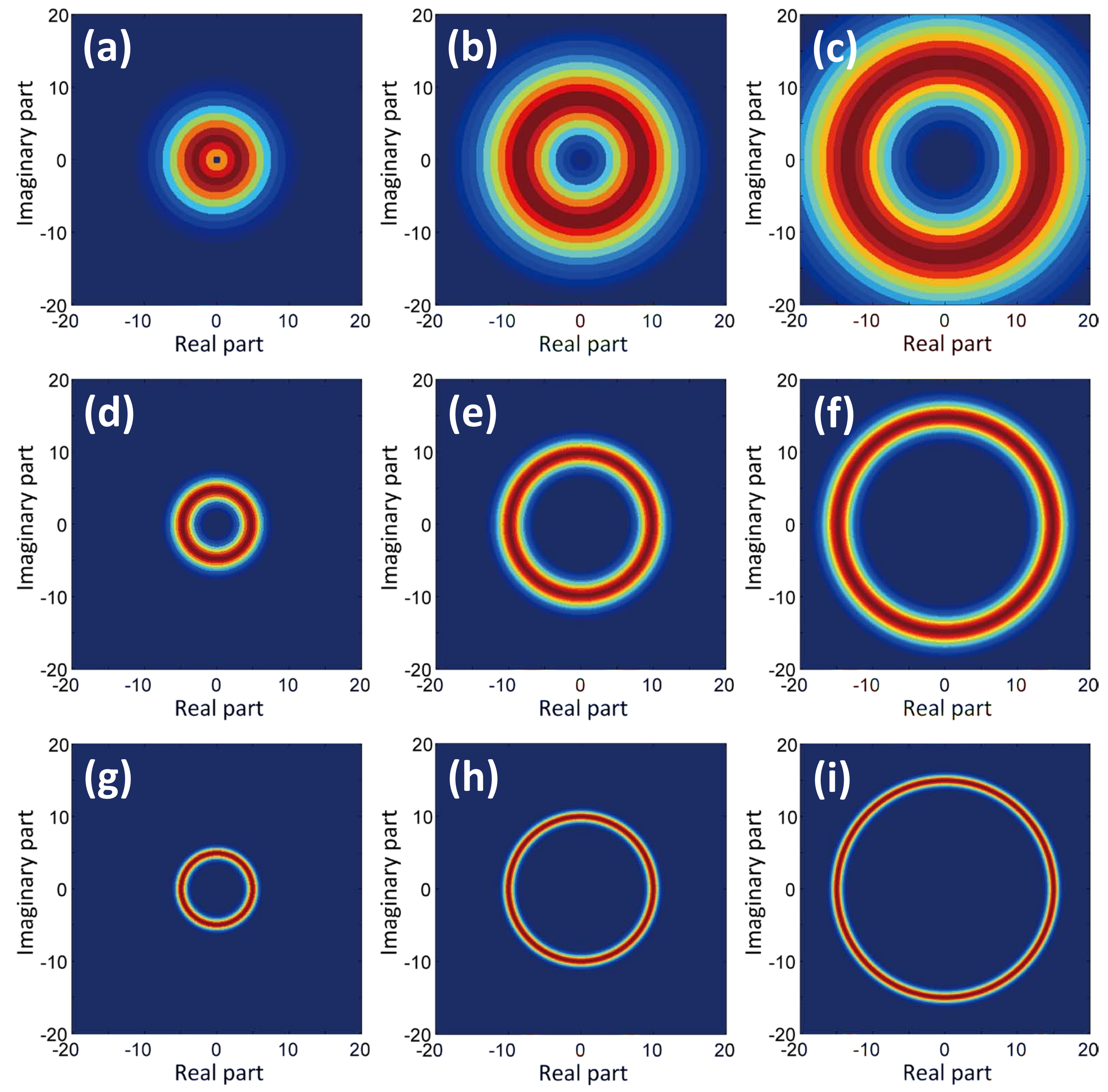
Property 2.

$z\sim\mathrm{cPois}(\lambda,p)$ のとき，z はCircular Complex Random Variable (CCRV) [3]である（右図参照）。

Property 3.

$z\sim\mathrm{cPois}(\lambda,p)$ のとき，$k=|z|^p$ は平均 λ のポアソン分布に従う。

$$k\sim\mathrm{Pois}(\lambda)=\frac{\lambda^k e^{-\lambda}}{k!}$$



複素ポアソン分布のカラーマップ表示図

$(\lambda,p)=$ **(a)** $(5,1)$ **(b)** $(10,1)$ **(c)** $(15,1)$ **(d)** $(5^{1.5},1.5)$ **(e)** $(10^{1.5},1.5)$ **(f)** $(15^{1.5},1.5)$ **(g)** $(5^2,2)$ **(h)** $(10^2,2)$ **(i)** $(15^2,2)$

4. 複素ポアソン分布の性質

Property 4.

$z\sim\mathrm{cPois}(\lambda,p)$ のとき, q 次モーメント $\beta_{n;m}:=\mathbb{E}[z^n z^{*m}]$ [3] は

$$\beta_{n;m}=\begin{cases}\mu'_{q/p} & (n=m)\\ 0 & (n\neq m)\end{cases}$$

で与えられる。ただし, n,m は自然数, $q=n+m$ である。また, μ'_ξ は平均 λ のポアソン分布の ξ 次のfractionalモーメントである。

Example 1.

Property 4 より，$z\sim\mathrm{cPois}(\lambda,p)$ の平均と分散は以下となる。

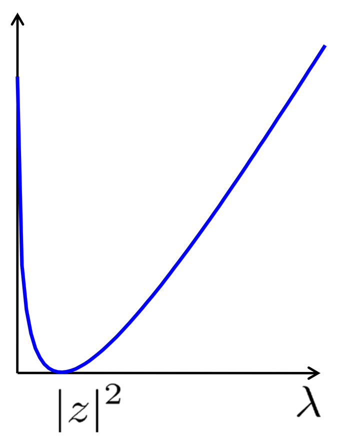
$$\mathbb{E}[z]=0$$

$$\mathbb{V}[z]=\mathbb{E}[|z|^2]=\begin{cases}\lambda^2+\lambda & (p=1)\\ \lambda & (p=2)\end{cases}$$

$$\because\quad \mu'_1=\lambda,\ \mu'_2=\lambda^2+\lambda$$

Property 5.

$z\sim\mathrm{cPois}(\lambda,p)$ のとき, λ の最尤推定は $|z|^p$ と λ の Iダイバージェンス最小化と等価である。



$\because$ 対数尤度関数 $L_{\mathcal{P}}(\lambda)\overset{c}{=}-\lambda+|z|^p\log\lambda$

$$\operatorname*{argmax}_{\lambda} L_{\mathcal{P}}(\lambda)=|z|^p\longrightarrow L_{\mathcal{P}}(|z|^p)\geq L_{\mathcal{P}}(\lambda)$$

$$L_{\mathcal{P}}(|z|^p)-L_{\mathcal{P}}(\lambda)=|z|^p\log\frac{|z|^p}{\lambda}-(|z|^p-\lambda)\geq 0$$

5. 確率過程の構築

周波数成分の確率密度関数の設定

- フーリエ成分 $\boldsymbol{z}=(z_1,\dots,z_I)^{\mathrm{T}}\in\mathbb{C}^I$ の各要素が独立に複素ポアソン分布に従う→ $z_i\sim\mathrm{cPois}(\lambda_i,p)$ $(i=1,\dots,I)$

$$\rightarrow\quad f_{\boldsymbol{z}}(\boldsymbol{z};\boldsymbol{\lambda},p)=\prod_{i=1}^{I}\frac{pe^{-\lambda_i}}{2\pi}\frac{|z_i|^{p-2}\lambda_i^{|z_i|^p}}{(|z_i|^p)!}$$ ただし, $\boldsymbol{\lambda}=(\lambda_1,\dots,\lambda_I)^{\mathrm{T}}$

信号の確率密度関数

- 変数変換 $\boldsymbol{x}=\boldsymbol{F}^{-1}\boldsymbol{z}$, $|\mathrm{det}\boldsymbol{F}|=1$

$\rightarrow\quad f_{\boldsymbol{x}}(\boldsymbol{x};\boldsymbol{\lambda},p)=f_{\boldsymbol{z}}(\boldsymbol{F}\boldsymbol{x};\boldsymbol{\lambda},p)$ ただし，サポートは $\mathcal{D}_x=\{\boldsymbol{x}\in\mathbb{R}^I\mid[\boldsymbol{F}\boldsymbol{x}]_i\in\mathcal{D},\forall i\}$

パラメータ $\boldsymbol{\lambda}$ の意味

- Example 1 より，$p=2$ のとき，λ_i は $\mathbb{E}[|z_i|^2]$ に対応
  周波数ビン i の
  パワースペクトル密度

最尤スペクトル推定

Property 5 より，$\boldsymbol{z}\sim f_{\boldsymbol{z}}(\boldsymbol{z};\boldsymbol{\lambda},2)$ のとき，
$\lambda_1,\dots,\lambda_I$ の最尤推定 ＝ 標本スペクトル $|z_1|^2,\dots,|z_I|^2$ と
パワースペクトル密度 $\lambda_1,\dots,\lambda_I$ の Iダイバージェンス最小化

まとめ（概要）

- Iダイバージェンス規準でスペクトル密度推定を行うことがどのような確率過程の信号モデルを仮定していることになっているかを明らかにした（下記参照）。

定常ガウス過程		今回提案の確率過程
複素正規分布 $z_i\sim\mathcal{N}_{\mathbb{C}}(0,\lambda_i)$ $(i=1,\dots,I)$	フーリエ成分の確率分布	複素ポアソン分布 $z_i\sim\mathrm{cPois}(\lambda_i,2)$ $(i=1,\dots,I)$
パワースペクトル密度	パラメータ $\boldsymbol{\lambda}$	パワースペクトル密度
$|z_1|^2,\dots,|z_I|^2$ と $\lambda_1,\dots,\lambda_I$ の板倉斎藤距離最小化と等価	$\boldsymbol{\lambda}$ の最尤推定	$|z_1|^2,\dots,|z_I|^2$ と $\lambda_1,\dots,\lambda_I$ の Iダイバージェンス最小化と等価

今後の展望

- 提案の確率過程に基づくブラインド信号処理（音源分離, 残響除去など）の枠組の構築（例えば，[4]の統計的枠組による定式化）
- 上記枠組に Iダイバージェンスと相性の良いパラメトリックなスペクトルモデル（例えば[5]）を導入
- 複素ポアソン分布のその他の性質を解明

参考文献

[1] F. Itakura and S. Saito, “Analysis synthesis telephony based upon the maximum likelihood method,” In Proc. 6th Int’l Cong. Acoust. (ICA’68), C-5-5, pp. C17–20, 1968.
[2] I. Csiszar, “Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems,” The Annals of Statistics, vol. 19, no. 4, 1991.
[3] P. O. Amblard, M. Gaeta, and J. L. Lacoume, “Statistics for complex variables and signals - part i: Variables,” Signal Processing, vol. 53, no. 1, 1996.
[4] 安良岡直希, 亀岡弘和, 吉岡拓也, 奥乃博, “補助関数法を用いた I ダイバージェンス規準残響抑圧," 日本音響学会2011年春季研究発表会, 3-1-9, 発表予定, 2011.
[5] H. Kameoka, “Statistical approach to multipitch analysis,” Ph.D. dissertation, The University of Tokyo, 2007.