

現代工学基礎 I

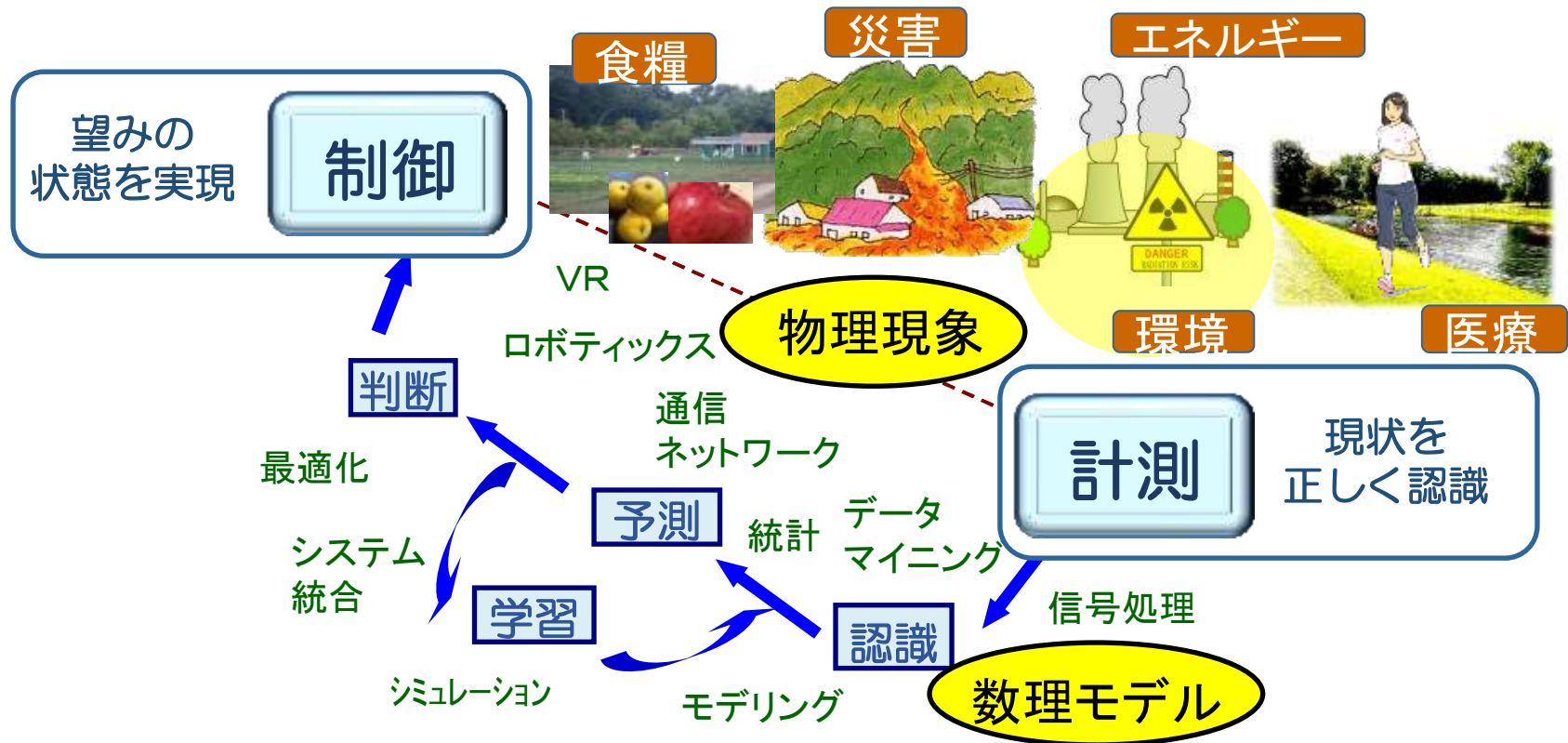
物理情報システム工学

工学部 計数工学科

猿渡 洋, 妹尾 拓, 近藤正章

物理情報システム工学

- 計測と制御による問題解決とシステム実現に向けて -

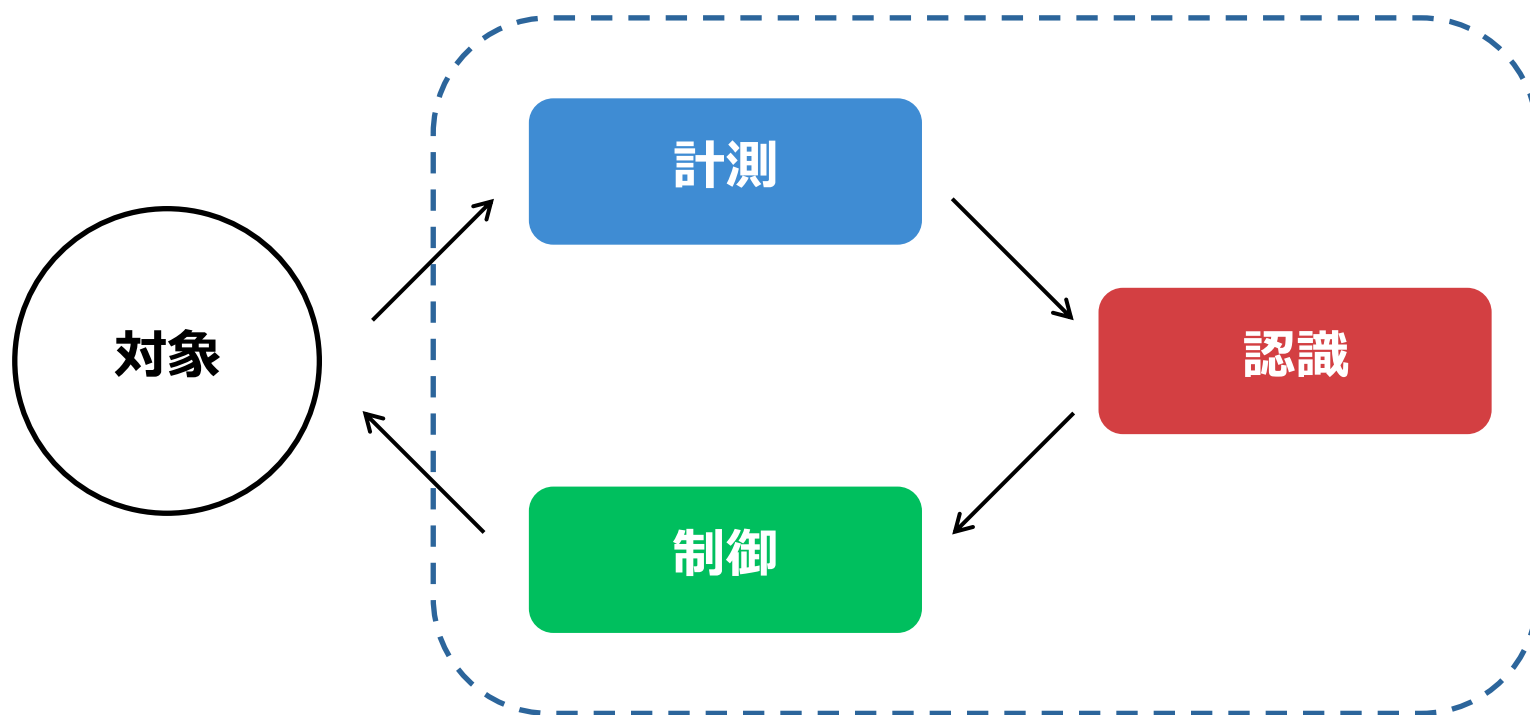


本講義の目的

- 物理的, 生物的, 人工的, 社会的, どんな対象であれ, そこに望みの機能を工学的に実現するために重要なこと
 1. 対象で生じる物理現象のモデリングと解析,
 2. 計測による物理現象の本質の情報化,
 3. 情報処理と制御による所望の機能の最適な実現
 4. 情報の認識とシステムインテグレーションによる実世界への働きかけ,
- この工学的方法論のエッセンスを講義

システムのデザイン

- 機能要素の統合
- 物理世界と情報世界を繋ぐ
- 「認識し、行動する」テクノロジーの創造



日程

- 9/26 ガイダンス 担当:猿渡、近藤、妹尾
- 10/3 休講
- 10/10 計測工学と信号処理の基礎:情報理解と実世界インテリジェントシステムへの応用(1) 担当:猿渡
- 10/17 計測工学と信号処理の基礎:情報理解と実世界インテリジェントシステムへの応用(2) 担当:猿渡
- 10/24 計測工学と信号処理の基礎:情報理解と実世界インテリジェントシステムへの応用(3) 担当:猿渡
- 11/ 1 制御工学の基礎:横断的制御理論とロボット制御への応用(1) 担当:妹尾
- (11/7 10/9↓へ振り替え)
- 11/ 9(金) 本郷研究室見学 (夕方～ 駒場5限後でも参加可能)

日程

- 11/21 制御工学の基礎：横断的制御理論とロボット制御への応用(2)
担当：妹尾
- 11/28 制御工学の基礎：横断的制御理論とロボット制御への応用(3)
担当：妹尾
- 12/ 5 情報の認識と計算システムの基礎：人工知能技術とそれを可能にする
高性能計算技術(1) 担当：近藤
- 12/12 先端科学技術研究センター生田・池内研究室見学会(仮)
- 12/19 情報の認識と計算システムの基礎：人工知能技術とそれを可能にする
高性能計算技術(2) 担当：近藤
- 1/ 9 情報の認識と計算システムの基礎：人工知能技術とそれを可能にする
高性能計算技術(3) 担当：近藤

猿渡 洋：計測工学と信号処理の基礎：

情報理解と実世界インテリジェント システムへの応用(1)

講義資料と成績評価

■講義資料

- <http://www.sp.ipc.i.u-tokyo.ac.jp/~saruwatari/>

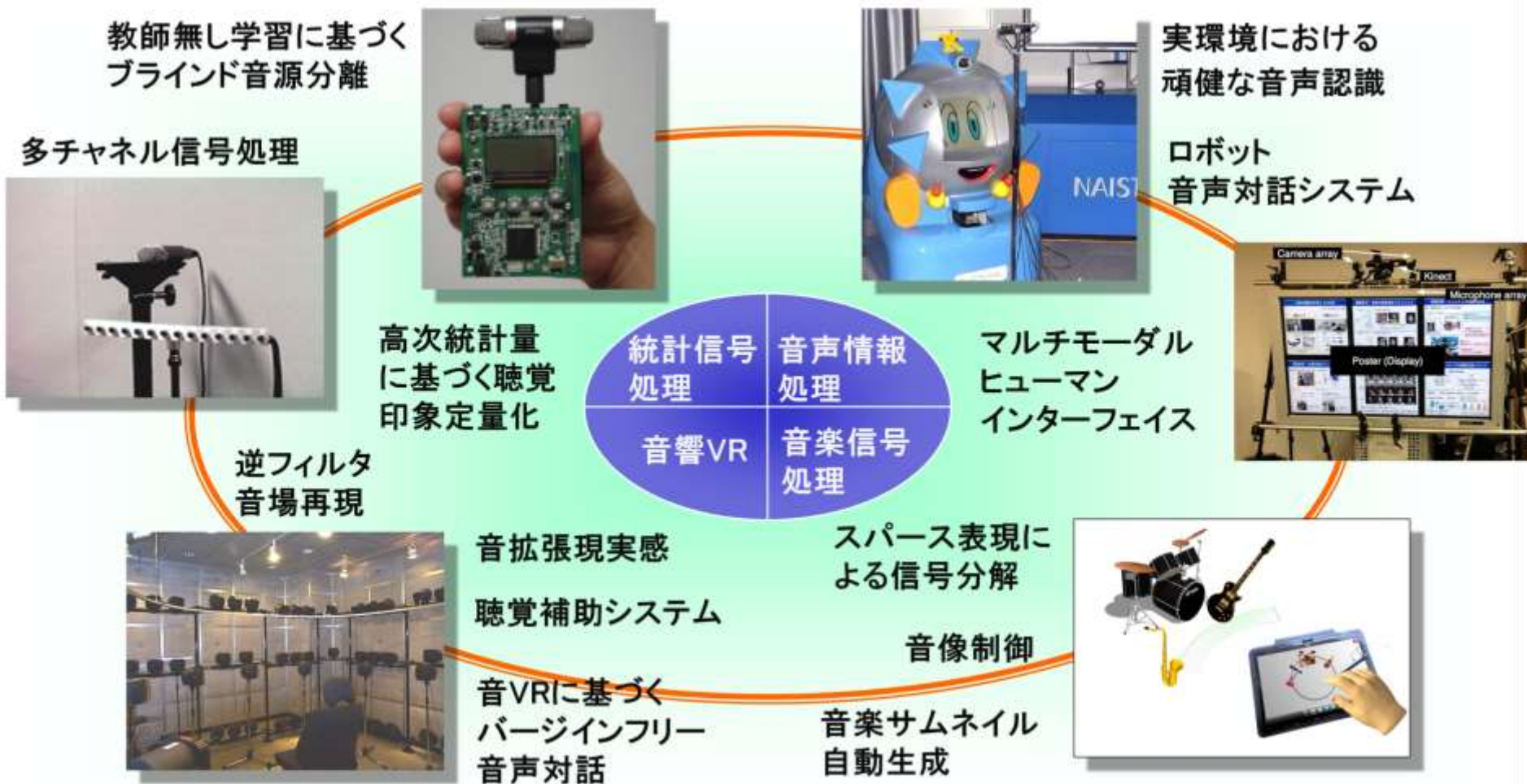
※計数工学科/情報理工学系研究科・システム情報第一研究室(猿渡研)からたどれるようにしておきます

■成績評価

- レポート課題
- 出席

自己紹介：私の研究俯瞰図

- 音声・音響・音楽メディアに関する信号処理・情報処理
- ヒューマンインターフェイス・コミュニケーションシステムの構築
- 統計的信号処理、機械学習、非線形システム解析、等



こんな機械に知性を感じる？



君の名前は何？
何ができるの？

えっ、人間型
ロボットなのに
スイッチで入
力するの？

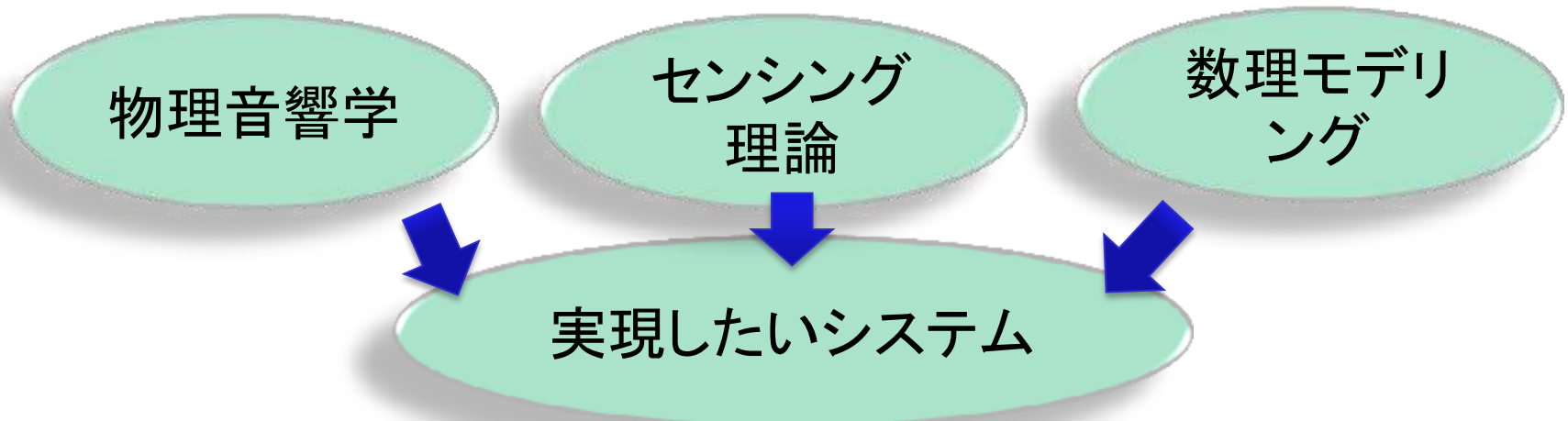


音メディアの重要性

- **音声は人間が持つ最も原始的なコミュニケーション手段**
 - ・ 特別な道具を必要としない誰でも使える意思伝達手段
 - ・ 人間同士もしくはコンピュータやロボットとのコミュニケーション手段の一つとして昔から注目されてきた
- **人間と人間のインタラクションにおける音声**
 - ・ 音声符号化⇒携帯電話等にて既に実用化
- **機械と人間のインタラクションにおける音声**
 - ・ 知性を感じさせる機械を実現するには音声対話メディアが必須である
 - ・ 必要とされる機能
 - ・ 必要な音を聞き分ける(雑音抑圧・音源分離)
 - ・ ユーザの声を言語として認識理解する(音声認識)
 - ・ ユーザに情報を発信する(音声合成・声質変換)

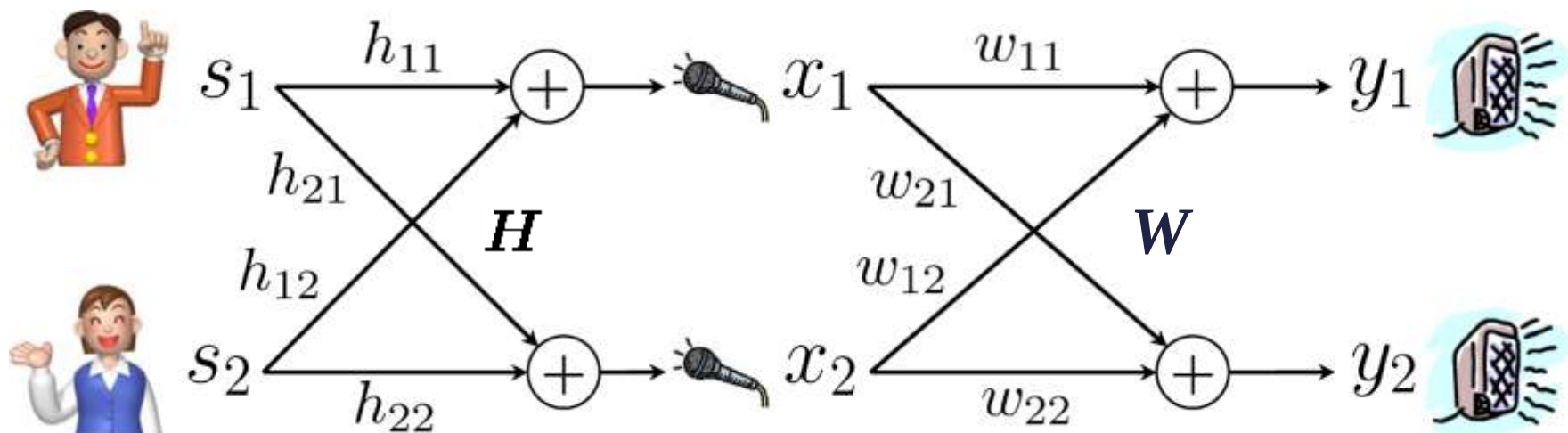
音メディアに関する信号処理研究の魅力

- 音メディアに関する信号処理研究の魅力とは？
 - 自然界の音が持つ無限の**多様性**（人の声・音楽・騒音 etc.）
 - 研究のアプローチに**多面性**あり（音は波？それとも確率？）
 - 最後は聴かせてなんぼの評価 ⇒ **芸術性**も併せ持つ
- 「**物理世界（波動）と情報世界（抽象）をまたぐ学問**」であり、かつそれを「**統一的に取り扱うシステム工学**」である。
- 対象の多様性ゆえに「**なんでもあり**」の分野でもある。



応用紹介: ブラインド音源分離 (BlindSourceSeparation)

- 混ざり合った信号 x_1, x_2 から元の信号を取り出す
- どのように混ざったかに関する空間情報 H は利用できない



実は上記は**2つのことを同時に推定**している

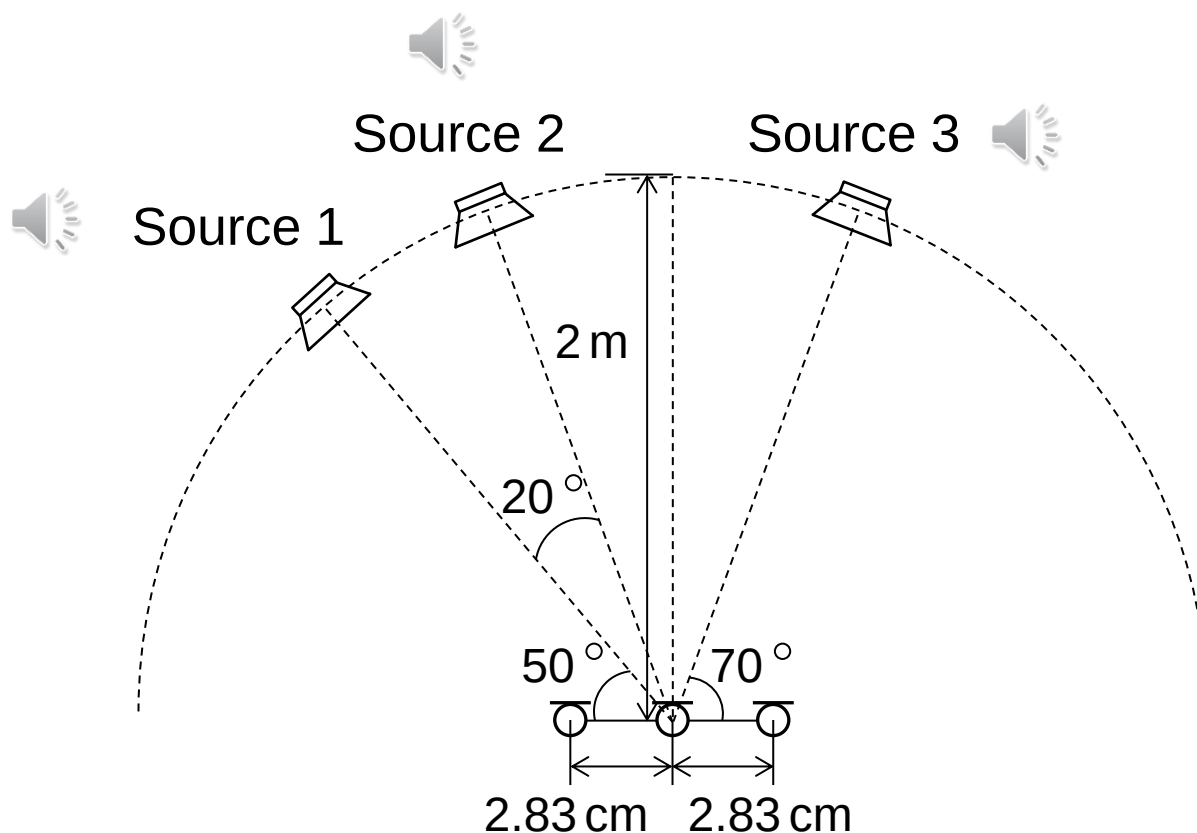
- [空間] 統計的に独立な音源の分類問題 (分離行列 W の推定)
- [信号] 各音源が属する確率分布 $p(y)$ の推定問題

上記を閉形式で解く方法は存在せず凸問題でもない⇒**大変困難!**

応用紹介：低ランク近似音源モデルに基づくBSS

[Kitamura, et al. 2015]

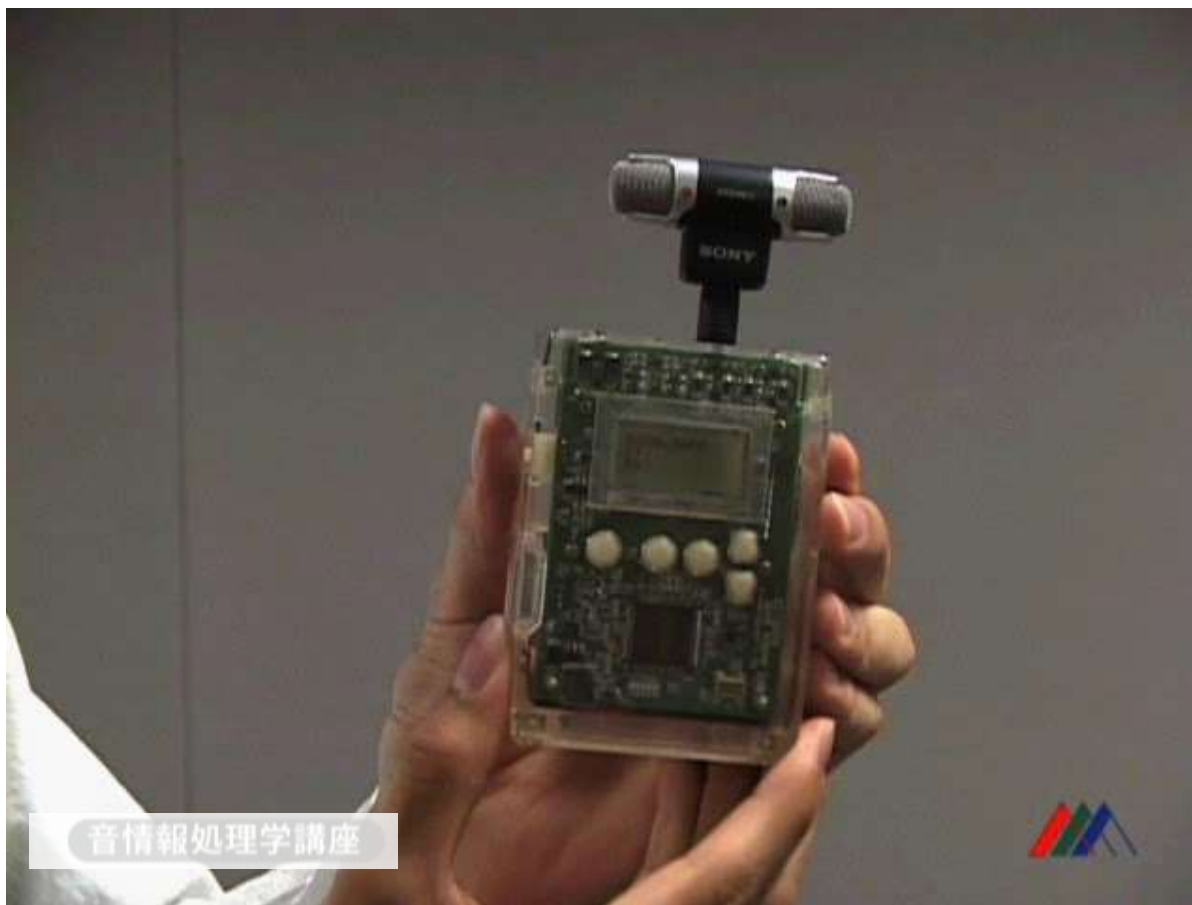
- ドラム、ストリング、音声からなる複合音の分離



応用紹介:「聖徳太子マイク」 リアルタイムBSS

[Saruwatari, et al. 2009]

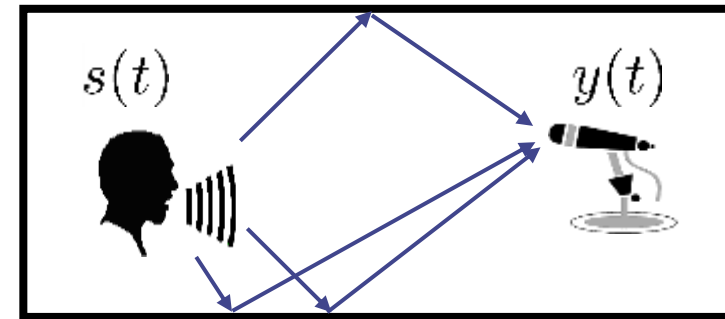
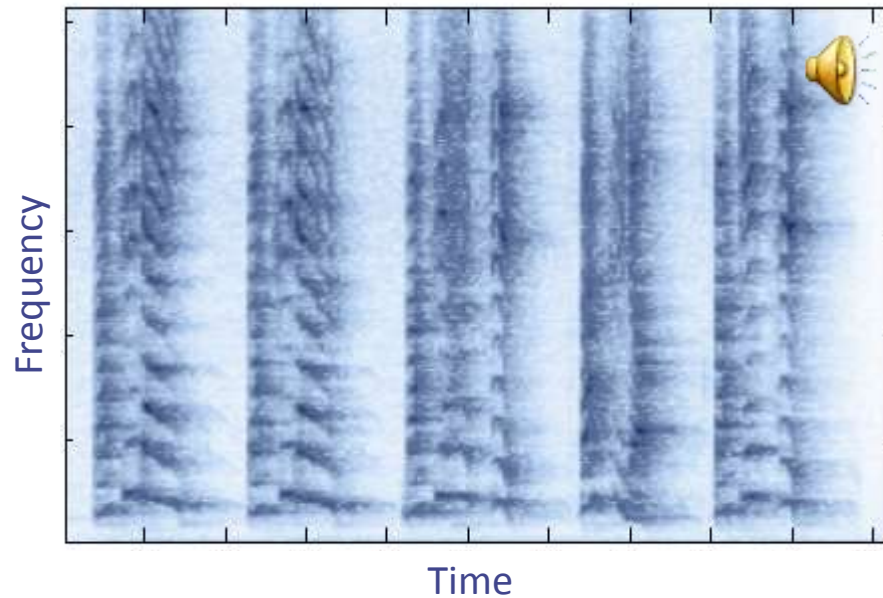
- 4chマイクとDSP(TI社製C67)から構成される独立成分分析BSS専用モジュールを2005年に開発
- 2009年に世界初の商用化(警察備品に採用)



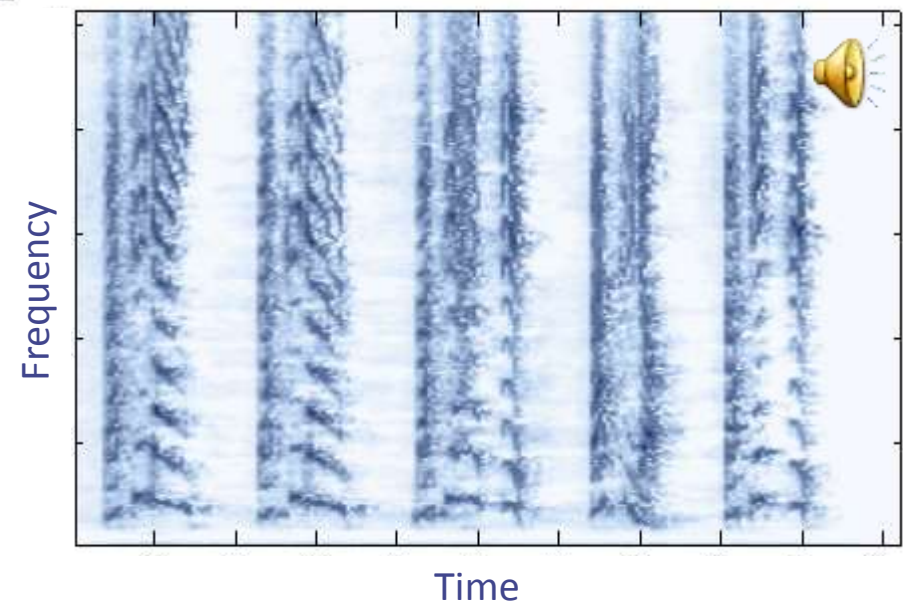
応用紹介: 残響抑圧

[Kameoka, et al.]

■ 観測信号のスペクトログラム



■ 残響除去信号のスペクトログラム



応用紹介: 音声合成



入力 x と出力 y の関係をどう記述するか？ → **統計的逆問題**

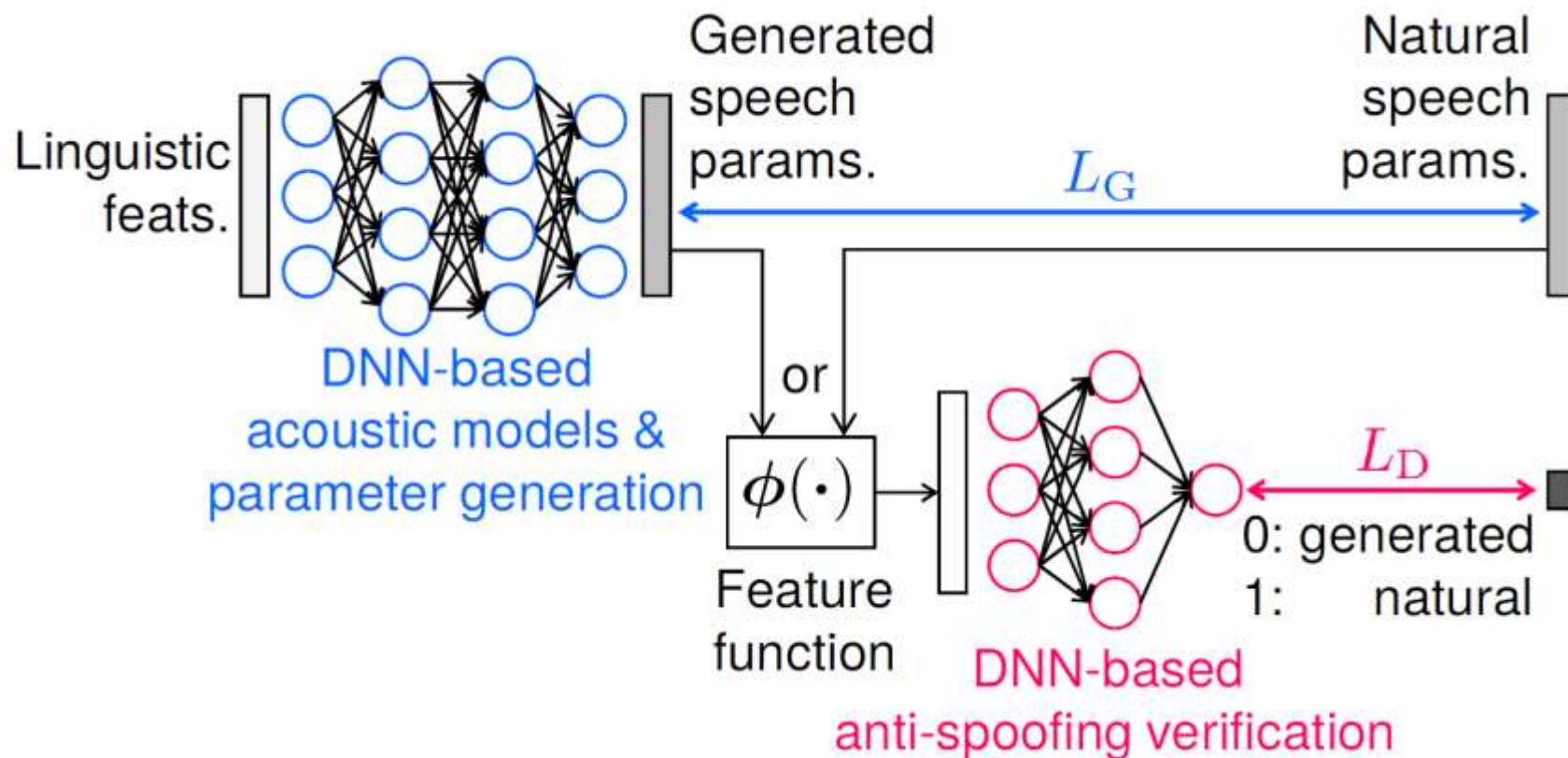
人間らしい声とは何か？ **「人間らしい声」の統計モデル化**

東京大学は世界一！
昼飯はとんこつラーメンに限る！



応用紹介: DNN-GAN音声合成

[Saito, Takamichi, et al. 2016]



人間の声に似せ
ようと努力

ウソ(合成音)に
騙されまいと攻防

応用紹介: 統計的時系列推定における音質の差

[Saruwatari, et al. 2012]

白色ノイズの場合

観測音



最尤推定



ベイズ推定



ミュージカル
ノイズフリー



人ごみノイズの場合

観測音



最尤推定



ベイズ推定



ミュージカル
ノイズフリー



どの推定方式が音として「自然」か？

良い統計的性質を持つ推定法 $\neq$ 音響的に優れた技術

応用紹介：音バーチャルリアリティ

[Koyama, et al. 2014]

物理的なスピーカ列はここ



信号処理における統計の数学的準備 ～特性関数、モーメント、最尤推定 etc.～

基礎事項

◆ 確率変数と確率分布

◆ 確率変数

- 「サイコロの目」のようにランダムに値が決まる対象を表すもの

◆ 確率分布

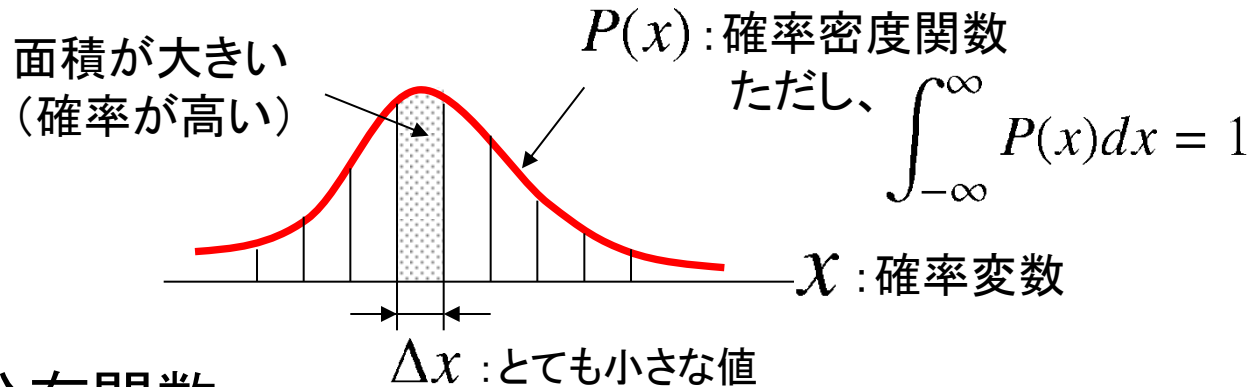
- 確率変数が持つ値が起こる確率のこと
- 確率変数が持ちうる値全てについての確率分布の総和は1
- 確率分布は必ず0以上

音の信号を確率変数として扱い
その特徴を確率分布により表現することを考える

基礎事項 — 確率分布を表す関数

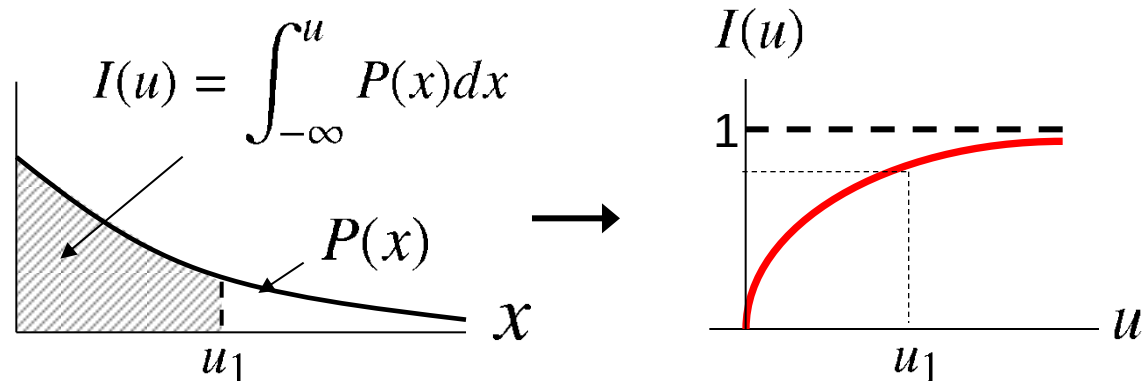
◆ 確率密度関数

確率の大きさを表す関数(Δx に乗ずる確率の密度)



◆ 累積分布関数

確率変数がある値以下の値を持つ確率を表す関数で、
確率密度関数の定積分で与えられる



基礎事項 — 期待値、分散

◆ 期待値

- ◆ 確率の重みつき平均
- ◆ ただの「平均」と同義語として使われたりする

$$\alpha_1 = \int_{-\infty}^{\infty} xP(x)dx = \underbrace{E(x)}_{\text{期待値演算}}$$

x : 確率変数

◆ 分散

- ◆ 確率分布の期待値からの「ばらつき度合い」を表す値

$$\alpha_2 = E\{(x - \underbrace{\mu}_{\text{期待値}})^2\}$$

- ◆ 分散の平方根を偏差と呼ぶ

基礎事項 — 確率分布形状と歪度・尖度

◆ 歪度 (Skewness)

- ◆ 確率分布形状の期待値についての「非対称度合い」を表す値

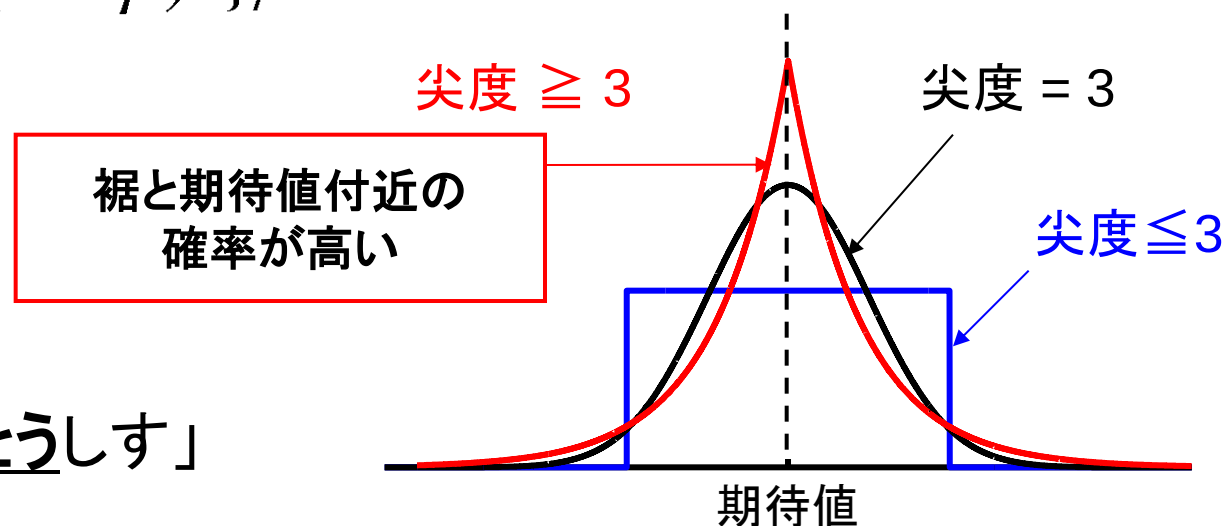
$$\alpha_3 = E\{(x - \mu)^3\} / \sigma^3$$

$$\sigma = \sqrt{\alpha_2} : \text{標準偏差}$$

◆ 尖度 (Kurtosis)

- ◆ 確率分布形状の裾と期待値周りの確率の重みを表す値

$$\alpha_4 = E\{(x - \mu)^4\} / \sigma^4$$



※発音は「かーとうしす」

モーメント

◆ モーメント(積率)

期待値、分散、歪度、尖度の式を眺めてみると、以下が各式に含まれていた

$$E\{(x - \mu)^n\}$$

分布の形状はこの値によって決まることがわかる

これを、「**期待値周りの n 次のモーメント**」と呼ぶ

$\mu = 0$ としたときは「**原点周りの n 次のモーメント**」と呼ぶ

◆ 特性関数

期待値、分散、歪度、尖度などの値を指定すれば、確率分布の候補は制限されてくる。その極限として、**全ての次数のモーメントを指定すれば、確率分布は一意に定まる(その逆も成り立つ)**。そのような、全ての次数のモーメントを生成する関数が以下の特性関数である

$$m_x(it) = \int_{-\infty}^{\infty} e^{itx} P(x) dx = E(e^{itx})$$

$$\left[\frac{\partial^{(n)} m_x(it)}{\partial it^{(n)}} \right]_{t=0} = \mu_n \quad : n \text{次モーメント}$$

特性関数の展開

◆ モーメント展開

モーメント展開は指数関数の展開により、以下で与えられる

$$\begin{aligned}m_x(it) &= E(e^{itx}) \\&= 1 + itE(x) + \frac{1}{2}(it)^2E(x^2) + \frac{1}{3!}(it)^3E(x^3) + \cdots \\&= \sum_{n=0}^{\infty} \frac{(it)^n}{n!}E(x^n)\end{aligned}$$

この展開によって任意の次数のモーメントが求まるが、 $(it)^n$ の級数は無限次まで続く。よって、この展開(モーメント展開)では、特定の次数までの確率分布の情報しか得られない。



確率分布の情報をより有効的に表現する展開が**キュムラント展開**

キュムラント展開 [1/2]

特性関数が以下のように展開されたとする

$$\begin{aligned}
 m_x(it) &= E(e^{itx}) \\
 &= \exp \left\{ (it)cum(x) + \frac{1}{2}(it)^2 cum(x^2) + \frac{1}{3!}(it)^3 cum(x^3) + \dots \right\} \quad (*) \\
 &= \exp \left\{ \sum_{n=0}^{\infty} \frac{(it)^n}{n!} cum(x^n) \right\} \\
 &= \exp\{cum(e^{itx}) - 1\}
 \end{aligned}$$

↑
特性関数をこの形に展開できるような
量 $cum(x^n)$ を導入

ここで、上式を次のように展開する

$$\begin{aligned}
 m_x(it) &= E(e^{itx}) \\
 &= \exp \{(it)cum(x)\} \exp \left\{ \frac{1}{2}(it)^2 cum(x^2) \right\} \exp \left\{ \frac{1}{3!}(it)^3 cum(x^3) \right\} \dots \\
 &= \left\{ 1 + (it)cum(x) + \frac{1}{2}(it)^2 cum(x)^2 + \dots \right\} \cdot \left\{ 1 + \frac{1}{2}(it)^2 cum(x^2) + \dots \right\} \dots \\
 &= 1 + (it) cum(x) + \frac{1}{2}(it)^2 \{cum(x^2) + cum(x)^2\} \\
 &\quad + \frac{1}{3!}(it)^3 \{cum(x^3) + 3cum(x^2)cum(x) + cum(x)^3\} \dots
 \end{aligned}$$

キュムラント展開 [2/2]

これまでに得られた以下の二つの展開を比較すると

$$\text{モーメント展開 : } E(e^{itx}) = 1 + itE(x) + \frac{1}{2}(it)^2 E(x^2) + \frac{1}{3!}(it)^3 E(x^3) + \dots$$

$$\begin{aligned} \text{前項の展開 : } E(e^{itx}) &= 1 + (it) \text{cum}(x) + \frac{1}{2}(it)^2 \{ \text{cum}(x^2) + \text{cum}(x)^2 \} \\ &\quad + \frac{1}{3!}(it)^3 \{ \text{cum}(x^3) + 3\text{cum}(x^2)\text{cum}(x) + \text{cum}(x)^3 \} \dots \end{aligned}$$

各次モーメントについて以下の関係が成り立つ

$$\begin{cases} E(x) = \text{cum}(x) \\ E(x^2) = \text{cum}(x^2) + \text{cum}(x)^2 \\ E(x^3) = \text{cum}(x^3) + 3\text{cum}(x^2)\text{cum}(x) + \text{cum}(x)^3 \end{cases}$$

よって前項で導入した謎の量は以下で表されることがわかる

$$\left. \begin{aligned} \text{cum}(x) &= E(x) \\ \text{cum}(x^2) &= E(x^2) - E(x)^2 \\ \text{cum}(x^3) &= E(x^3) - 3E(x^2)E(x) + 2E(x)^3 \end{aligned} \right\} \begin{aligned} &\text{これらを**キュムラント**と呼び、} \\ &\text{前項の式(*)を} \\ &\text{**キュムラント展開**と呼ぶ} \end{aligned}$$

**キュムラント展開は低次モーメントの情報を用いて
確率分布の情報をより有効的に表現できる方法**

正規分布(ガウス分布)のキュムラント

◆ 正規分布では3次以上のキュムラントは全てゼロ

正規分布の特性関数を求め、キュムラント展開すると

$$\begin{aligned} m_x(itx) &= \int_{-\infty}^{\infty} \exp(itx) P(x) dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp(itx) \exp\left\{-(x-\mu)^2/2\sigma^2\right\} dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp(it\mu - t^2\sigma^2/2) \int_{-\infty}^{\infty} \exp-[x - (\mu + it\sigma^2)]^2/2\sigma^2 dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp(it\mu - (it)^2\sigma^2/2) \sqrt{2\pi\sigma^2} \\ &= \exp(it\mu - (it)^2\sigma^2/2) \end{aligned}$$

上式をp.9のキュムラント展開式と比較し、正規分布のキュムラントを求めると

$$\text{1次キュムラント} \quad cum(x) = \mu$$

$$\text{2次キュムラント} \quad cum(x^2) = \sigma^2$$

$$\text{3次以上のキュムラント} \quad cum(x^k) = 0 \quad (k \geq 3)$$

[参考] 大数の法則、中心極限定理

◆ 大数の法則

観測された標本の平均は、標本の数が大きくなると、母集団（観測された標本が元々含まれていた集団）の平均に近づく

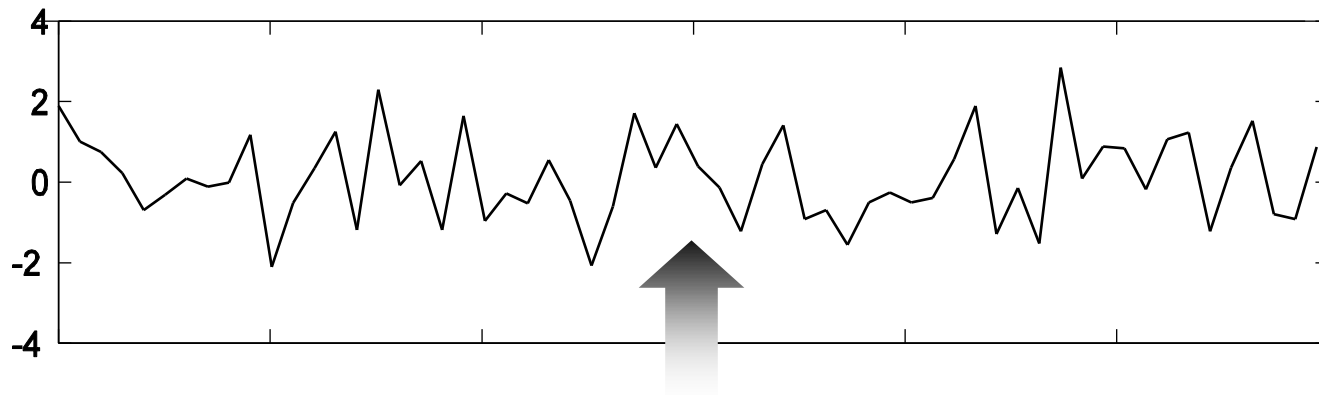
◆ 中心極限定理

同一の分布に従う、互いに独立な標本 $x^{(1)}, x^{(2)}, \dots, x^{(N)}$ の和の確率分布は、 $N \rightarrow \infty$ とともに以下の標準正規分布となる

標準正規分布

$$P(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \leftarrow \text{平均0, 分散1 の正規分布}$$

分布を用いた統計



一見複雑に見えるこの信号は、ある確率密度関数のたった2つの変数(これを**母数**と呼ぶ)だけを用いて生成された(この場合、平均と分散を与えただけ)

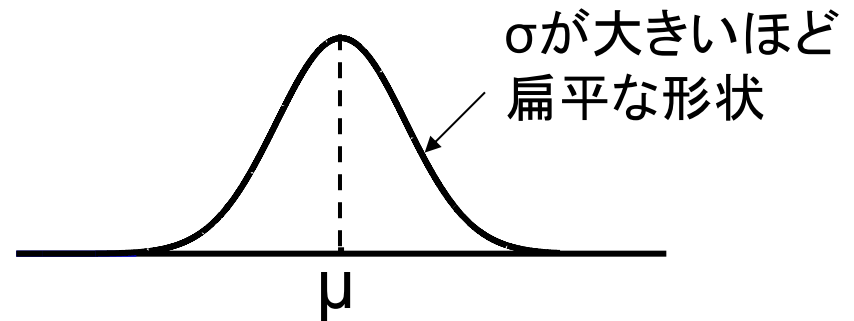
ならば、逆に、こんな複雑な信号の発生原因も2個のパラメータを用いて表現できるかも

そこで、一見複雑な信号を少数のパラメータを用いて表現することを目標とした上で、そのための確率密度関数をいくつか挙げ、その母数推定について述べる

正規分布(ガウス分布)

◆ 期待値 μ , 分散 σ^2 の正規分布の確率密度関数

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}$$



◆ 特徴

- ◆ 多くの物理現象を正規分布で表すことができると言われている
 - 例: 身長、実験誤差
- ◆ 数学的に扱いやすい性質を持つ場合が多い
- ◆ 確率のさまざまな定理に関係している(中心極限定理など)
- ◆ 正規分布の歪度は0、尖度は3
 - 歪度以上の高次の統計量は、正規分布による信号のモデル化の可否を決める判断基準となる
 - 尖度が3より大きな分布に従う確率変数を「優ガウス(スーパーガウス)的」、3より小さな分布を「劣ガウス(サブガウス)的」と呼んだりする

カイ2乗分布

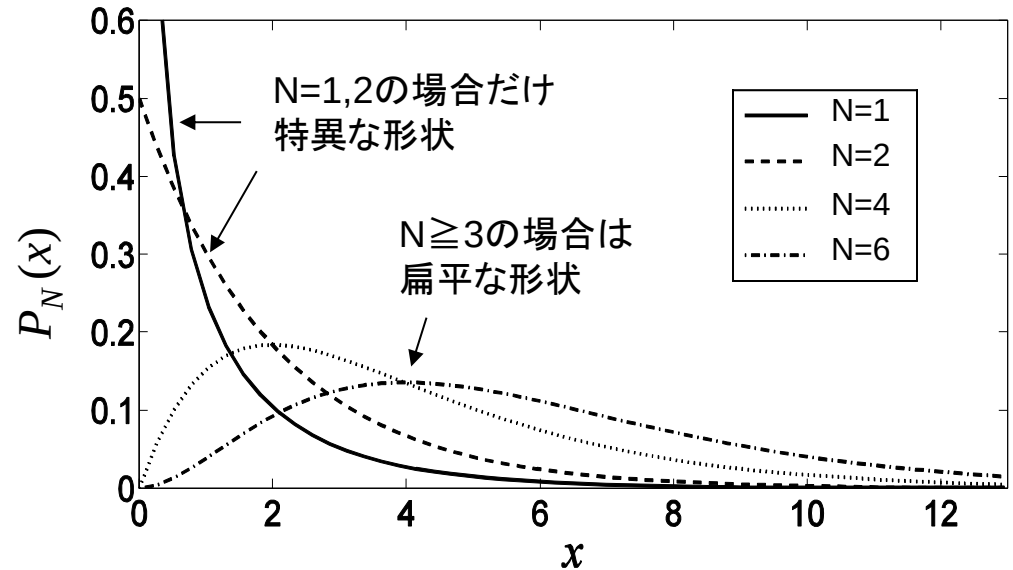
◆ 確率密度関数

$$P_N(x) = \frac{1}{2^{N/2}\Gamma(N/2)} x^{(N/2)-1} e^{-x/2}$$

ただし、 $x \geq 0$, $P(x) = 0$ ($x < 0$)

$$\Gamma(\alpha) = \int_{-\infty}^{\infty} x^{\alpha-1} e^{-x} dx$$

これをガンマ関数と呼ぶ



- ◆ 標準正規分布(※)に従う互いに独立な N 個の確率変数の2乗の和は、 N をパラメータに持つカイ2乗分布に従う

- ◆ 実際、この分布は物理的にどんな意味を持つのか

$N=1$ のカイ2乗分布に従う信号, つまり標準正規分布に従う信号のパワーは、標準正規分布に従う時間領域信号のパワースペクトルを意味する

ガンマ分布

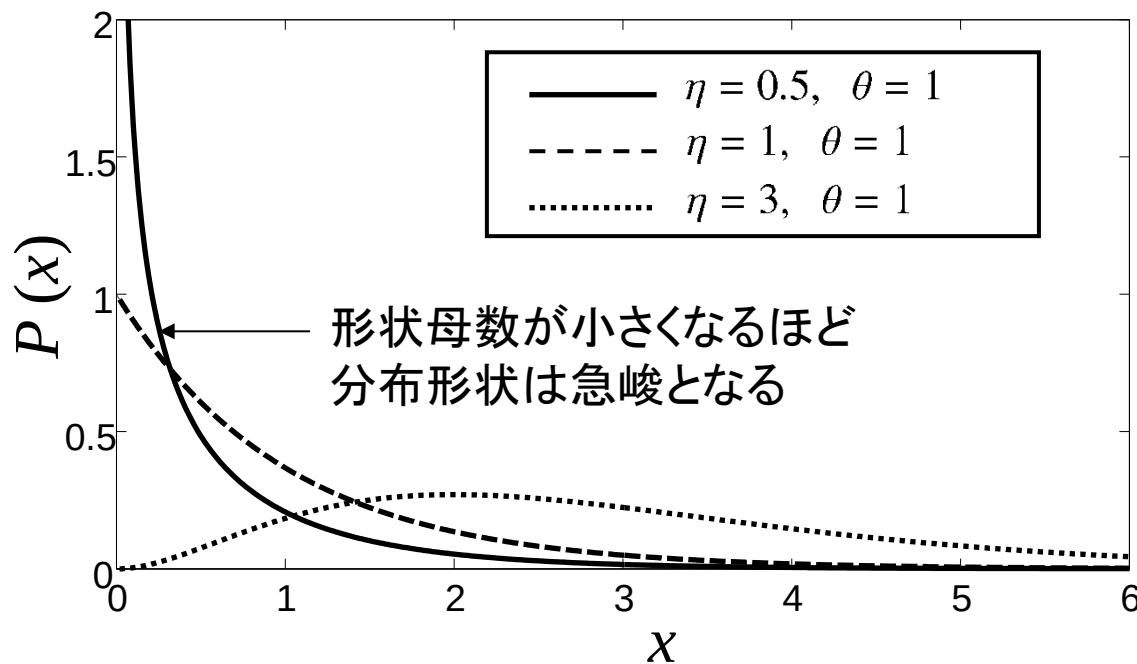
◆ 確率密度関数

$$P(x) = x^{\eta-1} \frac{e^{-x/\theta}}{\theta^{\eta} \Gamma(\eta)} \quad \begin{array}{l} \text{ただし、} x \geq 0, \quad P(x) = 0 \ (x < 0) \\ \eta > 0, \quad \theta > 0 \end{array}$$

◆ ガンマ分布の母数

η : 形状母数 (分布形状の急峻さを決めるパラメータ)

θ : 尺度母数 (分布形状の広がりを決めるパラメータ)



ガンマ分布と形状母数

ガンマ分布の特徴

◆ ガンマ分布の各平均量

- ◆ 平均: $\eta\theta$
- ◆ 分散: $\eta\theta^2$
- ◆ 尖度 (Kurtosis): $\frac{(\eta + 2)(\eta + 3)}{\eta(\eta + 1)}$

◆ 片側分布であるため、0以上の値を持つ信号の表現に適している

- ◆ 信号の「強さ」(振幅、パワーなど)を表現できる

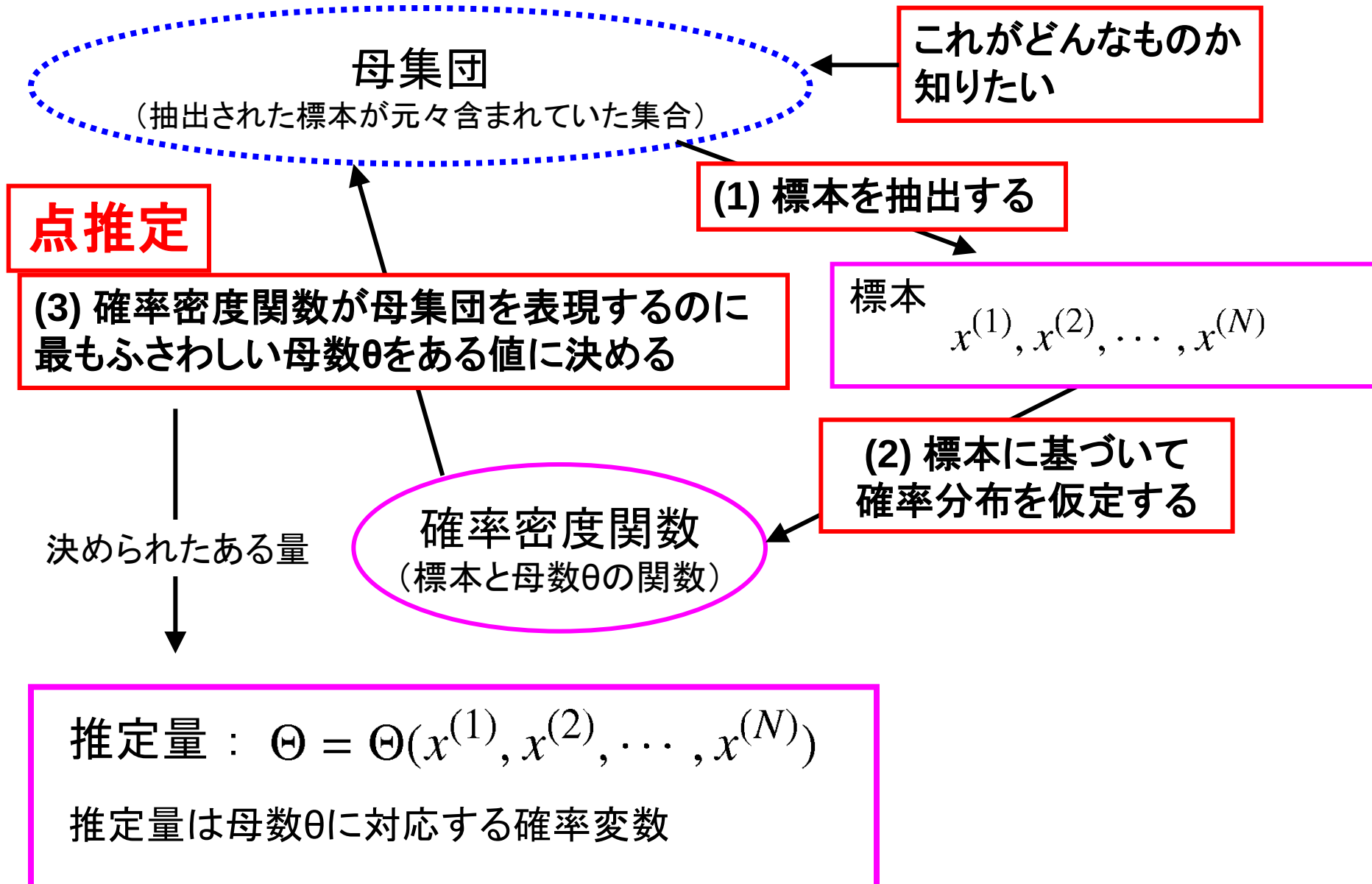
◆ さまざまな形状の分布をガンマ分布で表現できる

- ◆ 形状母数が1の場合、 $N = 1$ のカイ2乗分布となる
- ◆ 形状母数が1より小さい場合、優ガウ斯的な信号を2乗したものが従う分布となる
- ◆ 形状母数が1より大きい場合、劣ガウ斯的な信号を2乗したものが従う分布となる



ガンマ分布は信号のパワーの表現に適した分布

統計的推定の考え方



最尤(さいゆう)法の定義

母集団の確率密度関数を以下で表す

$$W(x, \theta) \quad x \text{ は確率変数、} \theta \text{ は母数}$$

同一の母集団から、無作為抽出によって少数の標本 $x^{(1)}, x^{(2)}, \dots, x^{(N)}$ が得られる確率は、**尤度関数**と呼ばれる以下の関数で表される

$$W(x^{(1)}, \theta) W(x^{(2)}, \theta) \cdots W(x^{(N)}, \theta) \equiv L(\theta) \leftarrow \begin{array}{l} \text{母数についての} \\ \text{関数} \end{array}$$

尤度関数が最も大きくなるように値が決められた θ は、標本の発生源(母集団)がどのような確率分布に従っているかを説明するのに尤もらしい

このように、最も尤度の大きい母数の推定値 $\hat{\theta}$ (最尤推定値)を求める方法を**最尤法**と呼ぶ

対数尤度関数とその用例

- ◆ 尤度関数の代わりに、以下の対数尤度関数を用いてもよい

$$l(\theta) \equiv \ln L(\theta) = \sum_{l=1}^N W(x^{(l)}, \theta)$$

- ◆ 対数尤度関数を最大にする θ は、以下の式の解となる

$$\frac{dl(\theta)}{d\theta} = 0$$

[例] 正規分布の最尤推定値

確率密度は $P(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$

この場合、対数尤度は $l(\theta) = \sum_{l=1}^N \ln P(x^{(l)}, \theta) = -\frac{1}{2\sigma^2} \sum_{l=1}^N \left[(x^{(l)} - \mu)^2 - \frac{N}{2} \ln 2\pi\sigma^2 \right]$

母数 θ は μ と σ^2 だから、対数尤度を偏微分して0とおいた方程式を解くと

$$\frac{dl(\theta)}{d\mu} = -\frac{1}{2\sigma^2} \sum_{l=1}^N 2(x^{(l)} - \mu) = 0 \quad \text{より、} \quad \hat{\mu} = \frac{1}{N} \sum_{l=1}^N x^{(l)}$$

$$\frac{dl(\theta)}{d\sigma^2} = -\frac{1}{2\sigma^4} \sum_{l=1}^N (x^{(l)} - \mu)^2 - \frac{N}{2\sigma^2} = 0 \quad \text{より、} \quad \hat{\sigma}^2 = \frac{1}{N} \sum_{l=1}^N (x^{(l)} - \hat{\mu})^2$$

標本を用いて
母数の最尤推定値が
直ちに求まる

結局、最尤法って何？

確率密度関数を $f(x|\theta)$ と表したときの対数尤度は

$$l(\theta) \equiv \ln L(\theta) = \sum_{l=1}^N \ln f(x^{(l)}|\theta) \quad (1)$$

ここで、大数の法則を思い出すと、

$$\frac{1}{N} \sum_{l=1}^N x^{(l)} \xrightarrow{N \rightarrow \infty} \mu = \langle z \rangle \quad (2)$$

Nを大きくとると母平均 μ に限りなく近づく
これを確率変数 z の期待値として表す
ことにする

z の確率密度関数を $W(z)$ (真の確率密度関数)とし、その期待値は以下となる

$$\mu = \int_{-\infty}^{\infty} z W(z) dz \quad (3)$$

確率密度関数 $f(x|\theta)$ の対数の平均を求めると、大数の法則より

$$\frac{l(\theta)}{N} = \frac{1}{N} \sum_{l=1}^N \ln f(x^{(l)}|\theta) \xrightarrow{N \rightarrow \infty} \langle \ln f(z|\theta) \rangle \equiv I_2(\theta) \quad (4)$$

期待値を I_2 とおく

期待値 I_2 も式(3)の z の期待値の形式で書くと

$$I_2(\theta) = \int_{-\infty}^{\infty} \ln f(z|\theta) W(z) dz \quad (5)$$

以下の対応関係から
 $z \longleftrightarrow \ln f(z|\theta)$

平均対数尤度と呼び、最尤推定量の評価などに用いられる

結局、最尤法って何？

$f(z | \theta) \leftarrow$ 手元にある標本だけをもとにして仮定する関数なので、
真の確率密度関数 $W(z)$ には一致しないことが多い

真の確率密度関数がもしわかった場合、平均対数尤度は

$$I_1 = \int_{-\infty}^{\infty} \ln(W(z)) W(z) dz \quad (6)$$

よって、真の確率密度関数に基づく I_1 と、最尤法に基づく $N \rightarrow \infty$ での I_2 の差は

$$I \equiv I_1 - I_2(\theta) = \int_{-\infty}^{\infty} W(z) \ln \frac{W(z)}{f(z | \theta)} dz \quad (7)$$



これはKL情報量と呼ばれ、真の分布と別の分布との「距離」を表す量

結局、最尤法って何？

KL情報量は必ず非負の値を持つ

$$I \geq 0 \quad (8)$$

また、 $p \geq 0$ という値について、以下の不等式が成り立つ

$$\ln p \leq p - 1 \quad (9)$$

ここで、式(7)のKL情報量を変形し、式(9)の関係を用いると

$$\begin{aligned} -I &= \int_{-\infty}^{\infty} W(z) \ln \frac{f(z | \theta)}{W(z)} dz \\ &\leq \int_{-\infty}^{\infty} W(z) \left\{ \frac{f(z | \theta)}{W(z)} - 1 \right\} dz \\ &= \int_{-\infty}^{\infty} f(z | \theta) dz - \int_{-\infty}^{\infty} W(z) dz = 0 \end{aligned} \quad (10)$$

したがって、 $I \geq 0$ となり、真の分布 $W(z)$ と $f(z | \theta)$ が一致するときだけ0となる



分布間距離 I をなるべく小さくするように母数 θ を決めてやれば
真の分布に近い分布を見つけることができる

結局、最尤法って何？

式(7)における I_1 は標本に依らないので、 I_2 を大きくすればするほど分布間距離は小さくなる



I_2 は対数尤度関数を N で割ったものであり、対数尤度関数に対応、
よって、 **I を最小化することは、対数尤度を最大化することと同じ**

結局、最尤法は

KL情報量(分布間距離)を最も小さくするように母数を決め、
真の分布になるべく近い $f(z | \theta)$ を探す方法

ガンマ分布の母数の最尤推定 [1/2]

ガンマ分布の確率密度関数は以下で与えられる

$$P(x) = x^{\eta-1} \frac{e^{-x/\theta}}{\theta^{\eta} \Gamma(\eta)} \quad \eta : \text{形状母数} \quad \theta : \text{尺度母数}$$

この場合の対数尤度関数は

$$l(\eta, \theta) = \ln \prod_{i=1}^N P(x_i) = (\eta - 1) \sum_{i=1}^N \ln x_i - \sum_{i=1}^N \frac{x_i}{\theta} - N\eta \ln \theta - N \ln \Gamma(\eta)$$

ここで、以下の関係を用いて対数尤度関数を形状母数のみで表すと

$$\frac{1}{N} \sum_{i=1}^N x_i = \eta \theta$$
$$l(\eta, \theta) = (\eta - 1) \sum_{i=1}^N \ln x_i - N\eta - N\eta \ln \left(\frac{1}{N\eta} \sum_{i=1}^N x_i \right) - N \ln \Gamma(\eta)$$

上式を形状母数で微分したものを0と置くと

$$\frac{\partial l(\eta, \theta)}{\partial \eta} = \sum_{i=1}^N \ln(x_i) - N \ln \left(\frac{1}{N} \sum_{i=1}^N x_i \right) - N \ln \eta - N\psi(\eta) = 0$$

ただし、 $\psi(\eta) = \frac{\Gamma(\eta)'}{\Gamma(\eta)}$

ガンマ分布の母数の最尤推定 [2/2]

よって、以下の方程式を解くことで形状母数が求まる

$$\frac{1}{N} \sum_{i=1}^N \ln(x_i) - \ln \left(\frac{1}{N} \sum_{i=1}^N x_i \right) = \ln \eta - \psi(\eta)$$

ただし、**解析的に求まらない**ので、上式の右辺を以下のように近似し、左辺を定数 s と置くと

$$s = \ln \eta - \psi(\eta) \approx \frac{1}{\eta} \left(\frac{1}{2} + \frac{1}{12\eta + 2} \right)$$

$$\longrightarrow s = \frac{3\eta + 6}{6\eta^2 + \eta}$$

よって、形状母数は以下の方程式を解くことで得られる

$$6s\eta^2 + (s - 3)\eta + 6 = 0$$

したがって、形状母数と尺度母数は以下で得られる

$$\left\{ \begin{array}{l} \hat{\eta} \approx \frac{3 - s + \sqrt{(s - 3)^2 + 24s}}{12s}, \\ \hat{\theta} \approx \frac{1}{\eta N} \sum_{i=1}^N x_i \end{array} \right. \quad s = \frac{1}{N} \sum_{i=1}^N \ln(x_i) - \ln \left(\frac{1}{N} \sum_{i=1}^N x_i \right)$$

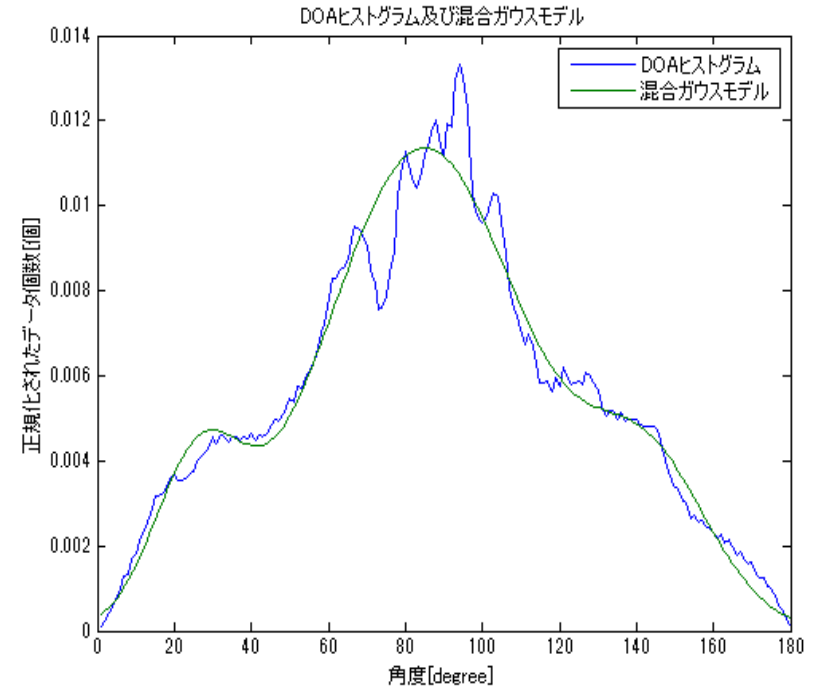
[参考] さらに進んだ分布推定方法: EMアルゴリズム

- Gaussian Mixture Model (GMM)を例にとって考える

平均 μ , 分散 σ^2 , 混合係数 π

混合音源の生成モデルを,
混合ガウスモデルと仮定

$$p(x) = \sum_{k=1}^K \pi_k \mathcal{N}(x | \mu_k, \sigma_k^2)$$



- 混合係数 π : k 番目のガウス分布が、全体の分布にどれだけ混合しているかの割合
- 負担率 γ_{nk} : データ x_n が k 番目のガウス分布から生成される確率
= データ x_n について k 番目のガウス分布の負担する比率

[参考] EMアルゴリズムの概略

- GMMパラメータの最尤推定

GMMのパラメータ: $\Theta = \{\mu_k, \sigma_k, \pi_k\}_{k=1, \dots, K}$

$\operatorname{argmax}_{\Theta} p(x|\Theta) = \operatorname{argmax}_{\Theta} \log p(x|\Theta)$ を求める

- 対数尤度関数

$\log p(x|\Theta) = \sum_n \log \sum_{k=1}^K \pi_k \mathcal{N}(x|\mu_k, \sigma_k^2)$ は解析的に解けないため補助関数法を用いる

- 補助関数 $Q(\Theta, \gamma)$

補助関数は $Q(\Theta, \gamma) = \sum_n \sum_{k=1}^K \gamma_{nk} \log \frac{\pi_k \mathcal{N}(x_n; \mu_k, \sigma_k^2)}{\gamma_{nk}}$

$\log p(x|\Theta) = \min_{\gamma} Q(\Theta, \gamma)$ を満たす

Θ と γ の交互反復更新アルゴリズムによる局所最大化を考える

[参考] 具体的なEMアルゴリズム

- Eステップ(負担率の更新)

データ x_n が k 番目の分布から生成されたということを $z_{nk} = 1$,
 $\sum_k z_{nk} = 1$ として, その期待値が負担率 γ_{nk}

$$\gamma(z_{nk}) = \frac{\pi_k \mathcal{N}(x_n | \mu_k, \sigma_k^2)}{\sum_{k=1}^K \pi_k \mathcal{N}(x_n | \mu_k, \sigma_k^2)}$$

- Mステップ(平均, 分散, 混合係数の更新)

$\sum_k \pi_k = 1$ の制約条件下での $\max_{\Theta} Q(\Theta, \gamma)$ を考えることにより,

$$\mu_k = \frac{1}{N_k} \sum_n \gamma_{nk} x_n, \quad \sigma_k^2 = \frac{1}{N_k} \sum_n \gamma_{nk} (x_n - \mu_k)^2, \quad \pi_k = \frac{N_k}{N}$$

但し, $N_k = \sum_{n=1}^N \gamma_{nk}$ であり, N は総データ数

統計的推定理論の応用例：音声信号処理

統計的システムモデリングとは？

まず生成モデルを立てる(順問題)

- ① 尤度関数の仮定
 - 観測データ Y を生成する確率的なプロセス $p(Y|\theta)$ をモデル化
- ② 事前分布の仮定
 - 生成モデルのパラメータ θ の生成プロセス $p(\theta)$ をモデル化



観測データから情報を抽出(逆問題)

- データ Y から θ と α を推論
- 最尤推定量 $\hat{\theta} = \underset{\theta}{\operatorname{argmax}} p(Y|\theta)$, MAP推定量 $\hat{\theta} = \underset{\theta}{\operatorname{argmax}} p(\theta, \alpha|Y)$
MMSE推定量 $\hat{\theta} = \mathbb{E}[\theta, \alpha|Y]$, ベイズ事後分布 $p(\theta, \alpha|Y)$

パラメータ推定

データの確率的な生成プロセスの仮定(順問題)

⇔ 観測データの確率モデル化

データから生成プロセスのパラメータの推定(逆問題)

⇔ 最尤推定, 最大事後確率推定,
最小平均二乗誤差推定, ベイズ推論

$p(y|\theta)$, $p(\theta)$ をモデル化

順問題

θ $\longrightarrow$ y

逆問題

ベイズの定理

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta}$$

ML推定量, MAP推定量, MMSE推定量

■ データ y が与えられた下でのパラメータ θ の推定量

■ 最尤(Maximum Likelihood)推定量

$$\hat{\theta}_{\text{ML}} = \underset{\theta}{\operatorname{argmax}} p(Y|\theta)$$

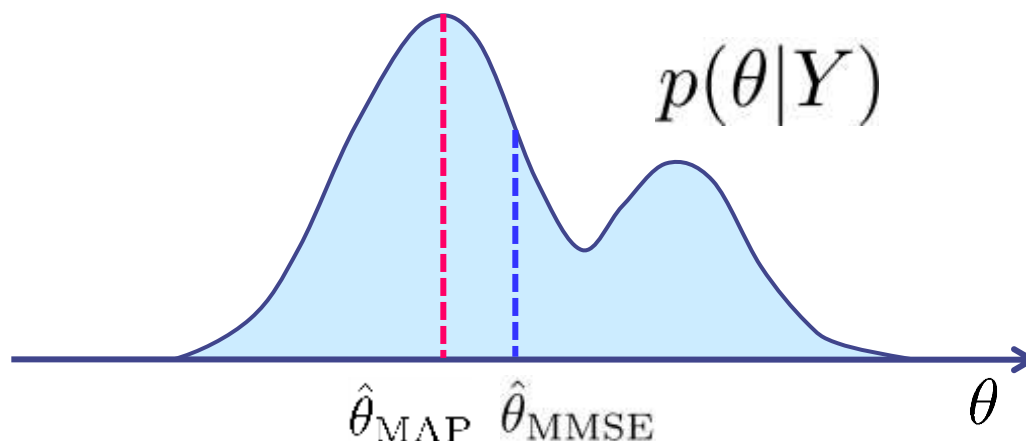
MAP推定で $p(\theta) \propto 1$
を仮定した場合に相当

■ 最大事後確率(Maximum A Posteriori)推定量

$$\hat{\theta}_{\text{MAP}} = \underset{\theta}{\operatorname{argmax}} p(\theta|y) = \underset{\theta}{\operatorname{argmax}} p(y|\theta)p(\theta)$$

■ 最小平均二乗誤差(Minimum Mean Squared Error)推定量

$$\hat{\theta}_{\text{MMSE}} = \underset{\hat{\theta}}{\operatorname{argmin}} \mathbb{E}[(\hat{\theta} - \theta)^2|Y] = \mathbb{E}[\theta|Y]$$

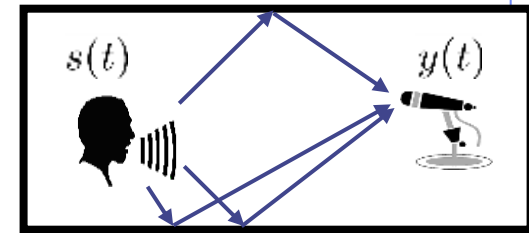
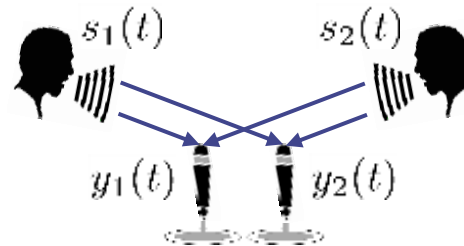


音声音響信号処理問題の多くは逆問題

■音響信号処理

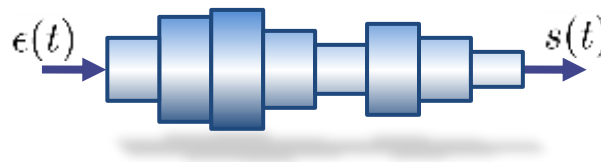
- ブラインド音源分離
- 残響除去

...



■音声情報処理

- 音素特徴抽出
- 音声認識
- イントネーション解析



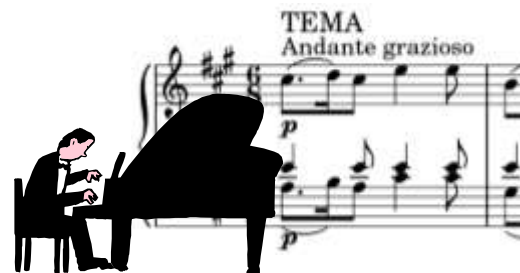
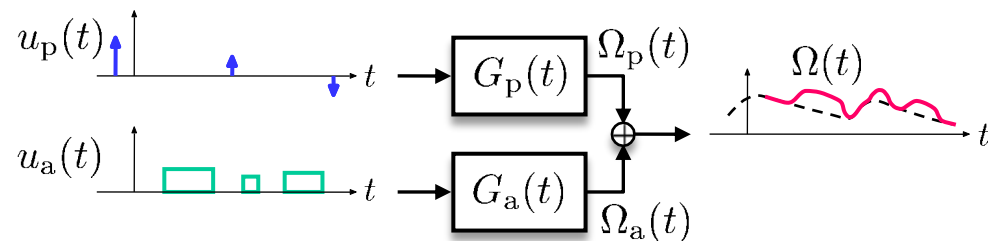
今日は晴天なり



■音楽情報処理

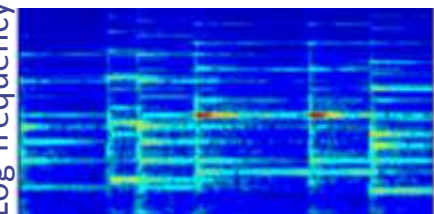
- 多重音解析
- 自動採譜

...



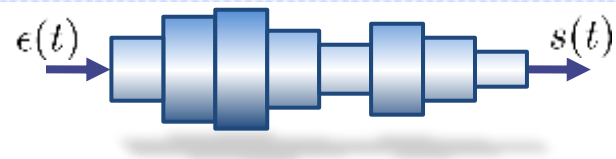
TEMA
Andante grazioso

Log-frequency



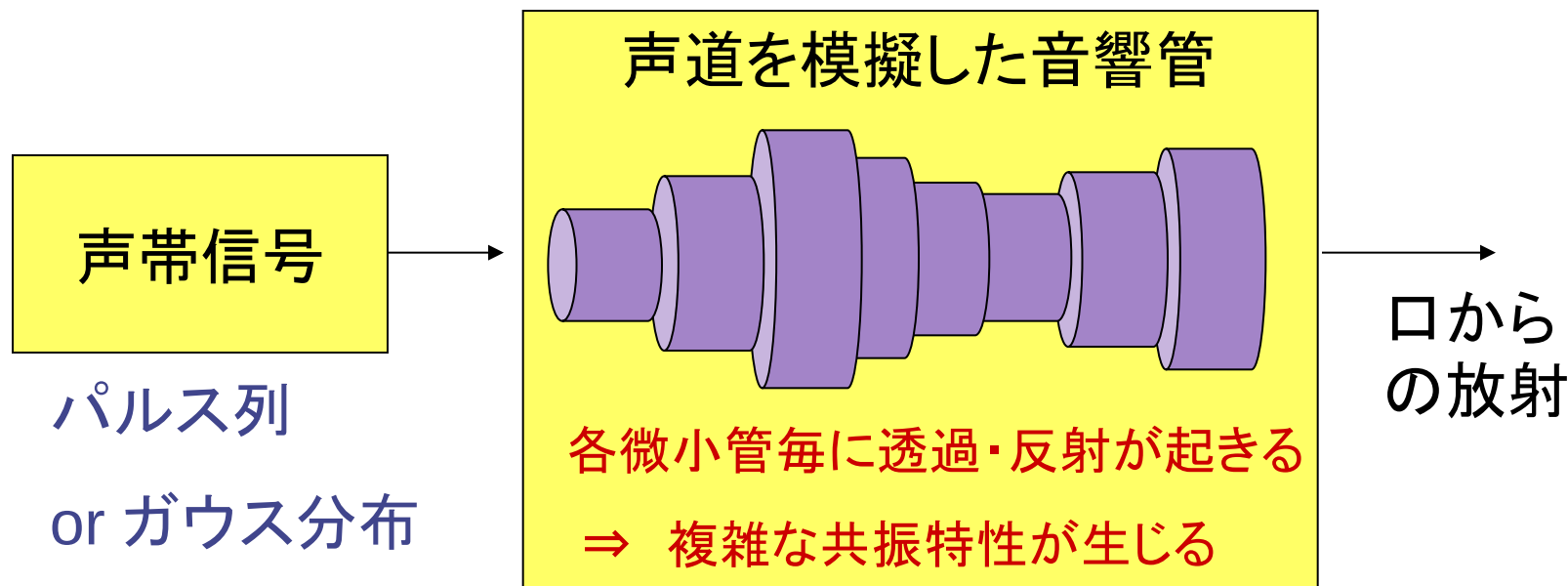
Time

携帯電話における逆問題



■ 人間の音声生成モデル

- 声帯での基本振動を声道で音色付ける。
- 声道
 - 位置によって太さの異なる音響管の連続と見なせる。
 - 音響管における共振現象 ⇒ 自己回帰 (AR) 過程



この生成モデルパラメータを求めて伝送する⇒大幅な情報圧縮

音声のスペクトル構造

■ 短時間スペクトル

- 音声は、短時間区間ごとの電圧スペクトル密度（周波数領域におけるパワー特性）で測ることが多い。

■ 音声スペクトル構造の2要素

- 周波数とともにゆるやかに変化する成分[スペクトル包絡]
⇒発声器官の共振・反共振特性を表す
(つまり人間の喉・口の形をあらわす特徴量)
- 細かく周期的(有声音; 母音などの場合)または非周期的(無声音の場合)に変化する成分 [スペクトル微細構造]
⇒音源の周期性
(つまり声帯の基本周期・声の高低を表す特徴量)

音声信号のスペクトルはこれら2つの要素の積で表される

音声のスペクトル構造

人間頭部の断面図

②

鼻腔

声道

の共振

音声の音色を与える
(スペクトル包絡構造)

① × ② = 最終的な音声

Nasal Cavity

Hard Plate

Soft Plate (Velum)

Pharynx

Epiglottis

Larynx

Vocal Cords
(Glottis)

Esophagus

Trac

Lungs

by Beng Tiong Tan

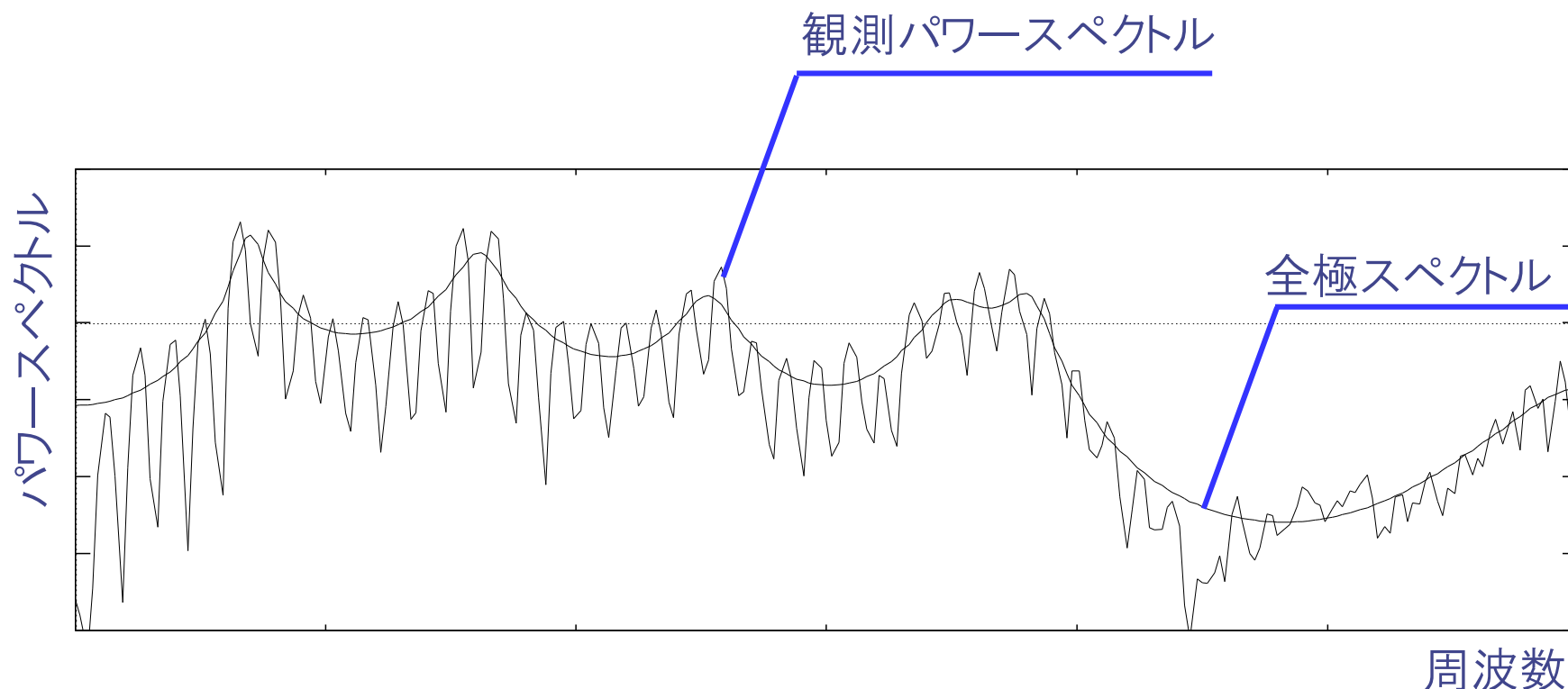
①

声帯の振動

音声の基本周期を与える
(スペクトル微細構造)

最尤推定によって得られた「声色」

- 音声信号だけを眺めてもその性質は分からない
 - ⇒ 滑らかな周波数特性を持つ確率的な信号であると仮定
 - ⇒ 最尤推定によってその周波数特性パラメータを推定
 - ⇒ これは母音・子音毎に異なるので音声認識や伝送に利用可能
- 観測パワースペクトルと全極スペクトルとの板倉斎藤距離最小化



ロボット対話 デモンストレーション

**Demonstration for
spoken-oriented hands-free
interactive robot.**