

月4 2013 冬学期 [4830-1032]

第4回

音声音響信号処理

(線形予測分析と自己回帰モデル)

亀岡弘和

東京大学大学院情報理工学系研究科

日本電信電話株式会社

NTTコミュニケーション科学基礎研究所

講義内容(キーワード)

- 信号処理、符号化、標準化の実用システム例の紹介
- 情報通信の基本(誤り検出、訂正符号、変調、IP)
- 符号化技術の基本(量子化、予測、変換、圧縮)
- 音声分析・合成・認識・強調、音楽信号処理
- 統計的信号処理の基礎(スペクトル、ガウス過程、最尤推定)
- ガウス性確率変数の基本性質
- 時間周波数分析(短時間フーリエ変換、ウェーブレット変換)
- ウィナーフィルタとカルマンフィルタ
- 音声生成過程のモデル(ソースフィルタ理論と藤崎モデル)
- 自己回帰モデルと線形予測分析
- 独立成分分析によるブラインド音源分離
- 非負値行列因子分解によるスペクトログラムの分解表現
- スペクトル間擬距離
- 最適化アルゴリズム(EMアルゴリズム、補助関数法)

講義スケジュール

10/ 7 守谷先生担当

10/15 (火) 守谷先生担当

10/21 守谷先生担当

10/28 休講

→ 11/ 5 (火) 線形予測分析と自己回帰モデル

11/11

11/18

11/25

12/ 2

12/ 9

12/16

12/23

1/13

1/20

1/27

成績評価

■レポート課題

- 本講義に関連する論文を1つ選び、発表資料形式(パワーポイント等)にまとめて学期末に提出してください。提出先は最終講義にてお知らせします。
- 「どの程度本質を理解しているか」「要点が分かりやすく記述されているか」「なぜその論文を重要と考えたか」を評価の規準にして採点します。
- 毎回の講義後にその回の講義に関連する論文を1つ挙げる予定です。それらの中から選んでも良いですし、自分で自由に探してきてもOKです。

■講義の感想

- レポートとともに講義に対する感想文も一緒に提出して下さい。
- ※講義資料は講義用ホームページにアップしていく予定です。

講義URL

- <http://hil.t.u-tokyo.ac.jp/~kameoka/SAP/>

本日の話題

■ 線形予測分析 (Linear Predictive Coding)

- 音声情報処理研究の歴史の幕開けとなった信号処理技術 (統計的手法を取り入れた初めての音声研究として有名)
 - 音声分析合成(ボコーダ)
 - 音声音響符号化
 - 音声認識のための音声特徴量
 - 音声強調(残響除去、ブラインド音声分離)
などへの応用
- 日本発の技術としても知られる
 - Levinson-Durbin-板倉アルゴリズム、偏自己相関(PARCOR)、線スペクトル対(Line Spectrum Pair)の発明や板倉齋藤距離の発見など、板倉文忠氏(名古屋大学名誉教授)の電電公社時代の活躍が世界的に有名

線形予測分析

■ 3つの観点から解説

- 「予測誤差」を最小化する観点
 - 最小二乗誤差推定
- 線形系としての観点
 - 自己回帰系 (AutoRegressive system)
 - 音声の生成過程モデル
 - 最尤推定
 - 白色化
- スペクトルマッチングとしての観点
 - 最尤スペクトル推定
 - 板倉斎藤距離

線形予測分析

■ 3つの観点から解説

■ 「予測誤差」を最小化する観点

■ 最小二乗誤差推定

■ 線形系としての観点

■ 自己回帰系 (AutoRegressive system)

■ 音声の生成過程モデル

■ 最尤推定

■ 白色化

■ スペクトルマッチングとしての観点

■ 最尤スペクトル推定

■ 板倉斎藤距離

「予測誤差」を最小化

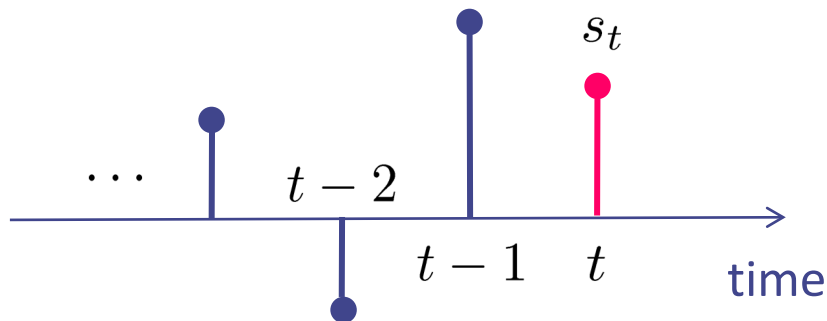
■ 動機: 符号化への応用

- 少ないパラメータで音声信号を表現したい

■ 問題: 線形予測誤差の最小化

- 時刻 t の信号のサンプル値 s_t を、過去のサンプル値 $s_{t-1}, s_{t-2}, \dots, s_{t-P}$ の線形結合で「予測」
- 「予測」の誤差を最小にするには結合係数(予測係数という)をどう置けば良い？

$$s_t \simeq \sum_{p=1}^P \underline{a_p} s_{t-p}$$



最小二乗誤差推定による定式化

■すべての t で $s_t \simeq \sum_{p=1}^P a_p s_{t-p}$ となる a_1, \dots, a_P を求めたい

■目的関数

$$\mathcal{J}(a_1, \dots, a_P) = \sum_t \left(s_t - \sum_{p=1}^P a_p s_{t-p} \right)^2 \rightarrow \text{minimize}$$

■最小解では $\frac{\partial \mathcal{J}}{\partial a_1} = \frac{\partial \mathcal{J}}{\partial a_2} = \dots = \frac{\partial \mathcal{J}}{\partial a_P} = 0$ を満たすため...

$$\begin{aligned} \longrightarrow \frac{\partial \mathcal{J}}{\partial a_q} &= 2 \sum_t \left(s_t - \sum_{p=1}^P a_p s_{t-p} \right) (-s_{t-q}) \\ &= 2 \left(\sum_p a_p \sum_t s_{t-p} s_{t-q} - \sum_t s_t s_{t-q} \right) \\ &= 0 \quad (q = 1, \dots, P) \end{aligned}$$

最小二乗誤差推定による定式化

■ 連立方程式に帰着

$$\left\{ \begin{array}{l} \sum_p a_p \sum_t s_{t-p} s_{t-1} = \sum_t s_t s_{t-1} \\ \sum_p a_p \sum_t s_{t-p} s_{t-2} = \sum_t s_t s_{t-2} \\ \vdots \\ \sum_p a_p \sum_t s_{t-p} s_{t-P} = \sum_t s_t s_{t-P} \end{array} \right. \longrightarrow c_{q,p} = \sum_t s_{t-p} s_{t-q}$$

以上より最適予測係数は
以下の方程式を満たす

$$\begin{bmatrix} c_{1,1} & \cdots & c_{1,P} \\ \vdots & \ddots & \vdots \\ c_{P,1} & \cdots & c_{P,P} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_P \end{bmatrix} = \begin{bmatrix} c_{0,1} \\ \vdots \\ c_{0,P} \end{bmatrix}$$

この方程式をYule-Walker方程式という

Levinson-Durbin-Itakuraアルゴリズム (1/6)

■ Yule-Walker方程式

$$\begin{bmatrix} c_{1,1} & \cdots & c_{1,P} \\ \vdots & \ddots & \vdots \\ c_{P,1} & \cdots & c_{P,P} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_P \end{bmatrix} = \begin{bmatrix} c_{0,1} \\ \vdots \\ c_{0,P} \end{bmatrix}$$

$$c_{p,q} = v_{|p-q|}$$

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_{P-1} \\ v_1 & v_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & v_1 \\ v_{P-1} & \cdots & v_1 & v_0 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_P \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_P \end{bmatrix}$$

Toeplitz行列

右辺と左辺
の関係が特殊

連立一次方程式の解き方

- 一般の場合：
Gaussの消去法
- 左辺が正値対称行列の場合：
Cholesky分解
- 左辺がToeplitz行列の場合：
Levinsonアルゴリズム

この場合の
解き方は？

Levinson-Durbin-Itakuraアルゴリズム (2/6)

■ 右辺を左辺に移項

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_{P-1} \\ v_1 & v_0 & \cdots & v_{P-2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{P-1} & v_{P-2} & \cdots & v_0 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_P \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_P \end{bmatrix} \quad P \times P \text{行列}$$

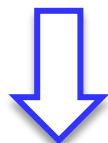
$$\sigma_P^2 = v_0 - \sum_{p=1}^P v_p a_p$$

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_P \\ v_1 & v_0 & \cdots & v_{P-1} \\ \vdots & \vdots & \ddots & \vdots \\ v_P & v_{P-1} & \cdots & v_0 \end{bmatrix} \begin{bmatrix} 1 \\ -a_1 \\ \vdots \\ -a_P \end{bmatrix} = \begin{bmatrix} \sigma_P^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (P+1) \times (P+1) \text{行列}$$

Levinson-Durbin-Itakuraアルゴリズム (3/6)

- P次の(最適な)予測係数から、(P+1)次の(最適な)予測係数を再帰的に解けないか？

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_P \\ v_1 & v_0 & \cdots & v_{P-1} \\ \vdots & \vdots & \ddots & \vdots \\ v_P & v_{P-1} & \cdots & v_0 \end{bmatrix} \begin{bmatrix} 1 \\ -a_1^{(P)} \\ \vdots \\ -a_P^{(P)} \end{bmatrix} = \begin{bmatrix} \sigma_P^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \cdots (*)$$



$$\begin{bmatrix} v_0 & v_1 & \cdots & v_{P+1} \\ v_1 & v_0 & \cdots & v_P \\ \vdots & \vdots & \ddots & \vdots \\ v_{P+1} & v_P & \cdots & v_0 \end{bmatrix} \begin{bmatrix} 1 \\ -a_1^{(P+1)} \\ \vdots \\ -a_P^{(P+1)} \end{bmatrix} = \begin{bmatrix} \sigma_{P+1}^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\hat{a}_1^{(P)}, \dots, \hat{a}_P^{(P)}$$

関係は？

$$\hat{a}_1^{(P+1)}, \dots, \hat{a}_P^{(P+1)}$$

Levinson-Durbin-Itakuraアルゴリズム (4/6)

■式(*)を変形

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_P \\ v_1 & v_0 & \cdots & v_{P-1} \\ \vdots & \vdots & \ddots & \vdots \\ v_P & v_{P-1} & \cdots & v_0 \end{bmatrix} \begin{bmatrix} 1 \\ -a_1^{(P)} \\ \vdots \\ -a_P^{(P)} \end{bmatrix} = \begin{bmatrix} \sigma_P^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

(P+1) × (P+1)行列

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_P & v_{P+1} \\ v_1 & v_0 & \cdots & v_{P-1} & v_P \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_P & v_{P-1} & \cdots & v_0 & v_1 \end{bmatrix} \begin{bmatrix} 1 \\ -a_1^{(P)} \\ \vdots \\ -a_P^{(P)} \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma_P^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_P & v_{P+1} \\ v_1 & v_0 & \cdots & v_{P-1} & v_P \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_P & v_{P-1} & \cdots & v_0 & v_1 \\ v_{P+1} & v_P & \cdots & v_1 & v_0 \end{bmatrix} \begin{bmatrix} 1 \\ -a_1^{(P)} \\ \vdots \\ -a_P^{(P)} \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma_P^2 \\ 0 \\ \vdots \\ 0 \\ w_P \end{bmatrix}$$

(P+2) × (P+2)行列

$$w_P \rightarrow v_{P+1} - \sum_{p=1}^P a_p v_{P-p+1}$$

Levinson-Durbin-Itakuraアルゴリズム (5/6)

■ 左辺は対称行列より

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_P & v_{P+1} \\ v_1 & v_0 & \cdots & v_{P-1} & v_P \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_P & v_{P-1} & \cdots & v_0 & v_1 \\ v_{P+1} & v_P & \cdots & v_1 & v_0 \end{bmatrix} \begin{bmatrix} 1 \\ -a_1^{(P)} \\ \vdots \\ -a_P^{(P)} \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma_P^2 \\ 0 \\ \vdots \\ 0 \\ w_P \end{bmatrix} \quad \dots \textcircled{1}$$

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_P & v_{P+1} \\ v_1 & v_0 & \cdots & v_{P-1} & v_P \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_P & v_{P-1} & \cdots & v_0 & v_1 \\ v_{P+1} & v_P & \cdots & v_1 & v_0 \end{bmatrix} \begin{bmatrix} 0 \\ -a_P^{(P)} \\ \vdots \\ -a_1^{(P)} \\ 1 \end{bmatrix} = \begin{bmatrix} w_P \\ 0 \\ \vdots \\ 0 \\ \sigma_P^2 \end{bmatrix} \quad \dots \textcircled{2}$$

■ ① - k_P × ② (k_P は任意の係数)

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_P & v_{P+1} \\ v_1 & v_0 & \cdots & v_{P-1} & v_P \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_P & v_{P-1} & \cdots & v_0 & v_1 \\ v_{P+1} & v_P & \cdots & v_1 & v_0 \end{bmatrix} \begin{bmatrix} 1 - 0 \\ -a_1^{(P)} + k_P a_P^{(P)} \\ \vdots \\ -a_P^{(P)} + k_P a_1^{(P)} \\ 0 - k_P \end{bmatrix} = \begin{bmatrix} \sigma_P^2 - k_P w_P \\ 0 \\ \vdots \\ 0 \\ w_P - k_P \sigma_P^2 \end{bmatrix}$$

Levinson-Durbin-Itakura アルゴリズム (6/6)

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_P & v_{P+1} \\ v_1 & v_0 & \cdots & v_{P-1} & v_P \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ v_P & v_{P-1} & \cdots & v_0 & v_1 \\ v_{P+1} & v_P & \cdots & v_1 & v_0 \end{bmatrix} \begin{bmatrix} 1 - 0 \\ -a_1^{(P)} + k_P a_P^{(P)} \\ \vdots \\ -a_P^{(P)} + k_P a_1^{(P)} \\ 0 - k_P \end{bmatrix} = \begin{bmatrix} \sigma_P^2 - k_P w_P \\ 0 \\ \vdots \\ 0 \\ w_P - k_P \sigma_P^2 \end{bmatrix}$$

- k_P は任意なので $w_P - k_P \sigma_P^2 = 0$ となるように k_P を選ぶと上式は下記の形になる

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_{P+1} \\ v_1 & v_0 & \cdots & v_P \\ \vdots & \vdots & \ddots & \vdots \\ v_{P+1} & v_P & \cdots & v_0 \end{bmatrix} \begin{bmatrix} 1 \\ -a_1^{(P+1)} \\ \vdots \\ -a_P^{(P+1)} \end{bmatrix} = \begin{bmatrix} \sigma_{P+1}^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- よって以下の再帰式を得る

$$k_P = \frac{w_P}{\sigma_P^2} \longrightarrow \text{偏自己相関(Partial Correlation; PARCOR)係数}$$

[Itakura1969]

$$\sigma_{P+1}^2 = \sigma_P^2 - k_P w_P$$

$$a_p^{(P+1)} = a_p^{(P)} - k_P a_{P-p+1}^{(P)}$$

なお、明らかに $a_1^{(1)} = \frac{v_1}{v_0}$

予測誤差

- 最適予測係数を $\hat{a}_1, \dots, \hat{a}_P$ とすると、「予測の誤差」は...

$$\epsilon_t = s_t - \sum_{p=1}^P \hat{a}_p s_{t-p}$$

$$\begin{bmatrix} 1 & & & & & & & & & & & & 0 \\ -\hat{a}_1 & 1 & & & & & & & & & & & \\ \vdots & \ddots & \ddots & & & & & & & & & & \\ -\hat{a}_P & & \ddots & \ddots & \ddots & & & & & & & & \\ & \ddots & & \ddots & \ddots & \ddots & & & & & & & \\ 0 & & & -\hat{a}_P & \cdots & -\hat{a}_1 & 1 & & & & & & \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_T \end{bmatrix} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_T \end{bmatrix}$$

- 予測誤差 $\{\epsilon_t\}_{1 \leq t \leq T}$ と予測係数 $\{\hat{a}_p\}_{1 \leq p \leq P}$ から元信号を復元可能

$$s_1 = \epsilon_1$$

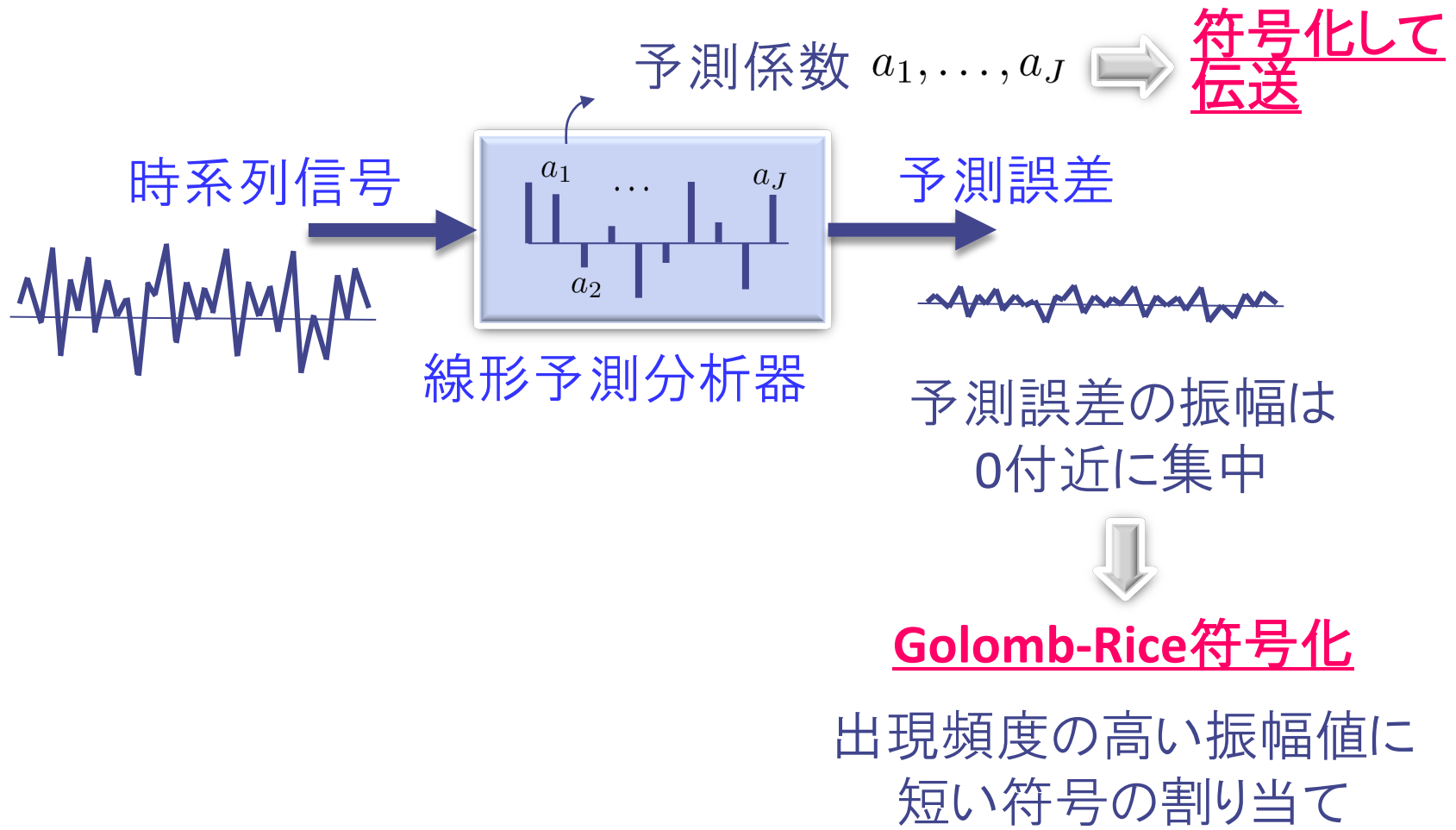
$$s_2 = \epsilon_2 + \hat{a}_1 s_1$$

$$s_3 = \epsilon_3 + \hat{a}_1 s_2 + \hat{a}_2 s_1$$

$$\vdots$$

線形予測符号化 (Linear Predictive Coding)

- 時系列信号の可逆圧縮符号化の標準的な方式



線形予測分析

■ 3つの観点から解説

■ 「予測誤差」を最小化する観点

■ 最小二乗誤差推定

■ 線形系としての観点

■ 自己回帰系 (AutoRegressive system)

■ 音声の生成過程モデル

■ 最尤推定

■ 白色化

■ スペクトルマッチングとしての観点

■ 最尤スペクトル推定

■ 板倉斎藤距離

線形予測分析

■ 3つの観点から解説

■ 「予測誤差」を最小化する観点

■ 最小二乗誤差推定

■ 線形系としての観点

■ 自己回帰系 (AutoRegressive system)

■ 音声の生成過程モデル

■ 最尤推定

■ 白色化

■ スペクトルマッチングとしての観点

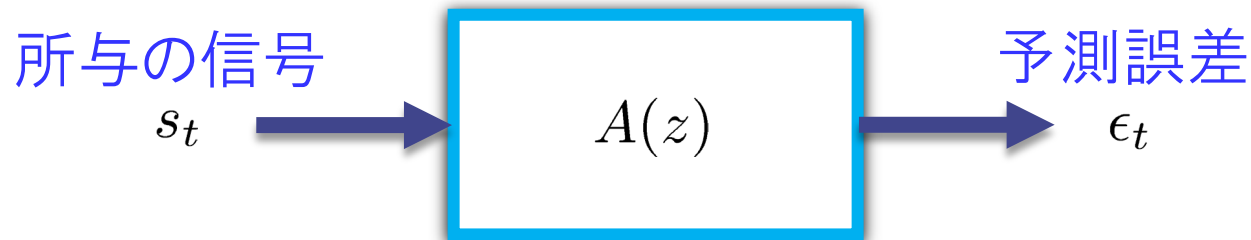
■ 最尤スペクトル推定

■ 板倉斎藤距離

線形系としての解釈

- 所与の信号から予測誤差を出力する線形システム

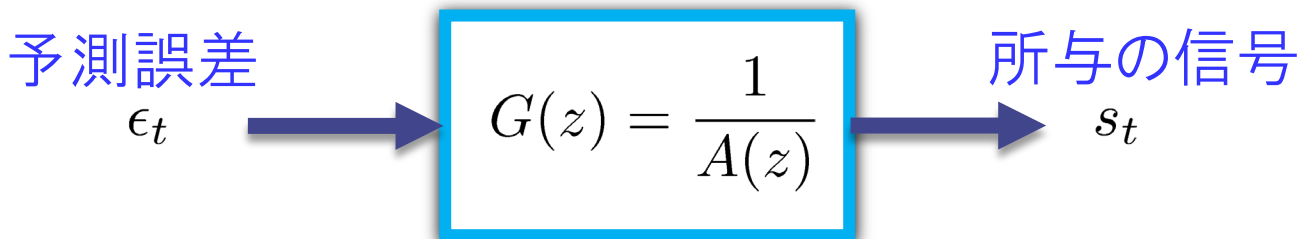
$$\epsilon_t = s_t - \sum_{p=1}^P a_p s_{t-p}$$



$$A(z) = 1 - a_1 z^{-1} - a_2 z^{-2} - \dots - a_P z^{-P}$$

移動平均システム(全零モデル)

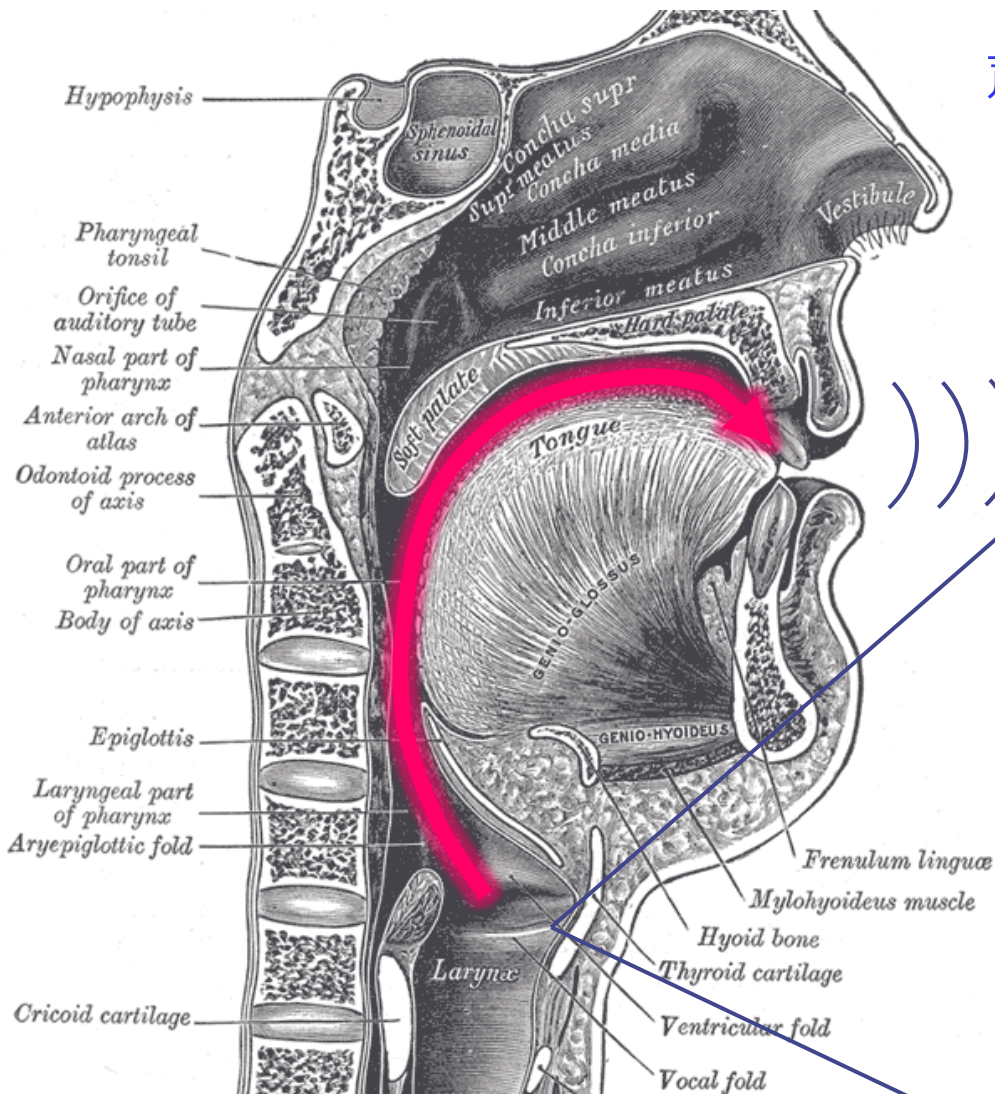
- 予測誤差を入力として所与の信号を出力する線形システムは？



自己回帰システム(全極モデル)

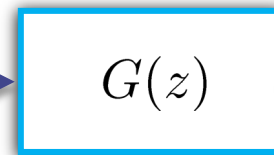
音声生成過程のモデルとして

- 声帯振動が声道で共振して音声波形となって口から発せられる



声帯振動

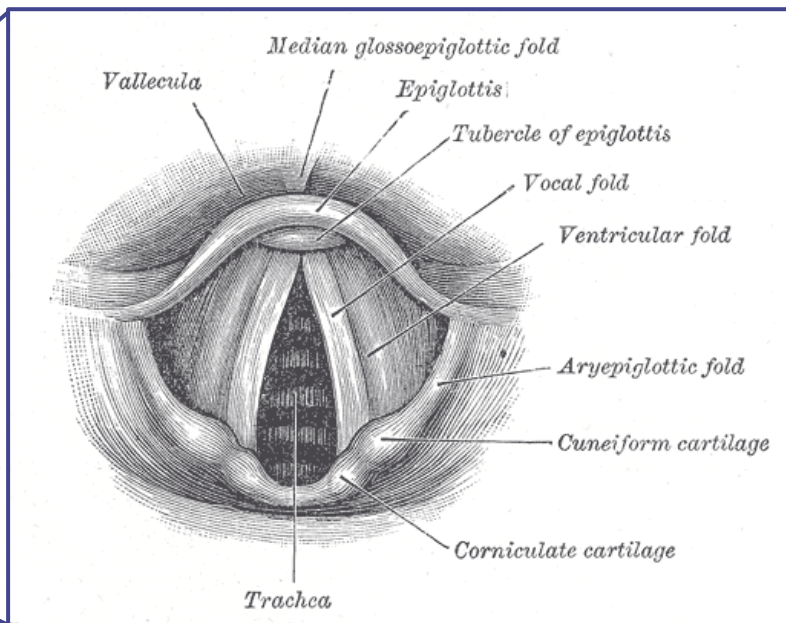
ϵ_t



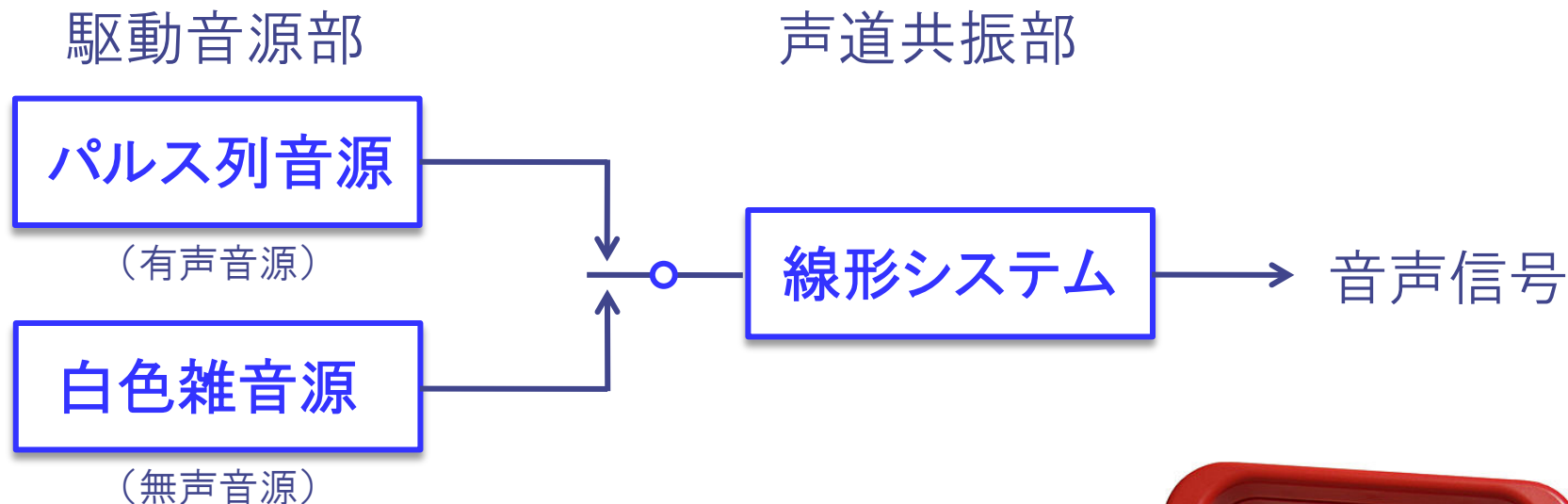
音声波形

S_t

自己回帰システムにより
声道特性を表現した場合
の音声生成過程モデル



音声生成の線形モデル



■ “Speak & Spell”

- LPCに基づく音声合成LSIを搭載
- 米国のTexas Instruments社開発
- 1978年発売



“Speak & Spell”のコマーシャル



統計モデルによる音声生成過程の表現

■ 声帯振動に関する仮定

- Gauss性 $\dots \epsilon = (\epsilon_1, \dots, \epsilon_T)^T \sim \mathcal{N}(\mu, \Sigma)$

- 定常性 $\dots \mu = \mathbf{0}, \Sigma : \text{Toeplitz行列}$

- 白色性 $\dots \Sigma = \sigma^2 I$

$$\left. \begin{array}{l} \epsilon_t \sim \mathcal{N}(0, \sigma^2) \\ (t = 1, \dots, T) \end{array} \right\}$$

■ 声道特性に関する仮定

- 自己回帰システム(全極モデル)

$$\underbrace{\begin{bmatrix} 1 & & & & & 0 \\ -a_1 & 1 & & & & \\ \vdots & \ddots & \ddots & & & \\ -a_P & & \ddots & \ddots & & \\ 0 & \ddots & & \ddots & \ddots & \\ 0 & & -a_P & \cdots & -a_1 & 1 \end{bmatrix}}_{\Psi} \underbrace{\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_T \end{bmatrix}}_{\mathbf{s}} = \underbrace{\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_T \end{bmatrix}}_{\boldsymbol{\epsilon}}$$

最尤推定

■ 今までの仮定をまとめると・・・

- $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$

- $\Psi \mathbf{s} = \epsilon$

未知パラメータは $\{a_p\}_{1 \leq p \leq P}, \sigma^2$

観測されるのは $\{s_t\}_{1 \leq t \leq T}$

■ 観測信号 \mathbf{s} の確率密度関数(尤度関数という)

- $\mathbf{s} = \Psi^{-1} \epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 \Psi^{-1} (\Psi^{-1})^T)$

- 対数尤度は

$$-\frac{T}{2} \log(2\pi) - \frac{1}{2} \log |\sigma^2 \Psi^{-1} (\Psi^{-1})^T| - \frac{1}{2\sigma^2} \mathbf{s}^T \Psi^T \Psi \mathbf{s}$$

\downarrow $\|\epsilon\|_2^2 =$

- logdet項: $\frac{\log |\sigma^2 \Psi^{-1} (\Psi^{-1})^T|}{= \log |\sigma^2 \mathbf{I}| - 2 \log |\Psi|}$
 $= T \log \sigma^2 \quad \because |\Psi| = 1$

$$\sum_t \left(s_t - \sum_{p=1}^P a_p s_{t-p} \right)^2$$

白色化効果

- 以上の統計モデルでは ϵ について白色性を仮定していたので、先の最尤推定では ϵ ができるだけ白色になるように $\{a_p\}_{1 \leq p \leq P}$ を決めようとしていたことになる
- このことをよりイメージしやすくするため、以上のモデルを周波数領域で定式化してみよう

ここら辺で一息

+

線形予測分析

■ 3つの観点から解説

■ 「予測誤差」を最小化する観点

■ 最小二乗誤差推定

■ 線形系としての観点

■ 自己回帰系 (AutoRegressive system)

■ 音声の生成過程モデル

■ 最尤推定

■ 白色化

■ スペクトルマッチングとしての観点

■ 最尤スペクトル推定

■ 板倉斎藤距離

線形予測分析

■ 3つの観点から解説

- 「予測誤差」を最小化する観点
 - 最小二乗誤差推定
- 線形系としての観点
 - 自己回帰系 (AutoRegressive system)
 - 音声の生成過程モデル
 - 最尤推定
 - 白色化
- スペクトルマッチングとしての観点
 - 最尤スペクトル推定
 - 板倉斎藤距離

周波数領域での定式化

■ 時間領域では...

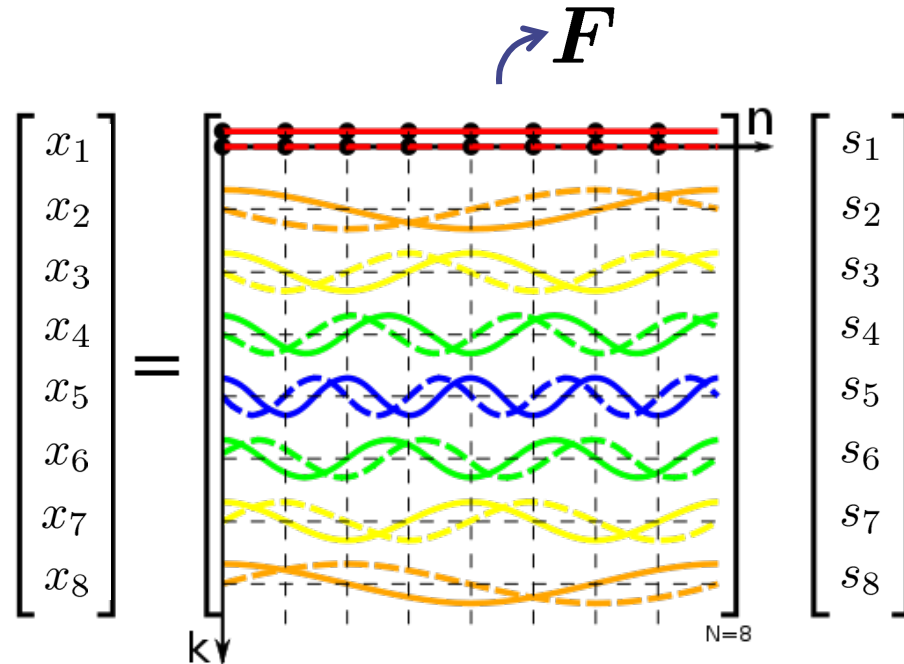
$$\left. \begin{array}{l} \blacksquare \Psi^{-1} \epsilon = s \\ \blacksquare \epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}) \end{array} \right\} \longrightarrow s \sim \mathcal{N}(\mathbf{0}, \sigma^2 \Psi^{-1} (\Psi^{-1})^T)$$

■ 周波数領域 (Fourier変換領域) では...

■ $x = F s$ (F は離散Fourier変換行列) の確率密度関数は？

$$x \sim \mathcal{N}_C(\mathbf{0}, \sigma^2 \mathbf{F} \Psi^{-1} (\Psi^{-1})^T \mathbf{F}^H)$$

$$\begin{aligned} & \mathbf{F} \Psi^{-1} (\Psi^{-1})^T \mathbf{F}^H \\ &= (\mathbf{F} \underline{\Psi^T \Psi} \mathbf{F}^H)^{-1} \end{aligned}$$



$\Psi^T \Psi$ について

- Ψ に関してここでは以下の巡回行列型を仮定

$$\Psi = \begin{bmatrix} 1 & & & -a_P & \cdots & -a_1 \\ -a_1 & 1 & & & & \vdots \\ \vdots & \ddots & \ddots & & & -a_P \\ -a_P & & \ddots & \ddots & & \\ & \ddots & & \ddots & \ddots & \\ 0 & & -a_P & \cdots & -a_1 & 1 \end{bmatrix}$$

- よって $\Psi^T \Psi$ も巡回行列

→ $\Psi^T \Psi$ は離散Fourier変換行列 F によって対角化される

$F \Psi^T \Psi F^H$: 対角行列

$$\sigma^2 (F \Psi^T \Psi F^H)^{-1} = \text{diag}(\lambda_1, \dots, \lambda_T)$$

$$\lambda_t = \frac{\sigma^2}{|A(e^{2\pi j(t-1)/T})|^2}$$

$$A(z) = 1 - a_1 z^{-1} - \cdots - a_P z^{-P}$$

周波数成分の確率密度関数

- 以上をまとめると・・・

$$\boldsymbol{x} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \text{diag}(\lambda_1, \dots, \lambda_T))$$

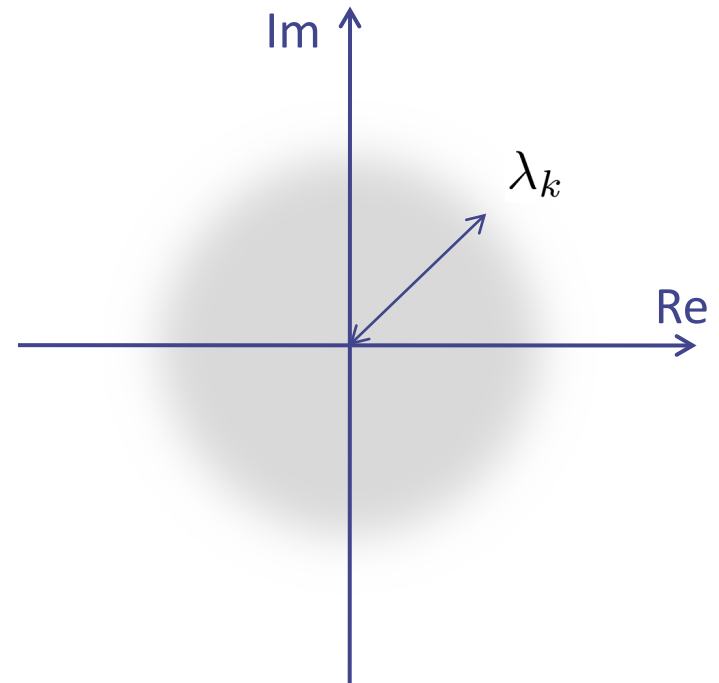
$$\lambda_t = \frac{\sigma^2}{|A(e^{2\pi j(t-1)/T})|^2}$$

$$A(z) = 1 - a_1 z^{-1} - \dots - a_P z^{-P}$$

- 周波数 $2\pi(k-1)/T$ の成分 x_k

- 他の周波数の成分と独立
- 分散が λ_k の複素正規分布に従う

$$x_k \sim \mathcal{N}_{\mathbb{C}}(0, \lambda_k)$$

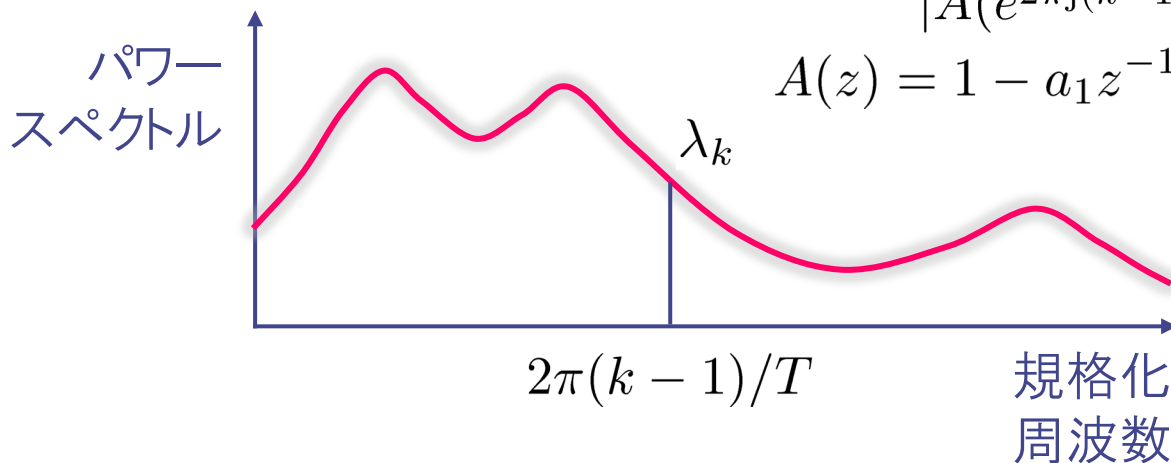


「スペクトルマッチング」としての見方

- 周波数成分 $\{x_k\}_{1 \leq k \leq T}$ が与えられた下での対数尤度

$$\mathcal{L}(\lambda_1, \dots, \lambda_T) = - \sum_k \left(\log \pi \lambda_k + \frac{|x_k|^2}{\lambda_k} \right)$$

$$\lambda_k = \frac{\sigma^2}{|A(e^{2\pi j(k-1)/T})|^2}$$
$$A(z) = 1 - a_1 z^{-1} - \dots - a_P z^{-P}$$



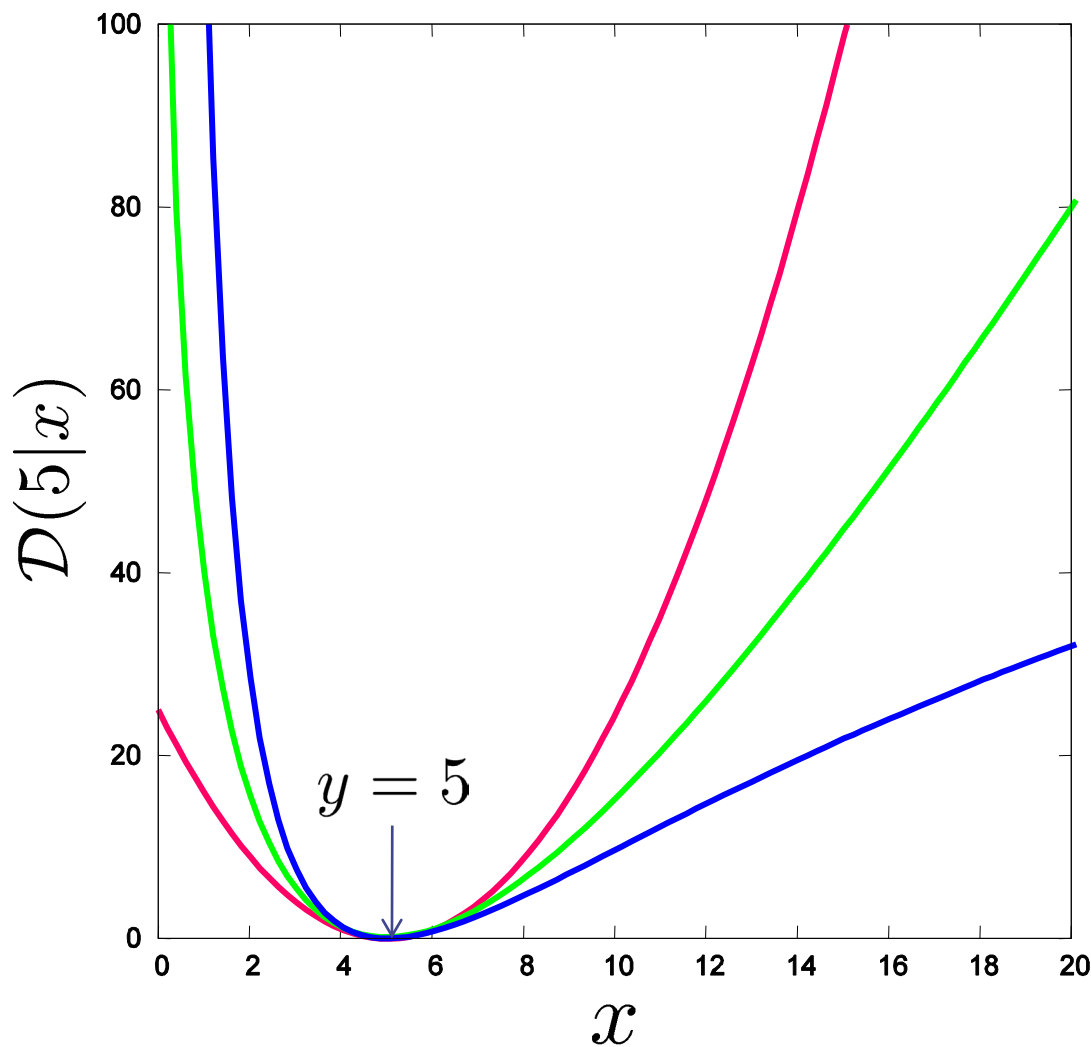
- 上記の対数尤度は、定数項を除けば以下と等しい

$$\mathcal{L}(\lambda_1, \dots, \lambda_T) \stackrel{c}{=} - \sum_k \left(\frac{|x_k|^2}{\lambda_k} - \log \frac{|x_k|^2}{\lambda_k} - 1 \right)$$

板倉斎藤距離

板倉齋藤距離

■他の擬距離尺度との比較



— 二乗誤差

$$D(x|y) = (y - x)^2$$

— Iダイバージェンス

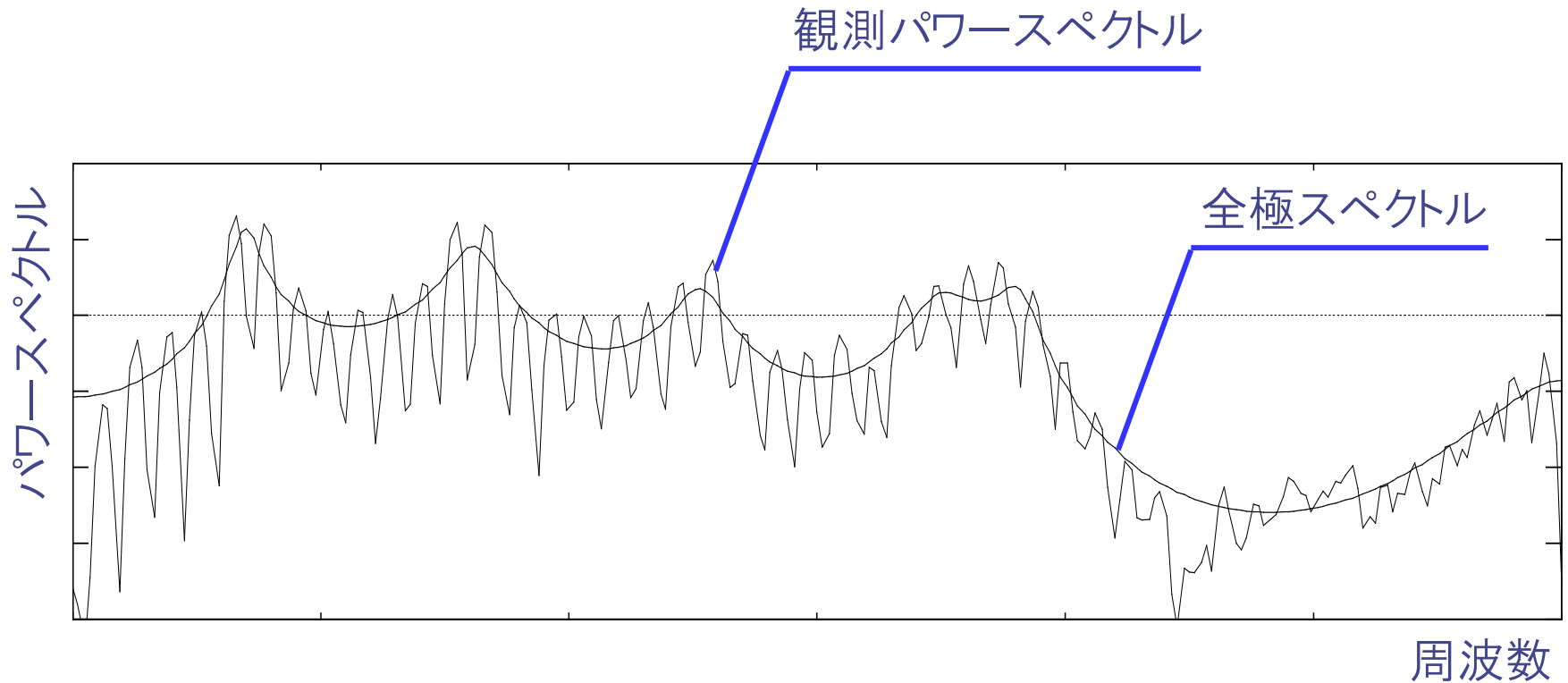
$$D(x|y) = y \log \frac{y}{x} - (y - x)$$

— 板倉齋藤距離

$$D(x|y) = \frac{y}{x} - \log \frac{y}{x} - 1$$

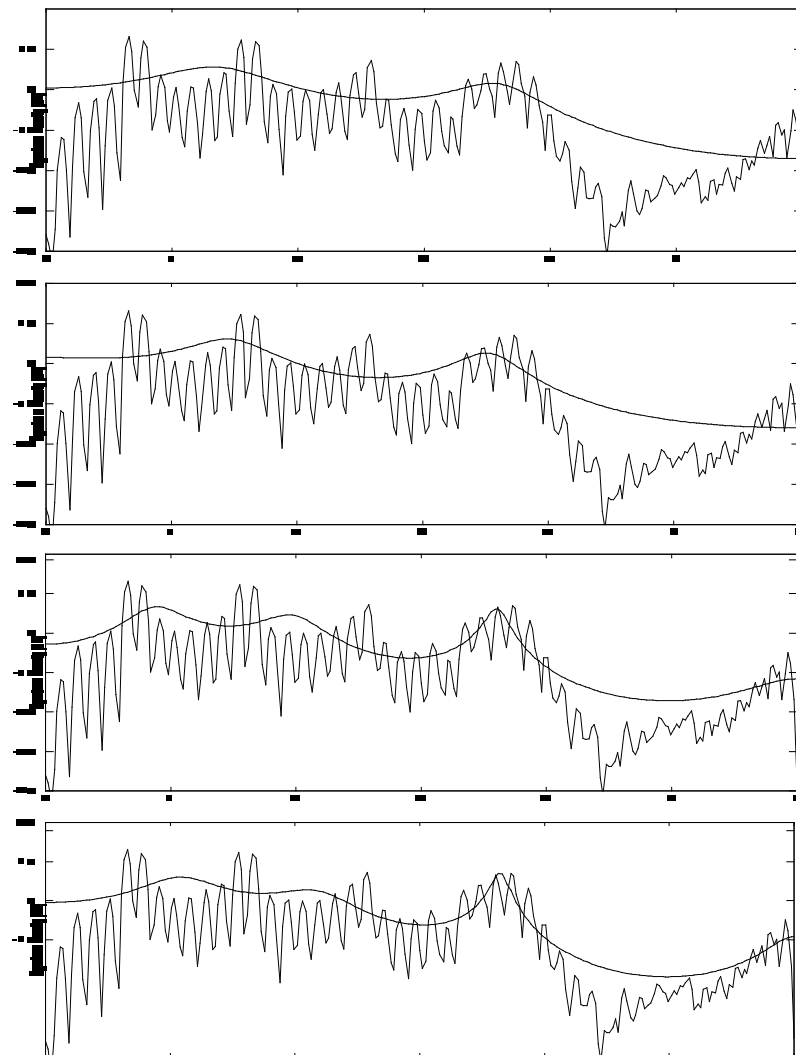
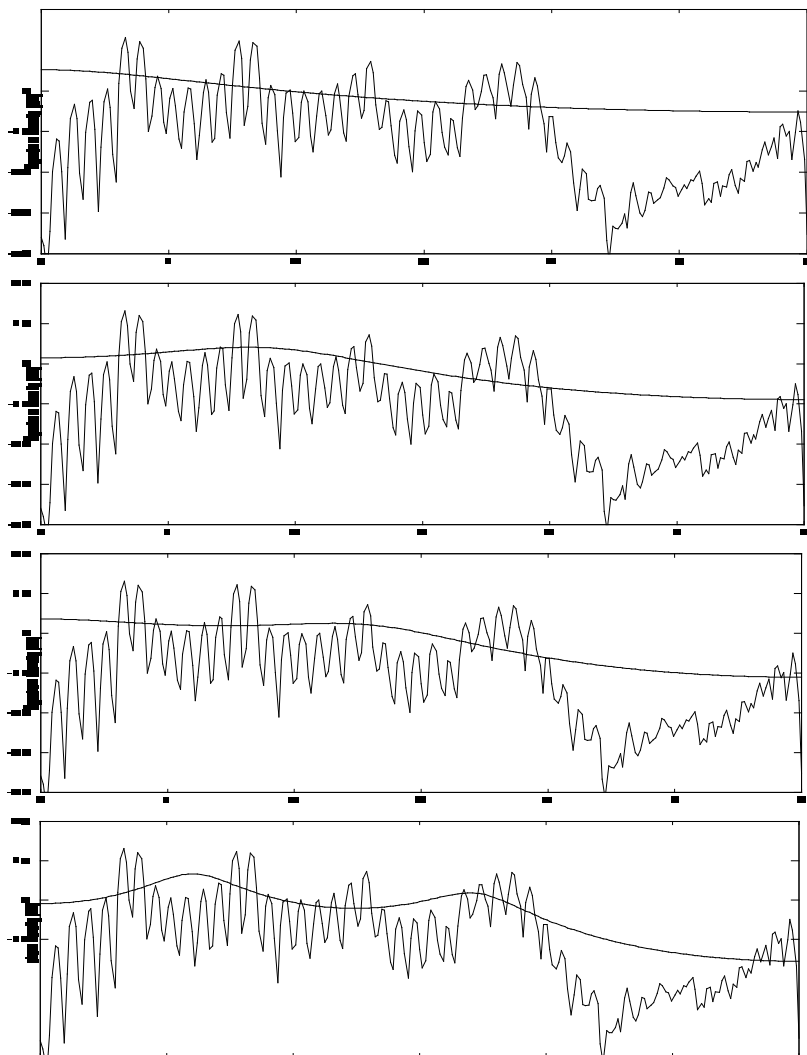
線形予測分析は「スペクトル包絡」の推定に相当

- 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



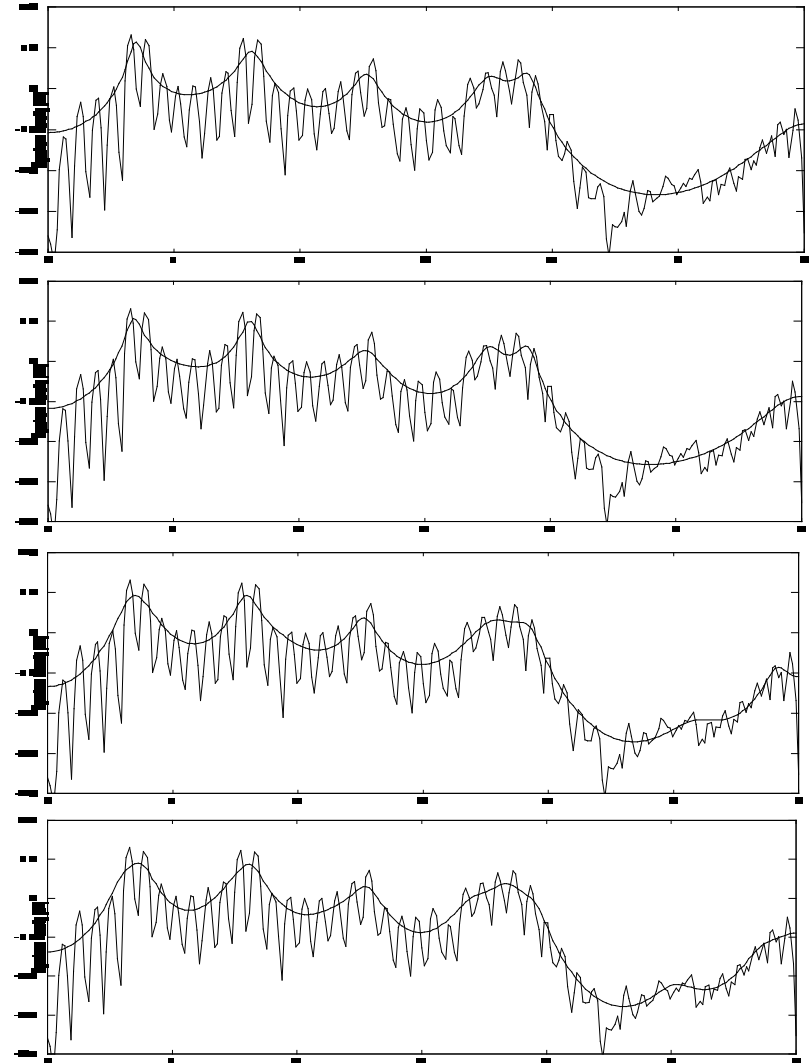
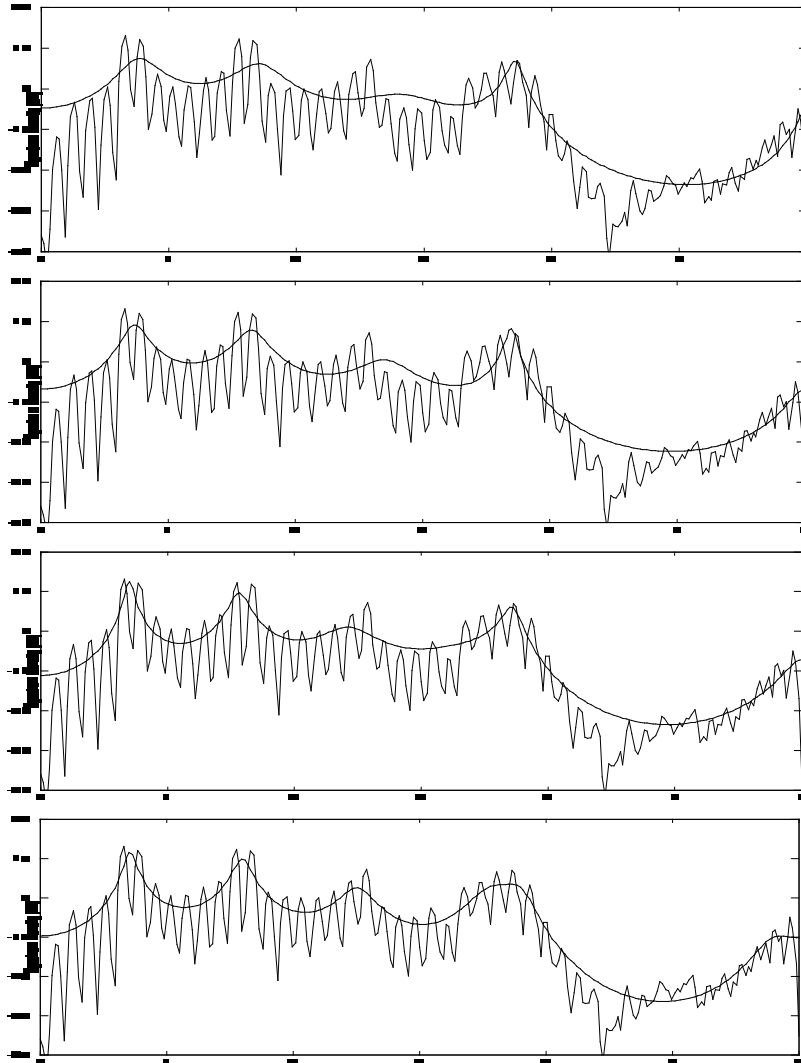
LPCによる音声スペクトル推定の例

■ 1~8次



LPCによる音声スペクトル推定の例

■ 9, 10, 12, 14, 16, 18, 22, 26次



線形予測分析

■ 3つの観点から解説

- 「予測誤差」を最小化する観点
 - 最小二乗誤差推定
- 線形系としての観点
 - 自己回帰系 (AutoRegressive system)
 - 音声の生成過程モデル
 - 最尤推定
 - 白色化
- スペクトルマッチングとしての観点
 - 最尤スペクトル推定
 - 板倉斎藤距離

レポート課題対象論文

- P. Kabal and R. P. Ramachandran, "The Computation of Line Spectral Frequencies Using Chebyshev Polynomials," IEEE Transactions on Acoustics, Speech, Signal Processing, vol. 34, no. 6, pp. 1419–1426, Dec. 1986.