マルチカーネル線形予測モデルによる音声分析*

○亀岡弘和，大石康智，持橋大地，ルルージョナトン

日本電信電話 (株) NTT コミュニケーション科学基礎研究所

1 序 論

音声信号から声道伝達特性と声帯駆動信号を推定するための古典的な手法として線形予測法が有名である [1]。線形予測法は，声帯による駆動音源信号が白色ガウス過程に従うという仮定に基づき，全極モデルによって表された声道伝達特性を推定する手法である。しかし，実際の音声における有声音の声帯振動波形は局所的に周期的であり，その周期が短い (基本周波数が高い) ほど白色性の仮定からの乖離が顕著となる。このため，従来の線形予測法では，声道伝達特性と声帯駆動信号の情報を必ずしも適切に分解することができず，声道伝達特性の推定結果が基本周波数の影響を陽に受けてしまうという問題があった。線形予測法における以上の基本周波数の影響を回避する目的で，これまで多数の改良アプローチが提案されている。その代表的なアプローチとして，声帯駆動信号が従う分布をラプラス分布と仮定したもの [2]，基本周波数推定を予め行い，抽出した調波成分のスペクトルピークだけに対して全極モデルをフィッティングするもの [3]，声帯駆動信号を隠れマルコフモデルにより確率モデル化したもの [4]，推定声門閉鎖区間を利用して全極モデルを推定するもの [5]，などが挙げられる。

本稿では，上述の改良アプローチと同様に従来の線形予測モデルによる声道スペクトル推定において問題となっていた基本周波数の影響を解消することを目的とするだけでなく，有声無声判定と基本周波数推定も同時に行うことができる新しい音声分析法を提案する。

2 マルチカーネル線形予測法

2.1 線形予測法 [1]

ここではまず従来の線形予測法を概説する。線形予測法では，短時間信号 $x_1, \cdots, x_I$ を，P 次の自己回帰過程

$$x_i = \sum_{p=1}^{P} a_p x_{i-p} + \epsilon_i \tag{1}$$

からの標本値系列と仮定する。$\boldsymbol{a} := (a_1, \cdots, a_P)^{\mathrm{T}}$ は全極モデルのパラメータであり，線形予測法においてはこれを予測係数と呼ぶ。このモデルは，ϵ_i を入力，$A(z) = 1/(1 - a_1 z^{-1} \cdots - a_P z^{-P})$ を伝達関数，x_i を出力とする線形系と見ることができ，声帯駆動信号に対応する入力 $\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_I)^{\mathrm{T}}$ を，定常な白色ガウス性雑音

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

と仮定する。$\boldsymbol{x} = (x_1, \cdots, x_I)^{\mathrm{T}}$,

$$\boldsymbol{\Psi} := \begin{bmatrix} 1 & & & & & 0 \\ -a_1 & \ddots & & & & \\ \vdots & \ddots & \ddots & & & \\ -a_P & & \ddots & \ddots & & \\ & \ddots & & \ddots & \ddots & \\ 0 & & -a_P & \cdots & -a_1 & 1 \end{bmatrix} \tag{3}$$

と置くと式 (1) は

$$\boldsymbol{\Psi x} = \boldsymbol{\epsilon} \tag{4}$$

と行列表現でき，式 (2) より

$$\boldsymbol{\Psi x} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}) \tag{5}$$

が言える。よって，

$$\boldsymbol{x} \sim \mathcal{N}\big(\boldsymbol{0}, \sigma^2 (\boldsymbol{\Psi}^{\mathrm{T}} \boldsymbol{\Psi})^{-1}\big) \tag{6}$$

となるため，$|\boldsymbol{\Psi}| = 1$ であることに注意すると，$\boldsymbol{x}$ の確率密度関数の対数は具体的に

$$\begin{aligned} \log p(\boldsymbol{x}; \boldsymbol{a}) &= -\frac{I}{2} \log 2\pi - \frac{1}{2\sigma^2} \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\Psi}^{\mathrm{T}} \boldsymbol{\Psi x} \\ &= -\frac{I}{2} \log 2\pi - \frac{1}{2\sigma^2} (\boldsymbol{x} - \boldsymbol{X a})^{\mathrm{T}} (\boldsymbol{x} - \boldsymbol{X a}) \end{aligned} \tag{7}$$

と書ける。ただし，$\boldsymbol{a} = (a_1, \cdots, a_P)^{\mathrm{T}}$,

$$\boldsymbol{X} := \begin{bmatrix} 0 & \cdots & 0 \\ x_1 & \ddots & \vdots \\ \vdots & \ddots & 0 \\ \vdots & & x_1 \\ \vdots & & \vdots \\ x_{I-1} & \cdots & x_{I-P} \end{bmatrix} \tag{8}$$

である。よって，最尤な予測係数 $\boldsymbol{a}$ は，正規方程式

$$\boldsymbol{X}^{\mathrm{T}} \boldsymbol{X a} = \boldsymbol{X}^{\mathrm{T}} \boldsymbol{x} \tag{9}$$

を解くことによって得られる。

2.2 ガウシアンプロセスによる声帯駆動のモデル化

以上のように線形予測法による音声生成モデルでは声帯駆動信号 $\{\epsilon_i\}_{i=1}^{I}$ を白色ガウス雑音と仮定したが，ここでは $\{\epsilon_i\}_{i=1}^{I}$ と標本時刻 $\{t_i\}_{i=1}^{I}$ との関係を表す回帰式をモデル化する見地から線形予測法を捉え直す。まず，

$$\phi_h(t) := \begin{cases} 1 & (t = c_h) \\ 0 & (t \neq c_h) \end{cases} \tag{10}$$

---

*Speech analysis with multi-kernel linear prediction. by Hirokazu Kameoka, Yasunori Ohishi, Daichi Mochihashi, Jonathan Le Roux (NTT Communication Science Laboratories)

と定義される基底関数を用いて表される線形回帰モデル

$$\epsilon = f(t) + \eta \tag{11}$$

$$f(t) = \sum_{h=1}^{H} \phi_h(t) w_h = \boldsymbol{\phi}(t)^{\mathrm{T}} \boldsymbol{w} \tag{12}$$

を考える。よって，$\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_I)^{\mathrm{T}}$ は，計画行列 $\boldsymbol{\Phi} := (\boldsymbol{\phi}(t_1), \cdots, \boldsymbol{\phi}(t_I))^{\mathrm{T}}$ を用いて

$$\boldsymbol{\epsilon} = \boldsymbol{\Phi} \boldsymbol{w} + \boldsymbol{\eta} \tag{13}$$

と表される。ただし，$\boldsymbol{\eta} := (\eta_1, \cdots, \eta_I)^{\mathrm{T}}$ である。ここで，重み $\boldsymbol{w}$ の事前確率を

$$\boldsymbol{w} \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_w^2 \boldsymbol{I}\right) \tag{14}$$

と仮定し，さらに $\boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{0}, \sigma_\eta^2 \boldsymbol{I})$ とすると，

$$\boldsymbol{\epsilon} \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_w^2 \boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}} + \sigma_\eta^2 \boldsymbol{I}\right) \tag{15}$$

が言える。H を十分大きくとり，$\{t_1, \cdots, t_I\} \subseteq \{c_1, \cdots, c_H\}$ となる状況を想定すると，

$$\boldsymbol{\phi}(t_i)^{\mathrm{T}} \boldsymbol{\phi}(t_j) = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases} \tag{16}$$

となるから，$\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}} = \boldsymbol{I}$ と書け，$\sigma^2 := \sigma_w^2 + \sigma_\eta^2$ と置くと，

$$\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}) \tag{17}$$

のように式 (2) の仮定に帰着する。以上のように，正規分布する結合係数による基底関数の線形和で表される回帰モデルをガウシアンプロセス [6] といい，その確率分布はカーネル関数

$$k(t, t') := \boldsymbol{\phi}(t)^{\mathrm{T}} \boldsymbol{\phi}(t') \tag{18}$$

をもとに構成されるグラム行列 $\boldsymbol{K} = (K_{i,j})_{I \times I} = \boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}}$

$$K_{i,j} = k(t_i, t_j) \tag{19}$$

によって規定される。本節では，従来の線形予測法における $\boldsymbol{\epsilon}$ の仮定が，基底関数を式 (10) と定義した場合のガウシアンプロセスによる $\boldsymbol{\epsilon}$ のモデル化に相当していることを見た。次節では，別のカーネル関数を導入して周期性をもつ雑音モデルを表現できることを示す。

2.3 周期カーネルの導入

先に述べたとおり，有声音の場合には声帯駆動信号は周期性をもつことが多い。本節では，その周期が既知であると仮定して議論する。

2.2 節で見たとおり，ガウシアンプロセスの確率分布はカーネル関数によって決まる。カーネル関数は，必ずしも基底関数の定義を介して導き出す必要はなく，半正定値性を満たすものであれば自由に定義できる。例えば，

$$k(t, t') := \exp\left(-2\sin^2\left(\pi \frac{t - t'}{T}\right) \Big/ \ell^2\right) \tag{20}$$

は半正定値性を満たすカーネル関数の一つであるが，これは，基底関数として周期 T の周期関数を仮定していることに相当する [6]。本稿では，調波成分が等しいパワーをもつ駆動信号を仮定して基底関数を

$$\phi_h(t) := \sum_{n=1}^{N} \sin\left(2\pi n \frac{t - c_h}{T}\right) \tag{21}$$

と置き，そのもとで式 (18) に基づいて算出されるカーネル関数を用いる。このような周期カーネルを用いることで，ガウシアンプロセスにより周期性をもつランダム雑音を表現することが可能となる。

2.4 マルチカーネル線形予測モデル

式 (13) を式 (4) に代入することによって得られる

$$\boldsymbol{\Psi}\boldsymbol{x} = \boldsymbol{\Phi}\boldsymbol{w} + \boldsymbol{\eta} \tag{22}$$

をカーネル線形予測モデルと呼ぶ。ここで式 (15) より，

$$\boldsymbol{\Psi}\boldsymbol{x} \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_w^2 \boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}} + \sigma_\eta^2 \boldsymbol{I}\right) \tag{23}$$

が言える。以後，$\boldsymbol{K} := \boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}}$ と置く。$\boldsymbol{\Phi}$ を，式 (21) で与えられる基底関数によって構成される計画行列とすると，このカーネル線形予測モデルは基本周期 T の音声信号を表現することができる。

さて，これまでは，声帯駆動信号の周期 T は既知と見なしていたが，音声信号の基本周期は実際には未知である。このため，予測係数だけでなく観測信号データからカーネル自体の推定も必要となる。そこで，ここではマルチカーネル学習の考え方に基づき，複数の周期カーネルの線形結合によって与えられるカーネル関数

$$k(t, t') = \sum_{m=1}^{M} \theta_m k_m(t, t') \tag{24}$$

を導入し，周期 T_m の周期カーネル $k_m(t, t')$ の優勢度 θ_m を推定すべき未知パラメータと見なす。ただし，

$$\sum_{m=1}^{M} \theta_m = 1 \tag{25}$$

とする。従って，$\{\theta_m\}_{m=1}^{M}$ の推定を通して音声信号に含まれる優勢な基本周期を知ることができる。$k_m(t_i, t_j)$ を i 行 j 列の要素にもつ行列を $\boldsymbol{K}_m$ とすると，

$$\boldsymbol{K} = \sum_{m=1}^{M} \theta_m \boldsymbol{K}_m \tag{26}$$

となり，これと式 (23) より，

$$\boldsymbol{x} \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_w^2 \sum_{m=1}^{M} \theta_m \boldsymbol{\Psi}^{-1} \boldsymbol{K}_m \boldsymbol{\Psi}^{-\mathrm{T}} + \sigma_\eta^2 (\boldsymbol{\Psi}^{\mathrm{T}} \boldsymbol{\Psi})^{-1}\right)$$

が言える。以後，モデルパラメータ $\boldsymbol{a}, \theta, \sigma$ をまとめて Θ で表す。以上より，音声信号データ $\boldsymbol{x}$ が与えられたと

きの，モデルパラメータ Θ の対数尤度は具体的に

$$\log p(\boldsymbol{x};\Theta) = -\frac{1}{2}\left(I\log 2\pi + \log|\boldsymbol{\Sigma}| + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}\boldsymbol{x}\right) \quad (27)$$

$$\boldsymbol{\Sigma} = \sigma_w^2\sum_{m=1}^{M}\theta_m\boldsymbol{K}_m + \sigma_\eta^2\boldsymbol{I} \quad (28)$$

と書ける。ただし，$\boldsymbol{\Sigma}$ は声帯駆動信号 $\boldsymbol{\epsilon}$ の分散共分散行列のモデルである。以上の音声信号モデルをマルチカーネル線形予測モデルと呼ぶ。

2.5　スパース正則化パラメータ学習

ここで，モデルの過適応を防ぐ目的で，θ をスパースに誘導する事前分布として一般化正規分布

$$p(\theta) = \prod_{m=1}^{M}\frac{\lambda\beta}{2\Gamma(1/\beta)}\exp\left\{-\lambda^\beta|\theta_m|^\beta\right\} \quad (29)$$

を仮定する。ただし，$0<\beta<2$ とする。以上より，σ_w^2 と σ_η^2 が周期成分と白色成分のエネルギーの割合を，θ_m が周期成分に含まれる特定周期の優勢度をそれぞれ表すパラメータとなり，予測係数 $\boldsymbol{a}$ とともにこれらの推定も同時に行うことで，基本周波数の影響を軽減した予測係数推定が行えることが期待される。

以下では，式 (27), (29) によって構成されるパラメータ事後確率を最大化する問題を考える。Θ についての同時最適解は解析的に解くことはできないが，$\boldsymbol{a}$ と θ と σ それぞれに関して順番に目的関数を最大化する操作を繰り返すことで局所最適解を探索することができる。まず，θ と σ を固定した下での $\boldsymbol{a}$ の最適解は正規方程式

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\boldsymbol{a} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{x} \quad (30)$$

を解くことにより解析的に求まる。θ と σ については，EM(Expectation-Maximization) 法と補助関数法 [7] を組み合わせることで目的関数の増加を保証する更新式を導くことができる。まず，観測信号 $\boldsymbol{x}$ を，各周期の成分 $\sqrt{\theta_1}\boldsymbol{\Psi}^{-1}\Phi_1\boldsymbol{w},\cdots,\sqrt{\theta_M}\boldsymbol{\Psi}^{-1}\Phi_M\boldsymbol{w}$ および白色成分 $\boldsymbol{\Psi}^{-1}\boldsymbol{\eta}$ に対応する $M+1$ 個の独立な確率変数

$$\begin{aligned}\boldsymbol{y}_m &\sim \mathcal{N}\left(\boldsymbol{0},\sigma_w^2\theta_m\boldsymbol{\Psi}^{-1}\boldsymbol{K}_m\boldsymbol{\Psi}^{-\mathrm{T}}\right),\ m=1,\cdots,M\\ \boldsymbol{y}_{M+1} &\sim \mathcal{N}\left(\boldsymbol{0},\sigma_\eta^2(\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi})^{-1}\right)\end{aligned} \quad (31)$$

の和に分解し，これらを完全データと扱う。よって，完全データ $\boldsymbol{y}=(\boldsymbol{y}_1^{\mathrm{T}},\cdots,\boldsymbol{y}_{M+1}^{\mathrm{T}})^{\mathrm{T}}$ に対する対数尤度は，

$$\log p(\boldsymbol{y};\Theta) \stackrel{c}{=} -\frac{1}{2}\left(\log|\boldsymbol{\Lambda}| + \boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{y}\right) \quad (32)$$

$$\boldsymbol{\Lambda} = \begin{bmatrix}\sigma_w^2\theta_1\boldsymbol{\Psi}^{-1}\boldsymbol{K}_1\boldsymbol{\Psi}^{-\mathrm{T}} & & & O\\ & \ddots & & \\ & & \sigma_w^2\theta_M\boldsymbol{\Psi}^{-1}\boldsymbol{K}_M\boldsymbol{\Psi}^{-\mathrm{T}} & \\ O & & & \sigma_\eta^2(\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi})^{-1}\end{bmatrix} \quad (33)$$

で与えられる。ただし，$\stackrel{c}{=}$ は定数項以外の等号を表す。上式に対し，$\boldsymbol{x}$, $\Theta=\Theta'$ が与えられたときの条件つき期待値をとり，事前分布項 $\log p(\theta)$ を加えると，Q 関数

$$Q(\Theta,\Theta') = \log p(\theta) - \frac{1}{2}\Big(\log|\boldsymbol{\Lambda}| + \mathrm{tr}\left(\boldsymbol{\Lambda}^{-1}\mathbb{E}\left[\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}\middle|\boldsymbol{x};\Theta'\right]\right)\Big) \quad (34)$$

を得る。ここで，$\boldsymbol{H}:=[\boldsymbol{I},\cdots,\boldsymbol{I}]$ と置くと，不完全データ $\boldsymbol{x}$ と完全データ $\boldsymbol{y}$ との間には

$$\boldsymbol{x} = \boldsymbol{H}\boldsymbol{y} \quad (35)$$

なる関係式が成り立つことから，$\mathbb{E}\left[\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}\middle|\boldsymbol{x};\Theta\right]$ は

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}\middle|\boldsymbol{x};\Theta\right] &= \boldsymbol{\Lambda} - \boldsymbol{\Lambda}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}^{\mathrm{T}})^{-1}\boldsymbol{H}\boldsymbol{\Lambda}\\ &+ \boldsymbol{\Lambda}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}^{\mathrm{T}})^{-1}\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{H}^{\mathrm{T}})^{-\mathrm{T}}\boldsymbol{H}\boldsymbol{\Lambda}^{\mathrm{T}}\end{aligned} \quad (36)$$

と具体的に与えられる。式 (33) の各対角ブロックに対応する $\mathbb{E}\left[\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}\middle|\boldsymbol{x};\Theta'\right]$ の対角ブロックを $\boldsymbol{R}_1,\cdots,\boldsymbol{R}_{M+1}$ と置くと，Q 関数は

$$\begin{aligned}Q(\Theta,\Theta') \stackrel{c}{=} &-\sum_{m=1}^{M}\lambda^\beta|\theta_m|^\beta - \frac{I}{2}\log\left(\sigma_w^{2M}\sigma_\eta^2\theta_1\cdots\theta_M\right)\\ &-\frac{1}{2\sigma_w^2}\sum_{m=1}^{M}\frac{1}{\theta_m}\mathrm{tr}\left(\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{K}_m^{-1}\boldsymbol{\Psi}\boldsymbol{R}_m\right)\\ &-\frac{1}{2\sigma_\eta^2}\mathrm{tr}\left(\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi}\boldsymbol{R}_{M+1}\right)\end{aligned} \quad (37)$$

と書ける。ここで，$|\theta_m|^\beta$ $(\beta<2)$ に関する以下の不等式

$$|\theta_m|^\beta \le \beta|\theta'_m|^{\beta-1}(\theta_m-\theta'_m) + |\theta'_m|^\beta \quad (\theta_m>0) \quad (38)$$

を用いて $Q(\Theta,\Theta')$ の下限関数を設計することができ，補助関数法 [7] に基づいて σ と θ の更新式が導ける。これにより，θ_m の更新式は二次方程式

$$2\sigma_w^2\lambda^\beta\beta|\theta'_m|^{\beta-1}\theta_m^2 + I\sigma_w^2\theta_m - \mathrm{tr}\left(\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{K}_m^{-1}\boldsymbol{\Psi}\boldsymbol{R}_m\right) = 0$$

の正の根で与えられる。なお，ここでは式 (25) の制約を考慮しなかったため，上式に基づいて θ を更新後，強制的に式 (25) を満たすように規格化するステップを設けることにする。σ_w^2 および σ_η^2 の更新式についても解析的に得ることができ，それぞれ

$$\sigma_w^2 = \frac{1}{IM}\sum_{m=1}^{M}\frac{1}{\theta_m}\mathrm{tr}\left(\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{K}_m^{-1}\boldsymbol{\Psi}\boldsymbol{R}_m\right) \quad (39)$$

$$\sigma_\eta^2 = \frac{1}{I}\mathrm{tr}\left(\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi}\boldsymbol{R}_{M+1}\right) \quad (40)$$

で与えられる。

2.6　従来法 [2] との関係

[2] では，予測係数の更新と各標本時刻における予測残差の分散の更新を反復的に繰り返すことで，予測残差

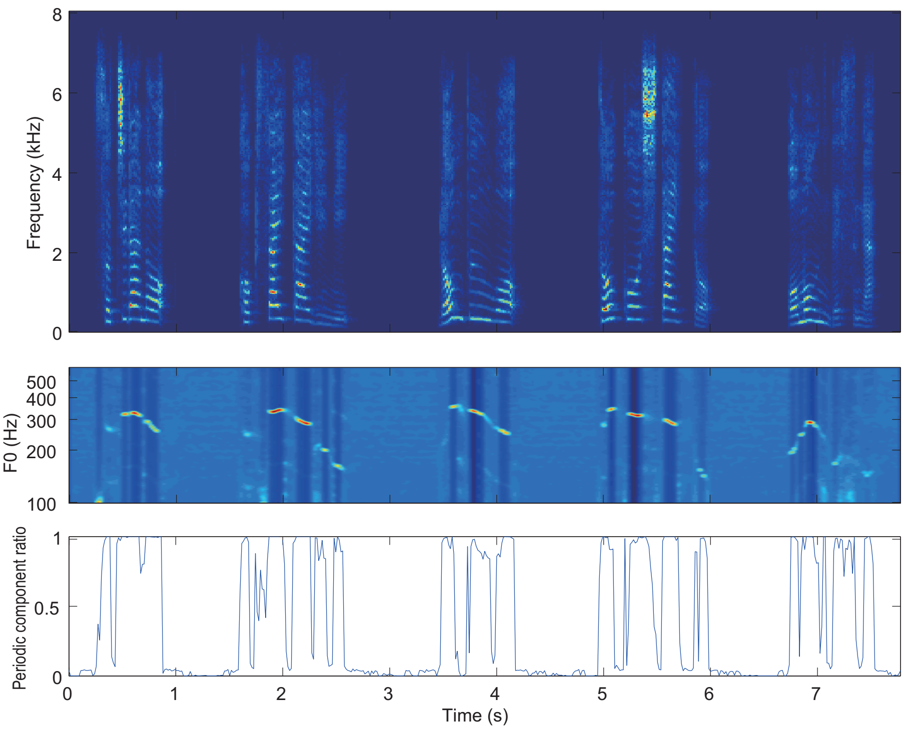

図 1　観測信号のスペクトログラム (上段) と推定した θ の分布 (中段) と周期・白色成分比 (下段)

の ℓ_1 ノルムを最小化する線形予測法を実現している。この手法は式 (30) において $\mathbf{\Sigma}$ を反復的に更新する点では提案法と共通しているが，その更新方法が異なる (従来法では $\mathbf{\Sigma}$ は対角行列であり，その対角要素は評価規準として各予測誤差値をどれだけ重く見積もるかを表したパラメータとなっている)。提案法とこの従来法の性質の違いについては稿を改めて議論したい。

3　動作実験

ATR 音声データベースの音声データ (女性話者) を用いて提案法の基本動作の確認を行った。標本化周波数 16kHz のディジタル信号に対し，フレーム長 64ms の Hanning 窓を 32ms シフトさせながら分析を行った。各周期カーネルは，100Hz から 600Hz までの間を 6cent 間隔で離散化した 518 個の基本周波数値 $1/T_m$ をもとに構成し，予測次数 P は 13 とした。また，アルゴリズムの反復計算の回数は 600 とした。

図 1 に，分析した音声信号のスペクトログラムと，提案法により推定された各時刻における $\theta_1, \cdots, \theta_M$ の分布，および，周期成分と白色成分のエネルギー比を表す $\sigma_w^2/(\sigma_w^2+\sigma_\eta^2)$ の値をグラフ化したものを示す。$\theta_1, \cdots, \theta_M$ の分布は各時刻の短時間信号においてどの基本周期が優勢であるかを表し，音声の基本周波数軌跡にピークが鮮明に現われていることが確認できる。

4　まとめ

本稿では，従来の線形予測モデルによる声道スペクトル推定において問題となっていた基本周波数の影響を解消することを目的とするだけでなく，有声無声判定と基本周波数推定も同時に行うことができるマルチカーネル線形予測分析法と呼ぶ新しい音声分析法を提案した。動作実験により，提案法の基本周波数推定機能，および，周期・白色成分比の推定機能を確認した。

謝辞　本研究に関して，有益な議論と情報提供を頂いた廣谷定男氏 (NTT CS 研) に感謝する。

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, C17–20, 1968.

[2] C.-H. Lee, “On robust linear prediction of speech,” IEEE trans. ASSP, **36**(5), 642–650, 1988.

[3] A. El-Jaroudi and J. Makhoul, “Discrete all-pole modeling,” IEEE trans. Signal Process., **39**(2), 411–423, 1991.

[4] 佐宗，田中，“HMM による音源のモデリングと高基本周波数に頑健な声道特性抽出，” 信学論 D-II, **J84-D-II**(9), 1960–1969, 2001.

[5] Y. Miyoshi et al., “Analysis of speech signals of short pitch period by a sample-selective linear prediction,” IEEE trans. ASSP, **35**(9), 1233–1240, 1987.

[6] C.E. Rasmussen and C.K.I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, Mass, USA, 2006.

[7] 亀岡，小野，柏野，嵯峨山，“複素 NMF: 新しいスパース信号分解表現と基底系学習アルゴリズム，” 音講論 (秋), 2-8-13, 657–660, 2008.