

予測誤差の Golomb-Rice 符号量を最小化する線形予測分析

亀岡 弘和[†] 鎌本 優[†] 原田 登[†] 守谷 健弘[†]

A Linear Predictive Coding Algorithm Minimizing the Golomb-Rice Code Length of the Residual Signal

Hirokazu KAMEOKA[†], Yutaka KAMAMOTO[†], Noboru HARADA[†], and Takehiro MORIYA[†]

あらまし 本論文では、予測誤差ベクトルの l_2 ノルムを最小化する従来の線形予測分析に基づく可逆圧縮符号化方式は予測誤差の符号量を直接的に小さくするものではなかった点に着目し、予測誤差の符号量をできるだけ小さくする線形予測分析の効率的なアルゴリズムの導出を目的とする。予測誤差の符号化に利用する Golomb-Rice 符号における振幅値と符号量との関係をよく表すコスト関数モデルを考え、これに基づいて立てられる目的関数（予測誤差ベクトルの l_1 ノルム等）の最小解を効率的に求める反復アルゴリズム（補助関数法）を提案する。評価実験により、提案法の動作と提案法による圧縮性能を確認した。

キーワード 可逆圧縮符号化, 線形予測分析, Golomb-Rice 符号, l_1 ノルム, 補助関数法

1. ま え が き

線形予測分析に基づく時系列信号の可逆圧縮符号化方式は、時系列信号を線形予測分析し、その結果求まる線形予測係数あるいは PARCOR 係数と予測誤差（残差）を符号化して伝送し、受信側で無歪に復号化できる仕組みとなっている。線形予測分析により生成される残差信号の振幅は通常 0 付近に集中するという性質を利用し、出現頻度の高い値ほど短い符号を割り当てる Golomb-Rice 符号 [1] をはじめとするエントロピー符号化を残差の符号化に用いることで全体の符号量を小さく抑えようとする点がこの方式の特徴である。Golomb-Rice 符号の場合、残差振幅に割り当てられる符号量は、振幅の絶対値とほぼ比例関係にあるため、残差符号量は残差振幅の絶対値和である程度よく近似される。したがって、残差振幅の絶対値和が小さいほど符号量はより小さくできる可能性がある。しかしながら、従来の線形予測分析による可逆圧縮符号化方式では、残差振幅の二乗和を最小化するように予測係数を求めるため、残差符号量を直接的に最小化する規準

での最適化となっていない。線形予測分析による時系列信号の可逆圧縮符号化の国際標準 MPEG-4 Audio Lossless Coding (ALS) [2] では、ビットストリームや復号化の手順の標準は規定されているものの、符号化の手段には自由度が残されているため、予測係数の算出方法には符号量削減の余地が少なからず残されている。もし予測係数の算出法の改善によって標準準拠の参照ソフトウェアに比べて圧縮率をたとえわずかも向上できれば、標準準拠の拘束から外れずに圧縮率向上が達成されたことになり大変価値が高い。

線形予測分析を包含する枠組である線形回帰分析において、誤差ベクトルの要素の絶対値和を最小化する回帰係数の解法は、その最適化問題が線形計画の問題として定式化できる [3] ことから、シンプレックス法に基づくもの [4] や内点法に基づくもの [5] や有効制約法に基づくもの [6] などがよく知られる。このほかに有効な解法として、重み付きメジアン法 [7]、反復再重み付き最小二乗誤差推定 (Iteratively Reweighted Least Squares, IRLS) 法 [8] などが提案されている。これらの手法は、線形予測分析による可逆圧縮符号化における残差符号量削減という本論文で提起する目的意識とは別に、最小絶対値推定量がロバスト推定量であることから線形予測分析に基づく音声分析の耐雑音性能を向上する目的として応用されている [9], [10]。

[†] 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, 厚木市
NTT Communication Science Laboratories, NTT Corporation, Atsugi-shi, 243-0198 Japan

我々の問題設定においては、残差の絶対値はあくまで Golomb-Rice 符号における残差振幅値と符号量との関係を近似した、いわば「符号量モデル」の一つという位置付けにすぎず、残差の絶対値とより更に良い目的関数がほかに立てられる可能性がある。上記従来法が他の目的関数に対しても必ずしも適用可能というわけではないため、このことが、より一般的な予測係数算出方法を開発する本研究の動機となっている。本論文では、残差符号量をできるだけ小さくする線形予測分析の効率的なアルゴリズムの確立を目的とし (1) 誤差ベクトルの要素の絶対値和を最小化する回帰係数を効率的に求める手法の導出を出発点とし、(2) Golomb-Rice 符号量関数を絶対値和に比べてよりよく表現する目的関数をいくつか導入し (3) これらを最小化する回帰係数も同様の原理に基づいて求められることを示す。そして以上の手法を線形予測分析に基づく可逆圧縮符号化に応用し、圧縮性能評価を行う。

なお、次章より明らかにするが、導出過程は異なるものの提案法は IRLS 法 [8] と形式的に非常に似た手法である。IRLS 法は、重み付き最小二乗誤差推定を反復的に重みを更新しながら行う手法であるのに対し、提案法は、重みだけでなく観測値を仮想的にシフトする補正項も更新しながら重み付き最小二乗誤差推定を反復的に行う手法である。IRLS 法は、重みの更新規範が目的関数の一次近似に基づいており、限られた目的関数に対してしか収束性は保証されない。これに対し提案法の原理的な特長は、重みと補正項の更新公式が目的関数の単調減少性を保証するように導き出されるため、IRLS 法では収束が保証されない目的関数に対しても反復アルゴリズムの収束性が保証される点にある。このため、提案法は IRLS 法より幅広いクラス の目的関数を扱うことができる。

2. 絶対誤差を最小化する線形回帰分析

2.1 問題設定

\mathbf{x}, \mathbf{s} をそれぞれ I 次元及び J 次元ベクトル、 \mathbf{A} を $I \times J$ 行列とし、次のような線形回帰モデル

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{z} \quad (1)$$

を仮定 (ただし、 \mathbf{z} は雑音を表す確率ベクトル) し、 \mathbf{x}, \mathbf{A} が与えられたもとで、 \mathbf{s} の最よう推定値を求める線形回帰分析の問題を考える。 \mathbf{z} にもしガウス性白色雑音を仮定すると、これは ℓ_2 ノルム $\|\mathbf{x} - \mathbf{A}\mathbf{s}\|_2^2$

を \mathbf{s} に関して最小化する問題であり、 $I \geq J$ ならば $\hat{\mathbf{s}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{x}$ が最適解となる。これに対し、我々がここで考えるのは、 \mathbf{z} の各要素が互いに独立にラプラス分布に従う場合を仮定した \mathbf{s} の最よう解を求める問題である。すなわち、 ℓ_1 ノルム (絶対誤差)

$$L(\mathbf{s}) \equiv \|\mathbf{z}\|_1 = \|\mathbf{x} - \mathbf{A}\mathbf{s}\|_1 = \sum_{i=1}^I |x_i - (\mathbf{A}\mathbf{s})_i| \quad (2)$$

を最小化する \mathbf{s} がこの場合の最よう解に対応する。この推定量のことは ℓ_1 推定量と呼び、 \mathbf{x} の外れ値に頑健なロバスト推定量の一つとして知られる。

2.2 線形予測モデルとの対応

線形予測モデルは式 (1) の特殊例である。線形予測モデルは、あるデジタル信号の時刻 i におけるサンプル y_i を、時刻 $i-1$ から過去 J 個のサンプル $y_{i-1}, y_{i-2}, \dots, y_{i-J}$ の線形結合で表現するモデル

$$y_i \simeq \sum_{j=1}^J s_j y_{i-j} \quad (3)$$

であり、自己回帰モデルとも呼ばれる。 I 個の連続するサンプル y_1, \dots, y_I をすべてを最もよく表す線形予測モデルの J 個の係数を推定することを線形予測分析と呼ぶ。式 (1) に対応させると、線形予測モデルは、

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_I \end{pmatrix} = \begin{pmatrix} y_0 & y_{-1} & \dots & y_{1-J} \\ y_1 & y_0 & \dots & y_{2-J} \\ \vdots & \vdots & & \vdots \\ y_{I-1} & y_{I-2} & \dots & y_{I-J} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_J \end{pmatrix} + \mathbf{z} \quad (4)$$

とおいたことに相当し、 \mathbf{z} が残差ベクトルである。

2.3 補助関数法の原理

提案法の特徴は、式 (2) の最小化に、補助関数法と呼ぶ原理を用いる点にある。補助関数法とは、適切に設計した補助関数を反復的に減少させていくことで目的関数を間接的に最小化する手法であり、EM アルゴリズムを一般化した原理である。EM アルゴリズム同様に補助関数法はいかなる最適化問題に適用可能というわけではないが、ある要件を満たす補助関数が設計できれば大変効率の良い最適化アルゴリズムが導ける場合がある。以下に、補助関数法の原理を示し、補助関数が満たすべき要件を明らかにする。

目的関数 $L(\mathbf{s})$ に対し、

$$L(\mathbf{s}) = \min_{\mathbf{B}} L^+(\mathbf{s}, \mathbf{B}) \quad (5)$$

が成り立つとき、 $L^+(s, \mathbf{B})$ を $L(s)$ の補助関数、 \mathbf{B} を補助変数と定義すると、次の定理が成り立つ。

定理 1 (補助関数法) 補助関数 $L^+(s, \mathbf{B})$ を、 \mathbf{B} に関して最小化するステップと、 s に関して最小化するステップを繰り返すことで、目的関数 $L(s)$ を単調減少させていくことができる。

(証明) $\mathbf{B}_{t+1} = \operatorname{argmin}_{\mathbf{B}} L^+(s_t, \mathbf{B})$ とし、 $s_{t+1} = \operatorname{argmin}_s L^+(s, \mathbf{B}_{t+1})$ とする。ただし、 t は反復計算のステップ数を表す。 $s = s_t$ 、 $\mathbf{B} = \mathbf{B}_t$ から $s = s_{t+1}$ 、 $\mathbf{B} = \mathbf{B}_{t+1}$ に更新されたときに、 $L(s)$ が非増加することを示す。明らかに $L(s_t) = L^+(s_t, \mathbf{B}_{t+1})$ であり、 $s_{t+1} = \operatorname{argmin}_s L^+(s, \mathbf{B}_{t+1})$ より、 $L^+(s_t, \mathbf{B}_{t+1}) \geq L^+(s_{t+1}, \mathbf{B}_{t+1})$ である。更に補助関数の定義より、 $L^+(s_{t+1}, \mathbf{B}_{t+1}) \geq L(s_{t+1})$ なのだから、結局、 $L(s_t) \geq L(s_{t+1})$ である。□

以上より、補助関数法を効果的に適用するには、次の 2 点を満たす補助関数を設計する必要がある。

(要件 1) $\operatorname{argmin}_{\mathbf{B}} L^+(s, \mathbf{B})$ が解析的に求められる。
 (要件 2) $\operatorname{argmin}_s L^+(s, \mathbf{B})$ が解析的に求められる。
 更に、要件 1, 2 を満たす $L^+(s, \mathbf{B})$ が複数存在する場合、その中で $L(s)$ に最も近いものを補助関数として採用することが望ましい。これを踏まえ、式 (2) の最小化問題を解くための補助関数を以下で設計する。

2.4 二次形式の補助関数の導出

式 (2) の最小化のための補助関数を設計するにあたり、絶対値関数より小さくなることなく (補助関数の定義より)、かつ、絶対値関数に接する (要件 1 より) 二次関数 (要件 2 より) を見つけたいが、このような二次関数は無数に存在する。二次関数を絶対値関数にできるだけ近づけるためには二次の係数を小さくとればよいので、条件を満たす二次関数の中で二次の係数が最小となるものを選べばよい。絶対値関数 $f_0: \mathbb{R} \rightarrow \mathbb{R}_+$

$$f_0(z) = |z| \quad (6)$$

に $b \neq 0$ で接する原点に関して対称な二次関数 (図 1)

$$h_0(z) = \frac{z^2}{2|b|} + \frac{|b|}{2} \quad (7)$$

が $f_0(z)$ を下回らず、かつ、 $f_0(z)$ に接する二次関数の中で二次の係数が最小となる。すなわち、

$$|z| \leq \frac{z^2}{2|b|} + \frac{|b|}{2} \quad (8)$$

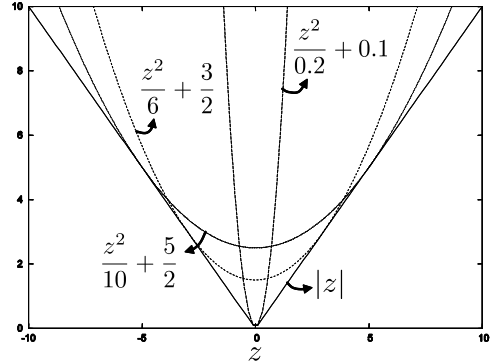


図 1 絶対値関数とそれに接する放物線 ($b = 0.1, 3, 5$)
 Fig. 1 Quadratic functions with $b = 0.1, 3, 5$ tangent to the absolute value function.

が成り立ち、等号は接点 $z = b$ において成立する。この不等式を式 (2) に当てはめると、

$$L(s) \leq \sum_{i=1}^I \left(\frac{(x_i - (\mathbf{A}s)_i)^2}{2|b_i|} + \frac{|b_i|}{2} \right) \quad (9)$$

$$= (\mathbf{x} - \mathbf{A}s)^T \mathbf{B} (\mathbf{x} - \mathbf{A}s) + d \equiv L^+(s, \mathbf{B})$$

のような不等式が立てられる。ただし、

$$\mathbf{B} = \begin{pmatrix} \frac{1}{2|b_1|} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{2|b_I|} \end{pmatrix}, \quad d = \sum_{i=1}^I \frac{|b_i|}{2} \quad (10)$$

である。以上より、 $L^+(s, \mathbf{B})$ は補助関数の定義を満たす。次に、 $L^+(s, \mathbf{B})$ が補助関数の要件を満たすかどうかを確認する。

2.5 予測係数と補助変数の更新則

まず、 $L(s) = L^+(s, \mathbf{B})$ となる \mathbf{B} は明らかに、

$$\mathbf{B} = \begin{pmatrix} \frac{1}{2|x_1 - (\mathbf{A}s)_1|} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{2|x_I - (\mathbf{A}s)_I|} \end{pmatrix} \quad (11)$$

である。一方、 \mathbf{B} 固定のときに $L^+(s, \mathbf{B})$ を最小化する s は、式 (11) より、大きい残差には小さい重みを、小さい残差には大きい重みを課す重み付き最小二乗誤差推定量に相当し、正規方程式の解

$$s = (\mathbf{A}^T \mathbf{B} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{B} \mathbf{x} \quad (12)$$

で与えられる。 $\mathbf{A}^T \mathbf{B} \mathbf{A}$ は正定値対称行列であるので、 s は Cholesky 分解で求めるとよい。以上より、 $L^+(s, \mathbf{B})$ は補助関数の要件を満たし、次の手順

- (1) s の初期値設定
- (2) s と B を次のように交互に更新

$$B \leftarrow \begin{pmatrix} \frac{1}{2|x_1 - (As)_1|} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{2|x_I - (As)_I|} \end{pmatrix}$$

$$s \leftarrow (A^T B A)^{-1} A^T B x$$

により、回帰係数の絶対誤差最小解を得ることができる。式 (2) は s に関して凸なので、収束解は大域最適解に一致する。ただし、大域最適解は一意とは限らず (大域最適解の集合は解空間上で凸集合をなし)、提案法はその中の 1 点を求めるものである。以上のアルゴリズムは、ロバスト推定の一つである Huber の M 推定 [11] に用いられる IRLS 法 [8] と同じ手続きをとる。 ℓ_1 推定量を得るための上記手順においては IRLS 法との間に違いはないが、次章で違いを明らかにする。

3. 残差符号量を最小化する線形予測分析

3.1 Golomb-Rice 符号量モデル

振幅 $z \in \mathbb{Z}$ に対する Golomb-Rice 符号量 $R(z)$ は、

$$R(z) = \begin{cases} \left\lfloor \frac{z}{2^{r-1}} \right\rfloor + r + 1 & (z \geq 0) \\ \left\lfloor \frac{-z-1}{2^{r-1}} \right\rfloor + r + 1 & (z < 0) \end{cases} \quad (13)$$

と表せる。ただし、 $\lfloor \cdot \rfloor$ は切捨て整数化演算子を表し、 $r \in \mathbb{N}$ は Golomb-Rice パラメータである。2. で最小化しようとした残差振幅の絶対誤差は、 $r = 1$ とおいた場合の Golomb-Rice 符号量のモデルに相当する。しかし Golomb-Rice 符号化では実際のところ $r > 1$ と設定することが多く、式 (13) は図 2 のように階段形の関数となる。ここで特に注目すべきは、 $z = 0$ 周辺の一定範囲内の値には等しい符号量が割り当てられる点であり、このことは、残差振幅の絶対値をできるだけ小さくすることだけに専念するよりも、0 付近の残差振幅には適度な「遊び」をもたせ、その自由度の分だけ極端な外れ値を少なくする方が全体の符号量をより小さくできることを示している。残差振幅が通常 0 付近に集中する点を考慮すると、残差振幅にかかる符号量に則した目的関数は、 $z = 0$ 周辺の $R(z)$ をよく近似したもので表された方がよいと考えられる。

以上の問題意識に基づき、絶対値関数に代わる次のような 3 種類の Golomb-Rice 符号量モデル $f_1(z), f_2(z), f_3(z)$

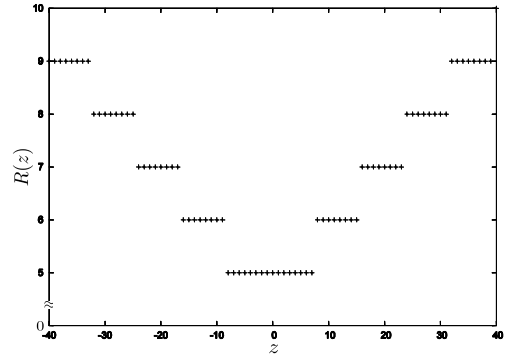


図 2 Golomb-Rice 符号量 $R(z)$ ($r = 4$)
Fig. 2 Golomb-Rice code length $R(z)$ with $r = 4$.

$$f_1(z) \equiv \begin{cases} \frac{z^2}{4\beta\gamma} + \frac{\beta}{\gamma} & (|z| \leq 2\beta) \\ |z|/\gamma & (|z| > 2\beta) \end{cases}$$

$$f_2(z) \equiv \frac{\sqrt{z^2 + \beta^2}}{\gamma}$$

$$f_3(z) \equiv \begin{cases} \beta/\gamma & (|z| \leq \beta - c) \\ \frac{(z - |\beta - c|)^2}{4c\gamma} + \frac{\beta}{\gamma} & (\beta - c < |z| \leq \beta + c) \\ |z|/\gamma & (|z| > \beta + c) \end{cases}$$

を導入する。 $\beta/\gamma = r + 1, \gamma = 2^{r-1}$ であり、 β, γ は Golomb-Rice パラメータに依存して一意に決まる定数である。 $f_1(z)$ は Huber 尺度、 $f_2(z)$ は ℓ_1 - ℓ_2 尺度と呼ばれる。 c は $0 < c \leq \beta$ の範囲で適当に定める定数であり、 $c \rightarrow 0$ で $f_3(z)$ は Vapnik's ϵ -insensitive [12] に漸近する。また、 $c = \beta$ で $f_1(z) = f_3(z)$ となる。各々の概形を図 3 に示す。以上の符号量モデルを用い、残差符号の全符号量を近似した 3 種類の目的関数

$$L_n(s) = \sum_{i=1}^I f_n(x_i - (As)_i) \quad (n = 1, 2, 3) \quad (14)$$

が立てられる。いずれの目的関数においても s の最適解は γ によらないので、以後簡単のため $\gamma = 1$ とする。なお、これらはいずれも s に関してやはり凸であり、局所解と大域最適解は一致する。

これらの目的関数はいずれも s に関して非線形であり、線形計画法をベースにした手法や、重み付きメジアン法は適用できない。また、いずれの目的関数においても、最適解近傍以外のあらゆる点で Hesse 行列がほぼ零行列になると考えられ、ニュートン法をはじめとするこの配法はうまく動作しないことが予想される。これに対し、2. で提案した原理はこれらの目的関数

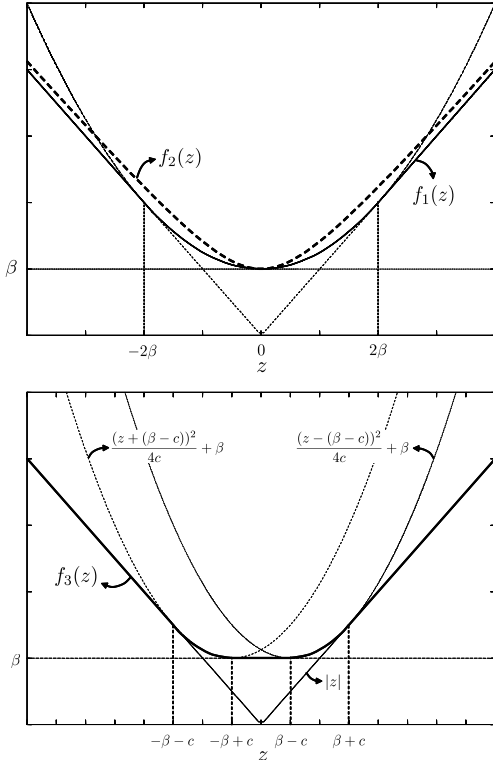


図 3 Golomb-Rice 符号量モデル $f_n(z)$ ($n = 1, 2, 3$)
Fig. 3 Golomb-Rice code length models $f_n(z)$.

の最小化に対しても同様に適用できる。

3.2 補助関数の導出

各々の符号量モデル $f_n(z)$ ($n = 1, 2, 3$) に対し、2. と同様の方針で、二次関数 $h_n(z)$

$$h_1(z) = \frac{z^2}{2 \max(|b|, 2\beta)} + \frac{\max(|b|, 2\beta)}{2}$$

$$h_2(z) = \frac{z^2}{2\sqrt{b^2 + \beta^2}} + \frac{b^2 + 2\beta^2}{2\sqrt{b^2 + \beta^2}}$$

$$h_3(z) = \begin{cases} \frac{z^2}{2|b|} + \frac{|b|}{2} & (|b| \geq 2\beta) \\ \frac{(z - (2\text{sign}(b)\beta - b))^2}{4(|b| - \beta)} + \beta & (\beta + c < |b| < 2\beta) \\ \frac{(z - \text{sign}(b)(\beta - c))^2}{4c} + \beta & (\beta - c < |b| < \beta + c) \\ \frac{(z - b)^2}{4(\beta - |b|)} + \beta & (0 \leq |b| < \beta - c) \end{cases}$$

がそれぞれ導出できる ($h_3(z)$ の導出は付録参照)。
 $h_n(z)$ ($n = 1, 2, 3$) はいずれも $f_n(z)$ と点 b で接し、 $f_n(z)$ より決して小さくなることのない二次関数の中で二次の係数が最小のものである。これにより、それぞれの目的関数について統一的に

$$L_n(s) \leq (\mathbf{x} - \mathbf{A}s - \mathbf{w}_n)^T \mathbf{B}_n (\mathbf{x} - \mathbf{A}s - \mathbf{w}_n) + d_n \equiv L^+(s, \mathbf{B}_n, \mathbf{w}_n) \quad (15)$$

なる不等式が立てられ、補助関数 $L^+(s, \mathbf{B}_n, \mathbf{w}_n)$ が設計できる。ここでは、 \mathbf{B}_n だけでなく \mathbf{w}_n も補助変数であることに注意が必要である。 \mathbf{w}_n は、コストが最小となる残差振幅ベクトル値を表す。すなわち、 \mathbf{w}_n が 0 以外の値をとる場合には、残差振幅値の絶対値を 0 に近づけさえすればよいというわけではなく、 \mathbf{w}_n を補助変数とすることは、反復計算の過程で残差振幅値がある程度小さな値になってさえいれば十分とみなし、その分、他の大きな残差振幅値を小さくするよう促す効果をもたらすのである。

さて、2. 同様、補助関数法の定理に基づき、

- (1) s の初期値設定
- (2) s と $\mathbf{B}_n, \mathbf{w}_n$ を次のように交互に更新

$$\mathbf{B}_n \leftarrow \begin{pmatrix} b_1^{(n)} & & 0 \\ & \ddots & \\ 0 & & b_I^{(n)} \end{pmatrix}, \quad \mathbf{w}_n \leftarrow \begin{pmatrix} w_1^{(n)} \\ \vdots \\ w_I^{(n)} \end{pmatrix}$$

$$s \leftarrow (\mathbf{A}^T \mathbf{B}_n \mathbf{A})^{-1} \mathbf{A}^T \mathbf{B}_n (\mathbf{x} - \mathbf{w}_n)$$

により、目的関数 $L_n(s)$ を最小化する回帰係数を得ることができる。行列 \mathbf{B}_n の対角要素 $b_i^{(n)}$ 及びベクトル \mathbf{w}_n の要素 $w_i^{(n)}$ の更新則は、符号量モデルの選択に応じて

$$b_i^{(1)} = \frac{1}{2 \max(|z_i|, 2\beta)}, \quad w_i^{(1)} = 0$$

$$b_i^{(2)} = \frac{1}{2\sqrt{z_i^2 + \beta^2}}, \quad w_i^{(2)} = 0$$

$$b_i^{(3)} = \begin{cases} \frac{1}{2|z_i|} & (|z_i| \geq 2\beta) \\ \frac{1}{4(|z_i| - \beta)} & (\beta + c \leq |z_i| < 2\beta) \\ \frac{1}{4c} & (\beta - c \leq |z_i| < \beta + c) \\ \frac{1}{4(\beta - |z_i|)} & (0 \leq |z_i| < \beta - c) \end{cases}$$

$$w_i^{(3)} = \begin{cases} 0 & (|z_i| \geq 2\beta) \\ 2\text{sign}(z_i)\beta - z_i & (\beta + c \leq |z_i| < 2\beta) \\ \text{sign}(z_i)(\beta - c) & (\beta - c \leq |z_i| < \beta + c) \\ z_i & (0 \leq |z_i| < \beta - c) \end{cases}$$

と与えられる。ただし、 $z_i = x_i - (\mathbf{A}s)_i$ である。

なお、IRLS 法では目的関数 $L_3(s)$ の単調減少は保証されない。この点が、補助関数法により導出された提案法と IRLS 法の原理的な違いと優位性を端的に示

している。

また、2. のアルゴリズムは、反復計算の過程で、

$$|x_i - (As)_i| = 0 \quad (16)$$

となる i が少なくとも一つ存在すると、行列 $A^T B A$ は非正則となり、反復計算が続けられなくなるという不安定さがあつたのに対し、本章で述べた B_n の更新式ではいずれも行列の対角要素が 0 より大きいため、 B_n が原因となって行列 $A^T B_n A$ が非正則になることはないという利点がある。

4. 評価実験

4.1 収束性の確認

まず、提案法の収束性を確認するための評価実験を行った。実験データは文献[13]に収録されている音声認識実験用の音声データ（サンプリング周波数：16 kHz，振幅語長：16 bit，モノラル，合計約 140 秒）を用い、線形予測分析の次数は 4，8，12 でそれぞれ固定とし、分析フレームは 1024 サンプルとして、2. の各反復計算における絶対値和（各フレームにおける残差振幅の絶対値和の合計値）の挙動を調べた。

図 4 は、各次数（四次：四角印，八次：三角印，十二次：丸印）において補助関数法の更新によりどれだけ絶対値和が変化していくかを表しており、初期値の Levinson-Durbin 法で求めたときの絶対値和で正規化している。いずれの次数でも 10 回程度の反復計算で収束していることが分かる。四次の場合は二乗誤差規準と比べて約 3% 絶対値和が小さくなっており、十二次では約 5% 小さくなっている。このことから、処理時間と最小化の度合のトレードオフを考えると、5 回程度の反復計算で十分な性能が得られると考えられる。

2. の収束解が大域最適解に一致することを確認するため、異なる初期値でも同じ絶対値和に収束するかどうかを調べた。図 5 は四次の予測において、初期値を Levinson-Durbin 法と Burg 法の 2 通りで求めた場合の収束性の違いを表している。前図と同様に Levinson-Durbin 法で求めたときの絶対値和で正規化している。いずれの初期値から反復計算を始めても、絶対値和がほぼ同一の値に収束していくのが分かる。八次や十二次の場合も、同様の結果であった。

両図において、反復計算を繰り返すごとに絶対値和はほぼ単調に減少している。単調な収束性をもつ提案法は、処理時間と圧縮性能のトレードオフのバランスを自由に調節できたり、初期値に依存しないことから

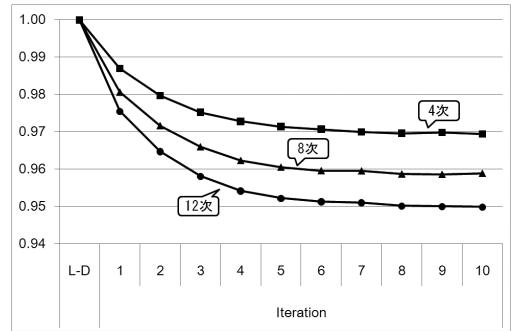


図 4 各次数における収束の様子（横軸：初期値及び反復計算回数，縦軸：初期値で正規化された絶対値和の値，初期値は Levinson-Durbin (L-D) 法の解）

Fig. 4 Transitions of the sum of the absolute value of predictive errors divided by the initial residue, obtained via Levinson-Durbin algorithm.

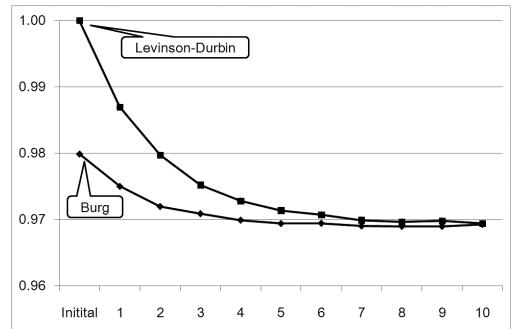


図 5 初期値を Levinson-Durbin 法と Burg 法で求めた場合の絶対誤差収束の様子

Fig. 5 Transitions of the absolute errors whose initial points were given respectively by Levinson-Durbin algorithm and Burg method.

既存アルゴリズムの後処理として組み込むこともできるため実用的である。

4.2 圧縮率の確認

ここでは、音響信号を対象として、提案法の圧縮性能の評価実験を行った。本実験の目的は、提案法の原理を実証することだけでなく、線形予測分析に基づく可逆圧縮符号化において残差の符号量削減の余地がどの程度残されているのを見極めることである。

評価実験用に、RWC 研究用音楽データベース [14] に含まれる、サンプリング周波数 44.1 kHz，16 ビットで収録されたステレオの音楽ファイル（WAV 形式）10 個（RWC-MDB-P-2001 No.33～42）を実験データ（合計約 432 MByte）として用いた。分析フレーム長は 2048 サンプルとした。

Levinson-Durbin 法に基づく二乗誤差規準の線形予

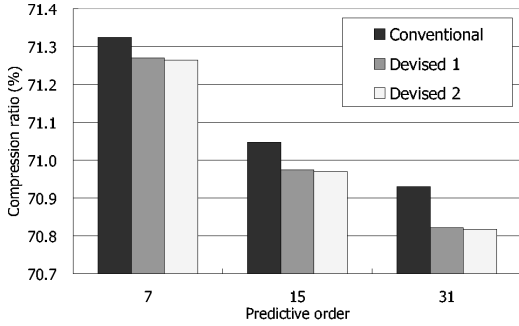


図 6 二乗誤差規準の線形予測符号化 (従来法) に対する, $L_2(s)$ 及び $L_3(s)$ を規準とした線形予測符号化 (提案法 1&2) による圧縮率の比較

Fig. 6 Comparison of the compression ratios among the devised methods and the conventional ℓ_2 -norm-based-LPC.

測分析の従来の方式を従来法, 従来法で得られた予測係数を初期値として, 3.1 で示した $L_2(s)$ 及び $L_3(s)$ を規準とした反復最適化アルゴリズムによる線形予測分析の方式をそれぞれ提案法 1, 提案法 2 と略称する. 提案法 1 及び提案法 2 の反復計算回数は 10 とした. なお, 各方式ではいずれもチャンネル間相関に基づく圧縮は行っていない. また, 予測次数 (式 (4) における J) は固定とした.

$L_2(s)$ は ℓ_2 ノルムに比べて Golomb-Rice 符号量に則した目的関数となっているため, 従来法に比べ, 提案法 1 の圧縮率は向上するものと予想される. 一方, $L_3(s)$ は $L_2(s)$ に比べてより Golomb-Rice 符号量に則した目的関数となっているため, 提案法 1 に比べ, 提案法 2 の圧縮率は若干向上するものと予想される.

従来法, 提案法 1, 提案法 2 それぞれについての, 予測次数が 7, 15, 31 の場合の圧縮率を図 6 に示す. 圧縮率は, 次のように算出した.

$$\text{圧縮率 (\%)} = \frac{\text{圧縮後のファイルサイズ}}{\text{もとのファイルサイズ}} \times 100$$

図 6 より, 予想どおり, 提案法 2 による圧縮率が最も高く, 次いで提案法 1 による圧縮率が高かった. 以上の結果から, 提案法の有効性と, 提案法により実現された残差符号量削減の度合を確認することができた. なお, $L_1(s)$ を規準とした場合と $L_2(s)$ を規準とした場合とで圧縮率の差はほとんど見られなかったため, $L_1(s)$ を規準とした場合の結果については省略した.

図 7 は予測次数が十五次の場合の圧縮率と演算速度の関係を示している. 圧縮にかかる時間は “timeit.exe” により計測し, リアルタイムファクタは

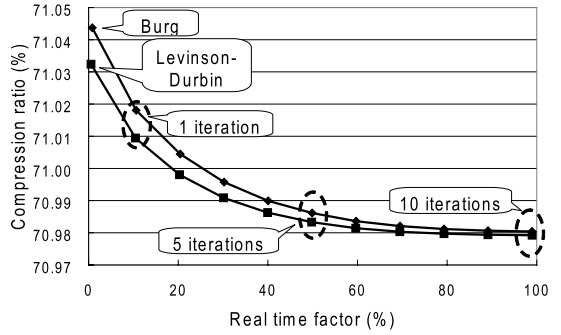


図 7 予測次数が十五次の場合の圧縮率と演算速度の関係
Fig. 7 Trade-off between the compression ratio and the computational cost when setting the number of LPC coefficients to 15.

リアルタイムファクタ (%)

$$= \frac{\text{圧縮にかかった総時間}}{\text{総演奏時間}} \times 100$$

とした. 100%のときに実時間で動作し, 10%であれば 10 倍速く動作する. つまりこの値が小さいほど性能が良い. また実験には Windows Server 2003, AMD Opteron プロセッサ 2.4 GHz, メモリ 2 GByte を搭載した計算機を用いた.

初期値を Levinson-Durbin 法で求めても, Burg 法で求めても, 10 回程度で収束していることが分かる. また, 10 回の反復計算を行っても, 実時間よりも高速に処理を終えることができることが分かる.

5. む す び

本論文では, 従来の線形予測分析による可逆圧縮符号化は残差符号量を直接的に小さくするものではなかった点に着目し, 残差符号量をできるだけ小さくする線形予測分析法の確立を目的とし (1) 残差の符号化に利用される Golomb-Rice 符号における振幅と符号量との関係をモデル化し, (2) Golomb-Rice 符号量モデルに基づいて立てられる目的関数の最小解を効率的に求める反復アルゴリズム (補助関数法) を考案した.

実験結果より, 提案法は初期値から 10 回程度の反復計算で単調に収束することと, 初期値には依存せずに収束することが確認された. ただし, ここでの初期値は二乗誤差最小化規準で求めたものであるという条件付きである. 単調な収束性をもつ提案法は, 処理時間と圧縮性能のトレードオフのバランスを自由に調節できたり, 容易に既存アルゴリズムに組み込むことが

できるため実用的である。また、提案法により圧縮率の向上を確認した。

一方、原理面では、提案法は、ある条件を満たす目的関数を規準とする場合には IRLS 法と同一の反復アルゴリズムを与えることが明らかになった。しかしながら、IRLS 法では収束性が保証されない目的関数に対しても、提案法の枠組みでは収束性が保証された反復アルゴリズムを導けることを例示した。このことで、提案法は IRLS 法より幅広いクラスの目的関数を扱える方法論であることが示される。

謝辞 本研究は、東京大学大学院小野順貴講師、嵯峨山茂樹教授との議論を通して進展したものです。有益な議論をして頂いた御両名に感謝致します。

文 献

- [1] S.W. Golomb, "Run-length encodings," IEEE Trans. Inf. Theory, vol.12, no.3, pp.399-401, 1966.
- [2] T. Liebchen and Y. Reznik, "MPEG-4 ALS: An emerging standard for lossless audio coding," Proc. IEEE Data Compression Conference 2004, pp.439-448, 2004.
- [3] T.E. Harris, "Regression using minimum absolute deviations," American Statistician, vol.4, no.1, pp.14-15, 1950.
- [4] I. Barrodale and F.D.K. Roberts, "An improved algorithm for discrete ℓ_1 linear approximation," SIAM J. Numer. Anal., vol.10, no.5, pp.839-848, 1973.
- [5] Y. Zhang, "Primal-dual interior point approach for computing ℓ_1 -solutions, and ℓ_∞ -solutions of overdetermined linear systems," J. Optim. Theory Appl., vol.77, no.2, pp.323-341, 1993.
- [6] P. Bloomfield and W. Steiger, "Least absolute deviations curve-fitting," SIAM J. Sci. Stat. Comp., vol.1, no.2, pp.290-301, 1980.
- [7] R.W. Hawley and N.C. Gallagher, Jr., "On edgeworth's method for minimum absolute error linear regression," IEEE Trans. Signal Process., vol.42, no.8, pp.2045-2054, 1994.
- [8] P. Green, "Iterative reweighted least squares for maximum likelihood estimation and some robust and resistant alternative (with discussions)," J. Roy. Statist. Soc., Series B, vol.46, pp.149-162, 1984.
- [9] E. Denoël and J.-P. Solvay, "Linear prediction of speech with a least absolute error criterion," IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-33, no.6, pp.1397-1403, 1985.
- [10] C.-H. Lee, "On robust linear prediction of speech," IEEE Trans. Acoust. Speech Signal Process., vol.36, no.5, pp.642-650, 1988.
- [11] P.J. Huber, Robust Statistics, Wiley, New York, 1981.
- [12] V. Vapnik, Statistical Learning Theory, Wiley, New

York, 1998.

- [13] 鹿野清宏, 伊藤克巨, 河原達也, 武田一哉, 山本幹雄(編), 音声認識システム, オーム社, 東京, 2001.
- [14] 後藤真孝, 橋口博樹, 西村拓一, 岡 隆一, "RWC 研究用音楽データベース: 研究目的で利用可能な著作権処理済み楽曲・楽器音データベース;" 情処学論, vol.45, no.3, pp.728-738, 2004.

付 録

ここでは、 $f_3(z)$ に点 b で接し、 $f_3(z)$ を下回らない二次関数の中で二次の係数が最小のものを求めたい。まず、 $|b| \geq 2\beta$ では式 (7) の頂点は β より下回らないため、式 (7) が上記要件を満たす。次に、 $\beta + c \leq |b| < 2\beta$ では、点 b で $|z|$ に接し、かつ、頂点が β に接する二次関数が要件を満たすので、

$$k(b-l)^2 + \beta = |b|, \quad 2k(b-l) = \text{sign}(b) \quad (\text{A.1})$$

となる k, l を連立して解けばよく、これを解けば

$$k = \frac{1}{4(|b| - \beta)}, \quad l = 2\text{sign}(b)\beta - b \quad (\text{A.2})$$

を得る。次に、 $\beta - c \leq |b| < \beta + c$ においては、 $f_3(z)$ が $\beta - c \leq z < \beta + c$ において二次関数であるため、

$$\frac{(z - |\beta - c|)^2}{4c} + \beta \quad (\text{A.3})$$

が要件を満たす。最後に、 $0 \leq |b| < \beta - c$ においては、頂点が (b, β) であり、かつ、 $\text{sign}(b)z$ に接する二次関数が要件を満たすので、

$$k(z-b)^2 + \beta = \text{sign}(b)z \quad (\text{A.4})$$

が重解をもつように k を決めればよく、これを解けば

$$k = \frac{1}{4(\beta - |b|)} \quad (\text{A.5})$$

を得る。

(平成 20 年 2 月 18 日受付, 6 月 11 日再受付)



亀岡 弘和 (正員)

2002 東大・工・計数卒。2004 同大大学院・情報理工システム情報修士課程了。2007 同博士課程了。情報理工博。同年日本電信電話(株)入社。以来、音楽、音声を対象とした音響信号処理の研究に従事。現在、NTT コミュニケーション科学基礎研究所勤務。第 20 回電気通信普及財団テレコムシステム技術学生賞、Best Student Paper Award Finalist at ICASSP2005、情報処理学会平 17 年度山下記念研究賞、日本音響学会第 2 回独創研究奨励賞板倉記念、平 18 年度東京大学大学院情報理工学系研究科研究科長賞、第 1 回 IEEE Signal Processing Society Japan Chapter Student Paper Award、日本音響学会第 25 回栗屋潔学術奨励賞を受賞。情報処理学会、日本音響学会、IEEE 各会員。



鎌本 優 (正員)

2003 慶大・理工・物理情報卒。2005 東大大学院・情報理工・システム情報修士課程了。同年日本電信電話(株)入社。以来、信号処理、情報理論に関する研究に従事。現在、NTT コミュニケーション科学基礎研究所勤務。平成 17 年度情報処理学会論文賞第 21 回電気通信普及財団テレコムシステム技術学生賞、第 22 回電気通信普及財団テレコムシステム技術賞(奨励賞)。日本音響学会、情報処理学会、情報理論とその応用学会、IEEE 各会員。



原田 登 (正員)

1995 九工大・情報工卒。1997 同大大学院・情報シス修士課程了。同年日本電信電話(株)入社。以来、音声・楽音の高効率符号化、音声情報処理、VoIP の研究に従事。現在、NTT コミュニケーション科学基礎研究所研究主任。第 22 回電気通信普及財団テレコムシステム技術賞(奨励賞)。日本音響学会、Audio Engineering Society、IEEE 各会員。



守谷 健弘 (正員)

1978 東大・工・計数卒。1980 同大大学院・工・計数修士課程了。同年日本電信電話公社(現 NTT)入社。以来、音声音響信号の符号化の研究開発、標準化に従事。1989 工学博士。同年 AT&T ベル研究所客員研究員。現在、NTT コミュニケーション科学基礎研究所特別研究室長、NTT フェロー。本会より学術奨励賞、論文賞、業績賞・小林記念特別賞受賞。IEEE(フェロー)、日本音響学会、情報処理学会、情報理論とその応用学会、Audio Engineering Society 各会員。