ISCA Archive
http://www.isca-speech.org/archive

6th International Conference on Spoken Language Processing (ICSLP 2000)
Beijing, China
October 16-20, 2000

RE-ESTIMATION OF LPC COEFFICIENTS IN THE SENSE OF L_∞ CRITERION

Zhang Sen+, Katsuhiko Shirai
Depart. of Information & Computer Science, School of Science & Engineering,
Waseda University, Tokyo, JAPAN
Email: zhangsen@yahoo.com

ABSTRACT

Now the generally used approaches such as auto-correlation method and covariance method for estimating LPC coefficients are to solve a set of linear equations by using of Levinson-Durbin recursion or lattice formulations, but the LPC coefficients computed are best only in the sense of L_2 criterion[1]. For speech processing, L_∞ criterion is a more suitable measure metric. The idea of our approach is: from the initial values of LPC coefficients, the residual errors could be reduced step by step by using Least Squares process iteratively until the LPC coefficients are approximately best in the sense of L_∞ criterion. Furthermore, this approach could be applied to other problems for estimating some parameters.

1. INTRODUCTION

There are two important methods to extract short-term speech features: one is Linear Predictive (LP) and other is Discrete Fourier Transform (DFT). The DFT is superior to represent the detailed structure of the speech waveform. On the other hand, the LP analysis can well represent the smoothed envelope of the spectrum and rather robust to noise. So most speech recognition systems selected LPC coefficients or the derived ones such as cepstrum or mel-cepstrum, as acoustic features.

Two methods have been used for estimating LPC coefficients, one is the auto-correlation method and the other is covariance method. Both of them have to solve a set of linear equations. Since the linear equations have a special structure, some effective computing algorithms have been developed, such as Levinson-Durbin recursion and lattice formulations, but the LPC coefficients computed are best only in the sense of L_2 criterion. For speech processing, L_∞ criterion is a more suitable metric[2][3]. Our approach is to refine LPC coefficients in the sense of L_∞ criterion. Our experiments showed that the residual errors could be reduced step by step by using Least Squares algorithms iteratively until the LPC coefficients are approximately best in the sense of L_∞ criterion. Furthermore, this approach could be applied to other problems for estimating some parameters.

2. LPC ANALYSIS

Let $\{x_t\}$ stand for the speech sample sequence, then the p dimensional LPC analysis is to find a set of coefficients $\{a_i \mid 1 \leqq i \leqq p\}$ that can minimize predictive errors in some sense such as L_2 criterion. That means the speech sample x_t can be represented by the past p samples with linear relation as,

$$a_1 x_{t-1} + \ldots + a_p x_{t-p} = x_t + \varepsilon_t$$

where $\{\varepsilon_t\}$ is a probabilistic variable with Gaussian distribution, and its averaged value is zero and its covariance is σ^2. The predictive errors are required minimum in the sense of L_2 criterion.

Now there are two most widely used methods for estimating the Linear Prediction Coefficients, i.e., Autocorrelation Method and Covariance Method. Note that the above methods can give LP coefficients which are only best or optimal in the sense of least squares metric or L_2 measure. We hope to get a set of LP coefficients which can minimize the error in sense of L_∞, which can assure the maximum errors of all predictive samples to be minimum.

3. RE-ESTIMATION OF LPC COEFFICIENTS

The traditional methods only consider the overall error of Linear Prediction in L_2 metric, but how about the errors of each predicted values and how to control them? We cannot get much information from the above Linear Prediction methods. Our aims are to solve such problems.

For the purpose of following description, let's introduce some definations first. Given a speech samples sequence $\{x(n)\}$, and a set of LPC coefficients $\{a_i \mid 1 \leqq i \leqq p\}$ in some sense, the predicted error E_n of sample $x(n)$ is defined as follows:

10.21437/ICSLP.2000-183

$$E_n = |x(n) - \underline{x}(n)|$$

Later we often refer E_n as the error at point x(n). If the errors at some points, such as E_k, satisfy the following condition:

$$E_k = \max_n |x(n) - \underline{x}(n)|$$

Then we refer E_k as the maximum error and x(k) as the maximum error point, or MEP for short. The Linear Prediction error in L_∞ metric is defined as follows:

$$E = \min_{\{a_i\}} \max_n |x(n) - \underline{x}(n)|$$

where $\underline{x}(n)$ is the predicted value of *x*(*n*), and $\{a_i\}$denote a set of LPC coefficients, the limits of n is within a frame. If a set of LPC coefficients are best in L_∞ metric, it must satisfy the above error conditon E. From the defination, it is clear that the predicted errors at all points would be uniformly minimum in L_∞ metric.

Now we present a method to re-estimate LPC coefficients. The idea is as follows: take the original LPC coefficients as initial values, then by reducing the maximum error at MEPs, we can get an new set of LPC coefficients. Take the new LPC coefficients as initial values, repeat this procedure at most p (LPC order) times, we can get LPC coefficients which are best in L_∞ metric.

Let $\{a_i^k\}$ denote the LPC coefficients computed after k-th iteration procedure, and E_i^k denote the predicted error at point x(i) with respect to the LPC coefficients $\{a_i^k\}$. Let $\{a_i^0\}$ denote the LPC coefficients computed by auto-correlation or covariance method. We now discuss our method in single-MEP case and poly-MEPs case respectively.

3.1 Single-MEP Case

In this case, there exists only one MEP. Suppose the MEP is x(k), $\underline{x}(k)$ is the predicted value of x(k), p is the order of Linear Prediction, $\{a_i^0\}$ is the original LPC coefficients, then

$$\underline{x}(k) = a_1^0 x(k-1) + \Lambda + a_p^0 x(k-p)$$

Let $e_k = x(k) - \underline{x}(k)$, then $E_k = |e_k|$

Suppose $x(k\text{-}l) \neq 0$, and it satisfies the following condition:

$$|x(k-l)| = \max\{|x(k-1)|, \ldots, |x(k-p)|\}$$

Set $\varepsilon\, x(k\text{-}l) = -\mathrm{sgn}(e_k)$ (3.1.1)

where sgn(.) is a sign function which means:

$$\mathrm{sgn}(x) = \begin{cases} 1 & if \quad x > 0 \\ -1 & if \quad x < 0 \\ 0 & if \quad x = 0 \end{cases}$$

and ε is a parameter to be determined, here ε is refered as direction parameter. From formulation (3.1.1) ε can be computed, $\varepsilon = -\mathrm{sgn}(e_k)/\, x(k\text{-}l)$.

Next we will determine parameter λ which is supposed greater than zero and refered as delt parameter. To decrease the maximum predicted error at point x(k), λ should meet the following condition:

$$|e_k + \lambda\varepsilon x(k-l)| = \max_{i \neq k} |e_i + \lambda\varepsilon x(i-l)|$$

Note that the formulation (3.1.1), the above equation can be rearranged as the following form:

$$E_k = \lambda + \max_{i \neq k} |e_i + \lambda\varepsilon x(i-l)|$$

Since we have

$$\max_{i \neq k} |e_i + \lambda\varepsilon x(i-l)| \\ \leq \max_{i \neq k} E_i + \lambda \max_{i \neq k} |\varepsilon x(i-l)|$$

Hence λ satisfies the following inequality,

$$\lambda \geq \left(E_k - \max_{i \neq k} E_i\right) \Big/ \left(1 + \max_{i \neq k} |\varepsilon x(i-l)|\right)$$

Note that the right of above inequality is greater than zero and smaller than E_k, so we choose the lower bound as the value of parameter λ ,

$$\lambda = \left(E_k - \max_{i \neq k} E_i\right) \Big/ \left(1 + \max_{i \neq k} |\varepsilon x(i-l)|\right) \qquad (3.1.2)$$

Considering the above formulations (3.1.1) and (3.1.2), the parameters ε and λ have been determined, so we can compute a set of new LPC coefficients by the following steps,

Step 1: For i=1 to p, but except l

$a_i^0 \rightarrow a_i^1$

Step 2: Set $a_l^0 + \lambda\,\varepsilon \rightarrow a_l^1$

Then $\{a_i^1, 1 \leq i \leq p\}$ is the new LPC coefficients which can decrease the maximum error at MEP, and the descending magnitude is λ. The above processing of single-MEP case results in Poly-MEPs.

3.2 Poly-MEPs Case

Now let's consider how to reduce the maximum errors in Poly-MEPs case. Suppose $\{a_i^0, 1 \leq i \leq p\}$ is a set of LPC coefficients, and the coresponding MEPs are $x(i_1), \ldots, x(i_k)$, 1<k<p. We hope to find a set of new LPC coefficients $\{a_i^1\}$ which meet the

following conditions:

$$E^1_{i_1} = \cdots = E^1_{i_k} = \max_{i \neq i_1,\ldots,i_k} E^1_i$$

According to the Linear Prediction defination, we have the following equations with respect to the MEPs $x(i_1), \ldots, x(i_k)$,

$$\begin{cases} \underline{x}(i_1) = a_1 x(i_1 - 1) + \cdots + a_p x(i_1 - p) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \underline{x}(i_k) = a_1 x(i_k - 1) + \cdots + a_p x(i_k - p) \end{cases}$$

where $\underline{x}(i_1), \ldots, \underline{x}(i_k)$ are predicted values of $x(i_1), \ldots, x(i_k)$ respectively. We can express this in matrix form:

$$\underline{X} = C_{kxp} A$$

where

$$\underline{X} = \begin{bmatrix} \underline{x}(i_1) \\ \vdots \\ \underline{x}(i_k) \end{bmatrix} \qquad A = \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix}$$

and ,

$$C_{kxp} = \begin{bmatrix} x(i_1 - 1) & \cdots & x(i_1 - p) \\ \cdots & \cdots & \cdots \\ x(i_k - 1) & \cdots & x(i_k - p) \end{bmatrix}$$

Suppose we can choose a sub-matrix C_{kxk} from C_{kxp} (Note 1<k<p),

$$C_{kxk} = \begin{bmatrix} x(i_1 - j_1) & \cdots & x(i_1 - j_k) \\ \cdots & \cdots & \cdots \\ x(i_k - j_1) & \cdots & x(i_k - j_k) \end{bmatrix}$$

And the inverse matrix of C_{kxk} exists, then we can decrease the maximum errors at MEPs further by method which is analogous to that in single-MEP case.

Let' s first determine the direction vector ε which is a k-dimensional vector, and then discuss how to compute the delt parameter λ which is a positive real number.

Considering the sub-matrix C_{kxk}, k-dimensional vector ε can be expressed as follows:

$$\varepsilon = \begin{bmatrix} \varepsilon_{j_1} \\ \vdots \\ \varepsilon_{j_k} \end{bmatrix}$$

Another k-dimensional vector S (which is refered as sign vector) is composed of 1 or –1, S can be expressed as follows:

$$S = \begin{bmatrix} -\operatorname{sgn}(e_{i_1}) \\ \vdots \\ -\operatorname{sgn}(e_{i_k}) \end{bmatrix}$$

Since C_{kxk} has an inverse matrix, the direction vector ε can be computed by the next equation:

$$C_{kxk}\, \varepsilon = S \qquad (3.2.1)$$

As we illustrated above, to obtain direction vector ε, we have to solve a kxk equations. Since the order k of such equations is less than prediction order p (usually less than 16), the additional computation load is not heavy.

Next we will compute the delt parameter λ which is a positive real number and indicates the error descending magnitude. Let K denote the number set of $\{i_1, \ldots, i_k\}$. The principle of determining λ is as following:

$$\max_{n \in K} |E_n - \lambda| = \max_{i \notin K} \left| e_i + \lambda \sum_{l=1}^{k} \varepsilon_{j_l} x(i - j_l) \right|$$

So λ must meet follows:

$$\lambda \geq \frac{\max_{n \in K} E_n - \max_{i \notin K} E_i}{1 + \max_{i \notin K} \left| \sum_{l=1}^{k} \varepsilon_{j_l} x(i - j_l) \right|}$$

Note that the right of above inequality is greater than zero, so we choose the lower bound as the value of λ ,

$$\lambda = \frac{\max_{n \in K} E_n - \max_{i \notin K} E_i}{1 + \max_{i \notin K} \left| \sum_{l=1}^{k} \varepsilon_{j_l} x(i - j_l) \right|} \qquad (3.2.2)$$

By the above formulations (3.2.1) and (3.2.2), the direction vector ε and parameter λ have been determined, now we can compute a set of new LPC coefficients by the following steps,

Step 1: For i=1 to p, except that $i \in K$
$a_i^0 \rightarrow a_i^1$
Step 2: compute vector ε .
Step 3: compute parameter λ .
Step 4: For all $i \in K$
Set $a_i^0 + \lambda \varepsilon_i \rightarrow a_i^1$

Then { $a_i^1, 1 \leq i \leq p$} is the new LPC coefficients which can decrease the maximum errors at MEPs, and the descending magnitude is λ.

For a given speech sample sequence {x(n)}, the re-estimation procedure of LPC coefficients is as follows:

Step 1: compute the LPC coefficients frame by frame in L_2 metric by auto-correlation or covariance method.
Step 2: for the computed LPC coefficients of a frame, check the MEPs. If only one MEP, apply the algorithm provided in Single-MEP case to re-compute LPC coefficients; if there are at least p (prediction order) MEPs, output the newest LPC coefficients as required and goto Step 3; otherwise, apply the algorithm provided in Poly-MEPs case to re-compute LPC coefficients.
Step 3: re-compute the LPC coefficients of next frame by Step 2 until all the frames have been processed.
Step 4: output the maximum errors of each frame and program stops.

The above discussion presented the main features and procedures of our approach to re-estimate LPC coefficients in L_∞ metric.

4. EXAMPLES AND RESULTS

To test the effectiveness of the algorithms for re-estimation of LPC coefficients, let's take an example. For instance, we will compute and reestimate the LPC coefficients of vowel /a/ frame by frame, and compare the predicted errors before and after reestimation.

Main information about vowel /a/: sampling rate is 16KHz, 16 bits, total samples 7397, duration 0.462 second. The frame length is 20ms or 320 samples, so the 7397 samples can be divided into 23 frames (we did not let the frames overlap). The prediction order is 16. The original speech waveform of /a/ is as follows:

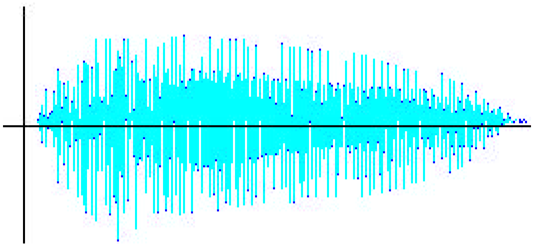

Figure 4.1 waveform of vowel /a/

By using autocorrelation method, we could compute the LPC coefficients of the first frame of /a/ as:
a^0 = (1.03640, 0.00357, 0.06287, -0.15509, 0.12480, -0.24308, 0.04208, 0.03400, 0.09100, 0.00430, -0.11282, 0.01295, 0.08555, 0.03283, -0.06486, 0.01465).
The corresponding overall errors in L_2 metric is E_2 = 0.00210
and the error at MEP is
E_∞ = 0.01746

After the application of Single-MEP method, the LPC coefficients of the first frame of /a/ becomes as:
a^1 = (0.95881, 0.00357, 0.06287, -0.15509, 0.12480, -0.24308, 0.04208, 0.03400, 0.09100, 0.00430, -0.11282, 0.01295, 0.08555, 0.03283, -0.06486, 0.01465).
The corresponding error at MEP is
E_∞ = 0.01630

After the application of Poly-MEPs method, the LPC coefficients of the first frame of /a/ becomes as:
a^{15} = (0.07546, -0.18726, 0.41583, 0.00626, -0.15000, -0.43083, 0.16563, 0.46284, 0.12172, 0.18506, -0.12460, 0.07848, -0.31432, 0.34452, 0.26376, 0.01465).
The corresponding error at MEP is
E_∞ = 0.00776

The errors corresponding to the 15 times of LPC coefficients reestimation are as follows respectively:

0.01746, 0.01630, 0.01585, 0.01407, 0.01355, 0.01331, 0.01329, 0.01229, 0.01189, 0.01167, 0.01143, 0.01137, 0.01123, 0.01087, 0.01068, 0.00776.

The above errors could be illustrated by the following figure 4.2. The experiments show that the maximum errors in L_∞ **metric decreased more than 50%** after reestimation of LPC coefficients.

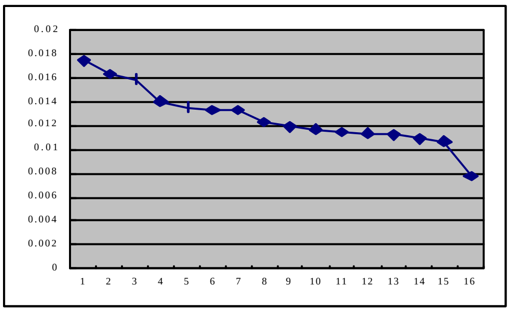

Figure 4.2 Error descending by reestimating of LPC coefficients of a frame

5. ACKNOWLEDGEMENT

Thanks for the encouragement of Prof. K.Shirai and the support of Prof. H. Fujisaki, and the help from my colleagues in Shirai Lab. in Waseda University. The discussion with them widened my knowledge and inspired my creative desire very much.

REFRENCES

[1] L.R. Rabiner and B.H. Juang, Fundamentals of Speech Recognition, Prentice-Hall International, Inc. 1993
[2] E. Denoel and Jean-Philippe Solvay, Linear prediction of speech with a least absolute error criterion, IEEE Transactions on ASSP, December,1985
[3] Yingbo Hua and T.K. Sarkar, A perturbation property of the TLS-LP Method, IEEE Transactions on ASSP, Nov., 1990