

Chapter 2

Overview of Standard Methods

Because of correlation between the speech signal samples, it is possible to predict the values of the speech signal based on a linear combination of the previous p samples, with an acceptable level of the total prediction error. The predicted value of signal in the n th moment of time $s(n)$ is determined by the relation

$$\hat{s}(n) = - \sum_{i=1}^p a_i s(n-i), \quad (2.1)$$

where $a_i, i = 1, 2, \dots, p$ are the coefficients of linear prediction, p is the order of the prediction model, and $s(n-i)$ signal samples in the $(n-i)$ th moment of time.

The prediction error $e(n)$ is the difference between the real and the predicted value of the signal samples, i.e.,

$$e(n) = s(n) - \hat{s}(n) = s(n) + \sum_{i=1}^p a_i s(n-i) = \sum_{i=0}^p a_i s(n-i), \quad a_0 = 1. \quad (2.2)$$

The relations (2.1) and (2.2) describe the model of linear prediction in time domain. By applying the z transformation to Eqs. (2.1) and (2.2) we obtain

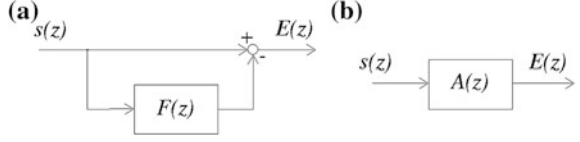
$$\hat{S}(z) = F(z)S(z) \quad (2.3)$$

where

$$F(z) = - \sum_{i=1}^p a_i z^{-i}, \quad (2.4)$$

$$E(z) = S(z)[1 - F(z)] = S(z)A(z), \quad (2.5)$$

Fig. 2.1 Model of linear prediction in z -domain
a direct, **b** inverse filter



while

$$A(z) = 1 + \sum_{i=1}^p a_i z^{-i} = 1 - F(z). \quad (2.6)$$

Equations (2.3)–(2.6) describe the model of linear prediction in the z -domain. The polynomial $A(z)$ has zeroes only and it is called the inverse filter, Fig. 2.1.

The set of parameters $a_i, i = 1, 2, \dots, p$, fully determining the filter $A(z)$, is a set of unknown linear prediction parameters to be determined. These coefficients can be obtained by minimizing the mean square prediction error at a predefined interval that is called the minimization interval.

The mean square error E can be expressed as

$$E = \sum_{n=n_0}^{n_1} e^2(n) = \sum_{n=n_0}^{n_1} [s(n) - \hat{s}(n)]^2 = \sum_{n=n_0}^{n_1} \sum_{i=0}^p \sum_{j=0}^p a_i s(n-i) s(n-j) a_j, \quad (2.7)$$

where E is the total mean square error, and $[n_0 - n_1]$ is the minimization interval. By introducing the notation

$$\phi_{ij} = \sum_{n=n_0}^{n_1} s(n-i) s(n-j), \quad (2.8)$$

the expression for the total mean square error of prediction can be written in the form

$$E = \sum_{i=0}^p \sum_{j=0}^p a_i \phi_{ij} a_j. \quad (2.9)$$

From the necessary condition for the minimum of the adopted criterion

$$\frac{\partial E}{\partial a_k} = 2 \sum_{i=0}^p a_i \phi_{ik} = 0 \quad (2.10)$$

the system of equations follows:

$$\sum_{i=1}^p a_i \phi_{ik} = -\phi_{0k}, \quad k = 1, 2, \dots, p, \quad (2.11)$$

from which one determines the linear prediction parameters. Depending on the manner of choice of the minimization interval of the mean square error, we distinguish the next methods of linear prediction:

- Autocorrelation method
- Covariant method.

2.1 Autocorrelation Method

Theoretically, the minimization interval of the mean square prediction error for the autocorrelation method is infinite, but using an adequate window function it is practically limited to the interval $0 \leq n \leq N - 1$. In that case, the signal covariance reduces to autocorrelation, i.e.,

$$\begin{aligned} \phi_{ij} &= \sum_{n=-\infty}^{\infty} s(n-i)s(n-j) = \sum_{n=-\infty}^{\infty} s(n)s(n+|i-j|) \\ &= \sum_{n=0}^{N-1-|i-j|} s(n)s(n+|i-j|) = R(|i-j|), \end{aligned} \quad (2.12)$$

where $R(|i-j|)$ is the autocorrelation function of the signal. The system of Eqs. (2.11) is transformed into

$$\sum_{i=1}^p a_i R(|i-j|) = -R(j), \quad j = 1, 2, \dots, p. \quad (2.13)$$

The relation (2.13) defines the autocorrelation method of linear prediction. In the matrix form, it is written as

$$\begin{bmatrix} R_0 & R_1 & R_2 & R_3 & \dots & R_{p-1} \\ R_1 & R_0 & R_1 & R_2 & \dots & R_{p-2} \\ R_2 & R_1 & R_0 & R_1 & \dots & R_{p-3} \\ R_3 & R_2 & R_1 & R_0 & \dots & R_{p-4} \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ R_{p-1} & R_{p-2} & R_{p-3} & R_{p-4} & \dots & R_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ \vdots \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \\ \vdots \\ \vdots \\ R_p \end{bmatrix} \quad (2.14)$$

where we introduced a short notation $R_i \equiv R(i)$. Here R_i is a Toeplitz matrix, since it is symmetrical and has equal elements on its diagonals. These properties of the matrix are used for the development of efficient algorithms for the solution of the system of Eq. (2.14). Among the most efficient ones is the Durbin algorithm. This algorithm is recursive and is given by the following relations:

$$E_0 = R(0), \quad a_0 = 1 \quad (2.15)$$

$$K_i = - \left[R(i) + \sum_{j=1}^{i-1} a_j^{(i-1)} R(i-j) \right] \frac{1}{E_{i-1}} \quad (2.16)$$

$$a_i^{(i)} = K_i \quad (2.17)$$

$$a_j^{(i)} = a_j^{(i-1)} + K_i a_{i-j}^{(i-1)}, \quad j = 1, 2, \dots, i-1 \quad (2.18)$$

$$E_i = (1 - K_i^2) E_{i-1} \quad (2.19)$$

where E_i represents the prediction error in the i th recursion step. Expressions (2.15)–(2.19) are solved recursively for $i = 1, 2, \dots, p$. The final solution has the form

$$a_j = a_j^{(p)}, \quad j = 1, 2, \dots, p. \quad (2.20)$$

Normalized coefficients of autocorrelation $r(i) = R(i)/R(0)$ can be used in the algorithm, so that the prediction error in the i th step of recursion is also normalized

$$V_i = \frac{E_i}{R(0)} = 1 - \sum_{k=1}^i a_k r(k). \quad (2.21)$$

The coefficients $\{K_i\}$, $i = 1, 2, \dots, p$ are called the reflection coefficients. In stable analysis, where the estimated transfer functions of the inverse filter $A(z)$ are stable (the corresponding zeroes of the transfer functions are located within the unit circle $|z| = 1$), the absolute values of these coefficients are lower than one

$$|K_i| < 1 \quad i = 1, 2, \dots, p. \quad (2.22)$$

2.2 Covariant Method

In the case of covariant method one predefines a finite interval on which minimization of the mean square prediction error is performed. It is determined by the system of equations

$$\sum_{i=1}^p a_i \phi_{ij} = -\phi_{0j}, \quad 1 \leq j \leq p, \quad (2.23)$$

where $\phi_{ij} = \sum_{n=p}^{N-1} s(n-i)s(n-j)$ is the estimation of the covariant function of the signal. In the matrix notation Eq. (2.23) has the form

$$\begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} & \cdots & \phi_{1p} \\ \phi_{21} & \phi_{22} & \phi_{23} & \cdots & \phi_{2p} \\ \phi_{31} & \phi_{32} & \phi_{33} & \cdots & \phi_{3p} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \phi_{p1} & \phi_{p2} & \phi_{p3} & \cdots & \phi_{pp} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} \phi_{01} \\ \phi_{02} \\ \phi_{03} \\ \vdots \\ \phi_{0p} \end{bmatrix} \quad (2.24)$$

i.e.,

$$\Phi a = \Psi, \quad (2.25)$$

where

- Φ —a matrix with dimensions $p \times p$ whose elements are the coefficients of the estimation of covariations ϕ_{ij}
- a —column vector of the coefficients of linear prediction $a_i, i = 1, 2, \dots, p$
- Ψ —column vector with elements $\{\Psi_i\} = \{-\phi_{0j}\}, j = i = 1, 2, \dots, p$

The matrix Φ is symmetric, but is not of Toeplitz type. The values of the matrix elements ϕ_{ij} approach the autocorrelation coefficients R_i for longer minimization intervals.

Equation (2.25) is most often solved using the procedure of Cholesky decomposition. A comparative presentation of the autocorrelation and covariant method is given in Table 2.1. One can observe from Table 2.1 that the basic advantage of the autocorrelation method is in lower computational complexity and in the fact that it theoretically guarantees the stability of the analysis, i.e., the stability of the resulting transfer function of the filter. The advantage of the covariant method is in its better estimation of parameters, especially on shorter segments of analysis. The basic disadvantage of the covariant method is that it does not guarantee the stability of the analysis.

The terms of stationary and nonstationary formulation of linear prediction are shown in Fig. 2.2. In the case of the autocorrelation method (stationary formulation) the minimization interval is theoretically infinite. Practically, since the duration of the signal is limited to the interval $[0, N - 1]$, it follows that the length of the minimization interval is $N + p - 1$, Fig. 2.2. At the beginning of the minimization interval the real values of the signal are predicted based on samples that are different from zero. The largest prediction errors are obtained at the ends of the minimization interval, which leads to a periodicity of the prediction error which is not a consequence of the signal nature. To decrease this effect, it is mandatory to use a window function during analysis. The most often used is Hamming window, defined by

Table 2.1 Comparative presentation of the autocorrelation and the covariant method

Method	Autocorrelation	Covariant
Criterion	Durbin algorithm	Cholesky decomposition
Length of the segment for analysis	N_1	$N_2 < N_1$
Stability	Theoretically guaranteed In practical situations instabilities occur because of calculations with finite word length	Not guaranteed Stability is tested based on the condition $ K_i < 1$
The number of multiplication/division operations	Forming $R(i) : N_1 p$ Solution of the equations: p	Forming $\phi(i, j) : N_2 p$ Solution of the equations: p
Use of window	Mandatory	Not mandatory
Order of the model	Does not depend on the method, but on the goal of the analysis. In the systems for speech analysis/synthesis the number is between 8 and 12. Informativeness of prediction parameters decreases with an increase of their index	
Computational complexity	Lower	Higher
Estimation of parameters for the same conditions of analysis (formants, formant widths, spectrum)	Worse	Better, especially on shorter analytical segments. On longer analytical segments the results are approximately equal

$$w_H(i) = \begin{cases} 0.54 + 0.46 \cos(2\pi i/N) & \text{for } i = 1, \dots, N \\ 0 & \text{other values of } i \end{cases} \quad (2.26)$$

In the case of covariant method, the length of the minimization intervals is theoretically and practically equal to $N - p$. The analysis interval is expanded with p more samples outside the minimization interval and the effect of ends does not appear as in the case of the autocorrelation method. Because of that the covariant method does not require the use of a window function. The choice of minimization interval determines the stability of the inverse filter $A(z)$. As previously mentioned, the autocorrelation method, because of the infinite minimization interval, theoretically guarantees stability, while the covariant method, because of the finite minimization interval, theoretically does not guarantee it.

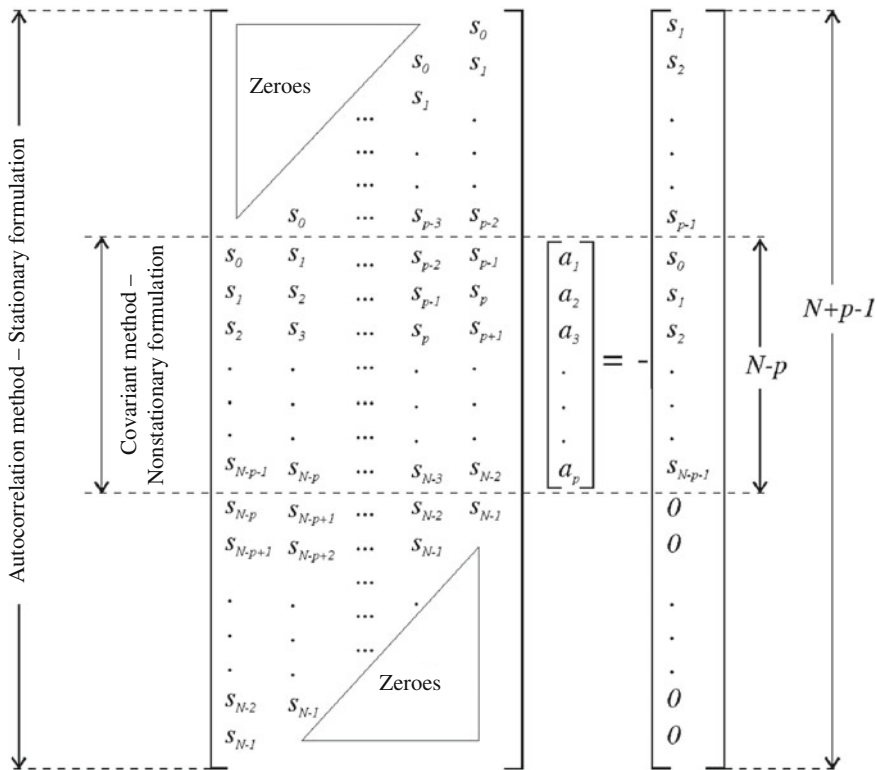


Fig. 2.2 Stationary and nonstationary formulation of linear prediction

2.3 Forward and Backward Prediction

Contrary to the autocorrelation and the covariant method, where each signal sample is predicted with only one predictor of p th order, in the case of partial correlation method each signal sample is predicted by p predictors forward and by p predictors backward, Fig. 2.3.

Let us denote by $f_m(n)$, the error of prediction of a signal sample $s(n)$ forward with a predictor of m th order, $m = 1, 2, \dots, p$, and by $b_m(n)$ the error of prediction of a signal sample $s(n-m-1)$ backward with a predictor of m th order, $m = 1, 2, \dots, p$. In that case according to definition

$$f_m(n) = s(n) - \left[- \sum_{i=1}^m a_i^{(m)} s(n-i) \right] = \sum_{i=0}^m a_i^{(m)} s(n-i), \quad (2.27)$$

i.e.,

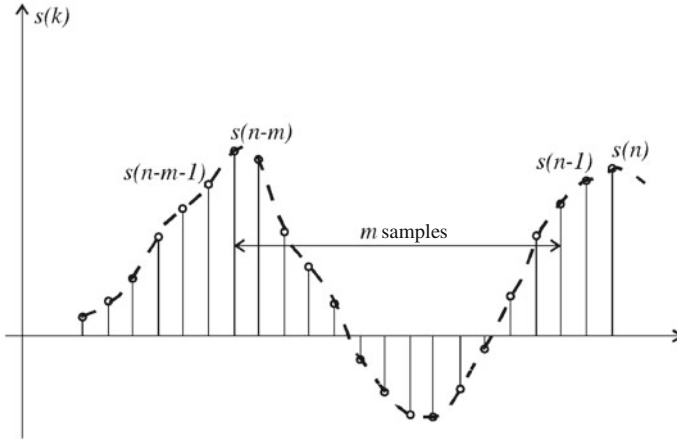


Fig. 2.3 Forward and backward prediction

$$b_m(n) = s(n-m-1) - \left[- \sum_{i=1}^m b_i^{(m)} s(n-i) \right] = \sum_{i=1}^{m+1} b_i^{(m)} s(n-i), \quad (2.28)$$

where $a_0^{(m)} = 1$ and $b_{m+1}^{(m)} = 1$. Since m changes from 1 to p , for each signal sample one obtains p prediction errors forward and p prediction errors backward. The errors $f_m(n)$ and $b_m(n)$ can be observed as signal output from filters with characteristics $A_m(z)$ and $B_m(z)$ to whose input a signal $s(n)$ is brought. The transfer functions of these filters can be written as

$$A_m(z) = \sum_{i=0}^m a_i^{(m)} z^{-i}, \quad \text{where } a_0^{(m)} = 1, \quad (2.29)$$

$$B_m(z) = \sum_{i=1}^{m+1} b_i^{(m)} z^{-i}, \quad \text{where } b_{m+1}^{(m)} = 1. \quad (2.30)$$

The mean square errors of forward and backward prediction F_m and B_m , respectively, for an m th order predictor are

$$F_m = \sum_{n=n_0}^{n_1} f_m^2(n), \quad (2.31)$$

$$B_m = \sum_{n=n_0}^{n_1} b_m^2(n). \quad (2.32)$$

The minimization interval $[n_0, n_1]$ can be theoretically finite or infinite, as in the case of the autocorrelation and the covariant method. The prediction parameters can be obtained by minimization of the total prediction errors F_m and B_m from the conditions

$$\frac{\partial F_m}{\partial a_i(m)} = 0, \quad \frac{\partial B_m}{\partial b_i(m)} = 0. \quad (2.33)$$

Since m assumes values from 1 to p , this method can be also understood as a step-by-step procedure of prediction error minimization.

There are different possibilities for the choice of minimization criterion. One can minimize F_m only or B_m only, or some of their combinations, for instance their sum. In the case of stationary formulation, i.e., an infinite minimization interval, all criteria give the same result. In the case of a finite minimization interval, for nonstationary formulation, this may not be the case.

2.4 Lattice Filter

Starting from the previously defined term of forward and backward prediction, the chosen interval and the minimization criterion, one arrives at the inverse filter $A(z)$ in the form of a lattice filter, Fig. 2.4.

Each filter cell is described by pair equations in time domain

$$f_{m+1}(n) = f_m(n) + K_{m+1}b_m(n-1), \quad (2.34)$$

$$b_{m+1}(n) = K_{m+1}f_m(n) + b_m(n-1), \quad (2.35)$$

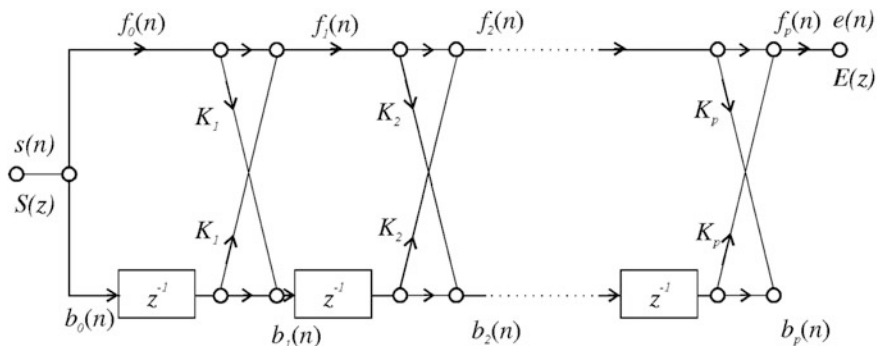


Fig. 2.4 Lattice-type structure of filter $A(z)$

or in the z -domain

$$F_{m+1}(z) = F_m(z) + z^{-1}K_{m+1}B_m(z), \quad (2.36)$$

$$B_{m+1}(z) = K_{m+1}F_m(z) + z^{-1}B_m(z). \quad (2.37)$$

The lattice filter structure is invariant regarding the chosen criterion and the minimization criterion. The differences in the choice of the interval and the minimization criterion influence the value of the coefficient K_m and, naturally, the stability of the analysis, i.e., the resulting estimated filter.

One can see from the above that instead of using a set of prediction coefficients $\{a_i\}$, $i = 1, 2, \dots, p$ the structure of the inverse filter $A(z)$ can be unambiguously described by a set of parameters $\{K_m\}$, $m = 1, 2, \dots, p$. These parameters are called the partial correlation coefficients. The partial correlation coefficients $\{K_m\}$, $m = 1, 2, \dots, p$ are unambiguously connected with the prediction parameters $\{a_j(m)\}$ $j = 1, 2, \dots, m$; $m = 1, 2, \dots, p$ by the relations

$$a_m^{(m)} = K_m, \quad (2.38)$$

$$a_j^{(m)} = a_j^{(m-1)} + K_m a_{m-j}^{(m-1)}, \quad j = 1, 2, \dots, m-1. \quad (2.39)$$

The coefficients $\{a_i(m)\}$ $i = 1, 2, \dots, m$ obtained after each recursion describe an m th order predictor, i.e., the structure of a filter $A_m(z)$. For $m = p$, the coefficients $a_j = a_j^{(p)}$, $j = 1, 2, \dots, p$ describe the final structure of the inverse filter $A(z)$.

The normalized autocorrelation function $r(i)$ of the signal $s(i)$ is defined by

$$r(i) = \frac{E\{s(n)s(n-i)\}}{[E\{s^2(n)\}E\{s^2(n-i)\}]^{1/2}} = \frac{R(i)}{R(0)}. \quad (2.40)$$

There is a high degree of redundancy in thus defined autocorrelation coefficients, since these coefficients are statistically highly correlated.

The partial correlation coefficients are defined as a cross-correlation measure of the error signal of forward and backward prediction using an i th order predictor

$$K_i = \frac{\sum_{n=-\infty}^{\infty} f_{i-1}(n)b_{i-1}(n)}{\left(\sum_{n=-\infty}^{\infty} f_{i-1}^2(n) \sum_{n=-\infty}^{\infty} b_{i-1}^2(n)\right)^{1/2}}, \quad i = 1, 2, \dots, p. \quad (2.41)$$

The redundancy of these coefficients is decreased because the signal samples of prediction error are much less correlated than the signal samples $s(i)$. Let us have (Table 2.2).

Now it is possible to introduce the definition of the partial correlation coefficients with regard to the chosen minimization criteria.

Table 2.2 The common criteria for deriving the partial correlation coefficients

$F_m(n) = E\{f_m^2(n)\}$	– Mean value of squared m th forward prediction error
$B_m(n) = E\{b_m^2(n)\}$	– Mean value of squared m th backward prediction error
$C_m(n) = E\{f_m(n)b_m(n-1)\}$	– Cross-correlation of forward and backward prediction error signals

2.5 Method of Minimization of Forward Prediction Error

In this case the partial correlation coefficients in the $m + 1$ st step, $m = 0, 1, \dots, p$, are obtained as a result of minimization of the $m + 1$ st mean square forward prediction error and can be expressed as

$$F_{m+1}(n) = E\{f_{m+1}^2(n)\}. \quad (2.42)$$

If one replaces the following expression into (2.42)

$$f_{m+1}^2(n) = [f_m(n) + K_{m+1}b_m(n-1)]^2, \quad (2.43)$$

and makes the derivative over the unknown K_{m+1} equal to zero, one arrives at the solution

$$K_{m+1}^f = -\frac{E\{f_m(n)b_m(n-1)\}}{E\{b_m^2(n-1)\}} = -\frac{C_m(n)}{B_m(n-1)}. \quad (2.44)$$

The symbol f denotes that these partial correlation coefficients are obtained as a consequence of the minimization of the forward prediction error.

2.6 Method of Minimization of Backward Prediction Error

In this case one minimizes the mean square of backward prediction error

$$B_{m+1}(n) = E\{b_{m+1}^2(n)\}. \quad (2.45)$$

If one replaces the following expression into (2.42)

$$b_{m+1}^2(n) = [K_{m+1}f_m(n) + b_m(n-1)]^2, \quad (2.46)$$

and makes the derivative over K_{m+1} equal to zero, one arrives at the solution

$$K_{m+1}^b = -\frac{E\{f_m(n)b_m(n-1)\}}{E\{f_m^2(n)\}} = -\frac{C_m(n)}{F_m(n)}. \quad (2.47)$$

Since $F_m(n)$ and $B_m(n-1)$ have nonnegative values and their numerators are equal, it follows that the signs of K_{m+1}^f and K_{m+1}^b are equal, i.e.,

$$S = \text{sign}[K_{m+1}^f] = \text{sign}[K_{m+1}^b]. \quad (2.48)$$

The symbol b denoted that these partial correlation coefficients are obtained as a consequence of the minimization of the backward prediction error.

2.7 Method of Geometric Mean

In this case one minimizes forward and backward prediction errors simultaneously, and the partial correlation coefficients have the form

$$K_{m+1}^I = -\frac{E\{f_m(n)b_m(n-1)\}}{[E\{f_m^2(n)\}E\{b_m^2(n-1)\}]^{1/2}} = -\frac{C_m(n)}{[F_m(n)B_m(n-1)]^{1/2}}. \quad (2.49)$$

The methods of forward and backward prediction theoretically do not guarantee the stability of the analysis for a finite minimization interval. The method of geometric mean theoretically also guarantees stability in the case of the choice of a finite minimization interval. It obtained its name because K_{m+1}^I represents a geometric means of K_{m+1}^f and K_{m+1}^b , i.e.,

$$K_{m+1}^I = S \left| K_{m+1}^f K_{m+1}^b \right|^{1/2} \quad (2.50)$$

where S is defined by expression (2.48). The radicand is always positive, since the signs of K_{m+1}^f and K_{m+1}^b are identical. A property of the geometric mean is that

$$\min \left\{ \left| K_{m+1}^f \right|, \left| K_{m+1}^b \right| \right\} \leq \left| K_{m+1}^I \right| \leq \max \left\{ \left| K_{m+1}^f \right|, \left| K_{m+1}^b \right| \right\}, \quad m = 0, 1, \dots, p. \quad (2.51)$$

Since K_{m+1}^I is always smaller than zero, if $K_{m+1}^f \geq 1$ then $K_{m+1}^b \leq 1$ and vice versa, the stability of the method of geometric mean is guaranteed, i.e., the transfer function of the linear predictor $1/A(z)$ is stable (the poles of the transfer function are within the unit circle $|z| < 1$).

2.8 Method of Minimum

This method is derived from the previous three. A new set of coefficients of partial correlation is defined using the criterion of minimum

$$K_{m+1}^M = S \min \left\{ \left| K_{m+1}^f \right|, \left| K_{m+1}^b \right| \right\}. \quad (2.52)$$

Such a criterion guarantees stable analysis, i.e., a stable transfer function of the linear predictor $1/A(z)$ determined in such manner.

2.9 General Method

A sequence of sets of partial correlation coefficients is defined in the following manner

$$K_{m+1}^r = S \left\{ \frac{1}{2} \left(\left| K_{m+1}^f \right|^r + \left| K_{m+1}^b \right|^r \right) \right\}^{1/r}, \quad m = 0, 1, \dots, p; \quad r \leq 0. \quad (2.53)$$

For $r = 0$ it follows that $K_{m+1}^0 = K_{m+1}^I$, and for $r = -\infty$ it follows that $K_{m+1}^{-\infty} = K_{m+1}^M$. In case that an infinite minimization interval is chosen $K_{m+1}^f = K_{m+1}^b$, which means that the results are independent on the choice of minimization criterion.

2.10 Method of Harmonic Mean

A set of partial correlation coefficients is defined as the harmonic mean of K_{m+1}^f and K_{m+1}^b

$$K_{m+1}^B = - \frac{2K_{m+1}^f K_{m+1}^b}{K_{m+1}^f + K_{m+1}^b} = - \frac{2C_m(n)}{F_m(n) + B_m(n-1)}. \quad (2.54)$$

It can be shown that the following is valid

$$\left| K_{m+1}^M \right| \leq \left| K_{m+1}^B \right| \leq K_{m+1}^I. \quad (2.55)$$

It is the only set of coefficients of partial correlation that can be derived based on the global minimization criterion. Namely, one can obtain expression (2.54) by minimizing the sum of forward and backward prediction errors

$$E_{m+1}(n) = F_{m+1}(n) + B_{m+1}(n-1). \quad (2.56)$$

By determining the derivative of the total prediction error over the unknown parameters K_{m+1}^B , making the obtained expression equal to zero from Eq. (2.56) one arrives to expression (2.54).

2.11 Lattice-Covariant LP Method

The partial correlation coefficients can be explicitly expressed as function of the mean square values of $F_m(n)$, $B_m(n)$ and $C_m(n)$.

While performing prediction using the m th order predictor, the forward $f_m(n)$ and the backward $b_m(n)$ prediction errors can be written in the form

$$f_m(n) = \sum_{k=0}^m a_k^{(m)} s(n-k), \quad (2.57)$$

$$b_m(n) = \sum_{k=0}^m a_k^{(m)} s(n-m+k). \quad (2.58)$$

By squaring the above expressions and determining the mean value at the minimization interval it is obtained that

$$F_m(n) = E\{f_m^2(n)\} = \sum_{k=0}^m \sum_{i=0}^m a_k^{(m)} a_i^{(m)} \Phi(k, i), \quad (2.59)$$

where $\Phi(k, i) = E\{s(n-k)s(n-i)\}$ is the covariant function. Further

$$B_m(n-1) = E\{b_m^2(n-1)\} = \sum_{k=0}^m \sum_{i=0}^m a_k^{(m)} a_i^{(m)} \Phi(m-1-k, m+1-i), \quad (2.60)$$

$$C_m(n) = E\{f_m(n)b_m(n-1)\} = \sum_{k=0}^m \sum_{i=0}^m a_k^{(m)} a_i^{(m)} \Phi(k, m+1-i). \quad (2.61)$$

$F_m(n)$, $B_m(n-1)$ and $C_m(n)$ can be explicitly expressed as functions of the product of prediction coefficients and covariant function. The coefficients of partial correlation are most often determined by the method of harmonic mean. The minimization of the sum of forward and backward prediction errors is performed according to that method, and the stability is guaranteed. When one uses the method of harmonic mean, $F_m(n)$ can be calculated recursively

$$F_m(n) = \left[1 - (K_M^B)^2\right] F_{m-1}(n); \quad F_0(n) = f_0 = R_0. \quad (2.62)$$

Expression (2.62) is correct when the minimization interval is infinite.

The algorithm of the covariant lattice method can be reduced to the following steps:

- calculation of the covariant function $\Phi(k, i)$ for $k, i = 0, 1, \dots, p$,
- $m = 0$,
- calculation of $C_m(n)$, $F_m(n)$ and $B_m(n-1)$ according to 2.59–2.61,
- calculation K_{m+1} according to (2.54),
- calculation of predictor coefficients $\{a_k^{(m+1)}\}$, $k = 0, 1, \dots, m+1$ from the predictor coefficients $\{a_k^{(m)}\}$, using (2.38) and (2.39),
- $m = m+1$,
- if $m < p$ return to the step of calculation of $C_m(n)$, $F_m(n)$ and $B_m(n-1)$, else the end of the calculation.

In the case of a stationary signal, or in the case of a choice of an infinite interval and the use of window function, covariation reduces to autocorrelation

$$\Phi(k, i) = R(k - i) = R(i - k). \quad (2.63)$$

If (2.63) is replaced to (2.59), (2.60) and (2.61), one obtains

$$F_m(n) = B_m(n-1) = \sum_{k=0}^m \sum_{i=0}^m a_k^{(m)} a_i^{(m)} R(i - k), \quad (2.64)$$

$$C_m(n) = C_m = \sum_{k=0}^m \sum_{i=0}^m a_k^{(m)} a_i^{(m)} R(m+1-i-k). \quad (2.65)$$

Starting from the system of equations

$$\sum_{i=0}^m a_i^{(m)} R(i - k) = 0; \quad k = 1, 2, \dots, m \quad (2.66)$$

and expression (2.62), one obtains the partial correlation coefficients

$$K_{m+1} = -\frac{C_m}{F_m} = -\frac{\sum_{k=0}^m a_k^{(m)} R(m+1-k)}{(1 - K_m^2) F_{m-1}}, \quad (2.67)$$

with an initial condition $F_0 = R(0)$. Expression (2.67) is mathematically fully equivalent to Durbin algorithm for autocorrelation method.

It follows from expression (2.67) that errors in the calculation of m th partial correlation coefficient significantly influence the accuracy of the calculation of the subsequent coefficients, since errors accumulate. To avoid this, the partial correlation coefficients are calculated according to

$$K_{m+1} = - \frac{\sum_{k=0}^m \sum_{i=0}^m a_k^{(m)} a_i^{(m)} R(m+1-i-k)}{\sum_{k=0}^m \sum_{i=0}^m a_k^{(m)} R(i-k)}. \quad (2.68)$$

Relation (2.68) defines the autocorrelation-lattice method of linear prediction.

Expressions for $F_m(n)$, $B_m(n-1)$ and $C_m(n)$ can be written in the form that, using the mathematical principle of association, minimizes the number of multiplication operations to approximately one half. These expressions become

$$\begin{aligned} C_m(n) = & \Phi(0, m+1) + \sum_{k=1}^m a_k^{(m)} [\Phi(0, m+1-k) + \Phi(k, m+1)] + \\ & + \sum_{k=1}^m (a_k^{(m)})^2 \Phi(k, m+1-k) + \\ & + \sum_{k=1}^{m-1} \sum_{i=k+1}^m a_k^{(m)} a_i^{(m)} [\Phi(k, m+1-i) + \Phi(i, m+1-k)], \end{aligned} \quad (2.69)$$

$$\begin{aligned} F_m(n) + B_m(n-1) = & \Phi(0,0) + \Phi(m+1, m+1) + \\ & + 2 \sum_{k=1}^m a_k^{(m)} [\Phi(0, k) + \Phi(m+1, m+1-k)] + \\ & + \sum_{k=1}^m (a_k^{(m)})^2 [\Phi(k, k) + \Phi(m+1-k, m+1-k)] + \\ & + 2 \sum_{k=1}^{m-1} \sum_{i=k+1}^m a_k^{(m)} a_i^{(m)} [\Phi(k, i) + \Phi(m+1-k, m+1-i)]. \end{aligned} \quad (2.70)$$

By analyzing relations (2.59)–(2.61) it can be shown that more efficient minimization algorithms exist, which not only minimize the number of multiplication operations, but also minimize the total number of arithmetic operations. From (2.69) and (2.70) it follows that minimization of the number of multiplication operations results in an increase of the number of addition operations and in complex indexing, so that some beneficial properties of expressions (2.59), (2.60) and (2.61) are lost.

2.12 Basic Properties of Partial Correlation Coefficient

In the case of stable analysis the modulus of the partial correlation coefficients is smaller than zero, i.e., $|K_m| < 1$, $m = 1, 2, \dots, p$. One can control the stability of analysis by testing this condition. Instabilities may occur even in algorithms that theoretically guarantee stability, if implemented using arithmetic with fixed decimal mark and finite word length.

The value of a partial correlation coefficient is independent on signal amplification. This property is absent from some other digital signal analysis methods, for instance Fast Fourier Transform (FFT).

Spectral sensitivity of partial correlation parameters

$$\frac{\partial S}{\partial K_m} = \lim_{\Delta K_m \rightarrow 0} \left| \frac{\Delta S}{\Delta K_m} \right| \quad (2.71)$$

is largest for lowest order parameters and decreases with an increase of the index of the coefficient, where the spectrum variation is

$$\Delta S = \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} |10(\log S(K_m, \omega) - \log S(K_m + \Delta K_m, \omega))^2| d\omega \right]^{1/2}, \quad (2.72)$$

and ΔK_m is the change of value of m th partial correlation coefficient, causing a variation of ΔS .

Partial correlation coefficients of are nonuniformly distributed within the interval $(-1, 1)$ and are grouped around a mean value. Their scattering is smaller for a larger index.

2.13 Equivalence of Discrete Model and Linear Prediction Model

Speech signal is described as an output of a discrete model, i.e., digital filter, with a z -function transfer

$$H(z) = \frac{G}{A(z)} \quad (2.73)$$

which is excited by a sequence of periodical pulses or by white noise samples. $H(z)$ describes the envelope of speech signal spectrum. Samples of speech signal at the filter output can be expressed in the form

$$s(n) = \sum_{k=1}^p \alpha_k s(n-k) + Gu(n), \quad (2.74)$$

where p is the order of the filter, $Gu(n)$ —excitation signal, α_k —filter coefficients $H(z)$. Equation (2.74) describes discrete model in time domain.

As shown, the model of the analysis of speech signal using the linear prediction method is described by the difference equation

$$s(n) = \sum_{i=1}^p a_i s(n-i) + e(n). \quad (2.75)$$

If the filter coefficients (2.74) are equal to the linear prediction coefficients (2.75), i.e., if $\alpha_k = a_k$ and with the approximation of the excitation function $Gu(n)$ by the prediction error $e(n)$, these two models are equivalent. The adequacy of the use of linear prediction method for the analysis and synthesis of speech signal is based on this fact.

The coefficients of recursive filter $H(z)$ are actually linear prediction coefficients and are obtained directly using one of the quoted methods for their estimation. Different methods for estimation of linear prediction parameters (autocorrelation, covariant, lattice-covariant) can be understood as different statistical estimators of the coefficients of the filter $H(z)$.

The amplification constant G is the excitation parameter obtained from the condition that the excitation energy is equal to the prediction error energy

$$G^2 \sum_{m=0}^{N-1} u^2(m) = \sum_{m=0}^{N-1} e^2(m) = E_n, \quad (2.76)$$

where $u(n)$ are unit impulses with period T_0 for voiced segments, or samples of white noise with unit dispersion in the case of voiceless speech.

In the autocorrelation method the procedure to obtain the amplification constant is a constitutive part of the algorithm for the solution of matrix equation and is calculated as

$$G^2 = E_n = R(0) + \sum_{k=1}^p a_k R(k). \quad (2.77)$$

2.14 Speech Synthesis Based on Linear Prediction Model

Synthesis is a reverse process to analysis. A synthesizer of speech signal based on the LP model is basically a digital recursive filter with a form $H(z) = G/A(z)$, having poles only. Synthesized speech is obtained by filtering excitation function

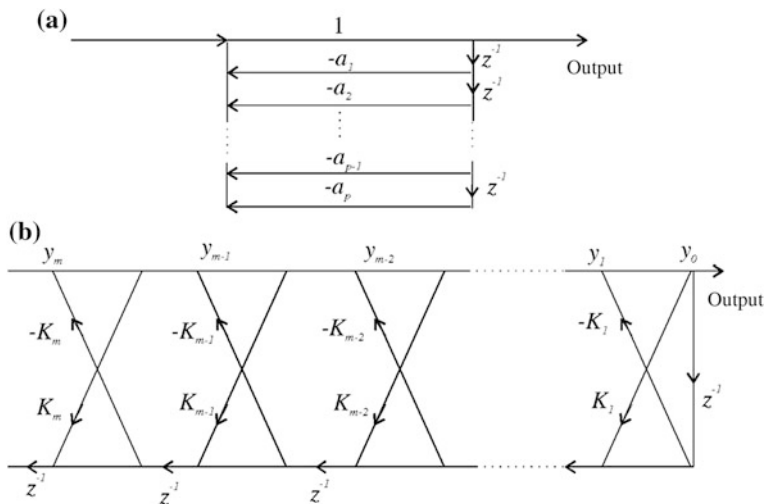


Fig. 2.5 Synthesizer structure. **a** Direct canonical structure, **b** lattice structure

using this filter. The parameters of the synthesizer filter are, in dependence on its structure, the linear prediction coefficients, the coefficients of partial correlation or some their transformations. These parameters are obtained by the analysis of speech signal and are changed each 15–20 ms. The basic problem with the implementation of synthesizer filters is stability. The condition of synthesizer filter stability reduces to the condition that the poles $H(z)$ should be located within the unit circle. A special aspect of the synthesis is implementation of synthesizer filter in arithmetic with fixed decimal point and a finite length of machine word.

There are different structures of synthesizer filters. The most often met ones in practical implementations of synthesizers are the lattice structure and the direct realization. These two structures can be regarded as the starting ones from which others are derived. Their application is especially convenient because the parameters of these structures are directly partial correlation coefficients in the lattice implementation, and linear prediction coefficients in the direct implementation. On the other hand, they represent the optimum from the point of view of the number of computational operations and the necessary memory space for their implementation.

The lattice filter of synthesizer is especially interesting because of the fact that it is possible to directly control the stability of a synthesizer filter by testing the values of the coefficients of partial correlations. Fig. 2.5a shows a block scheme of direct realization, and Fig. 2.5b shows a lattice structure of a synthesizer filter.

Robust Digital Processing of Speech Signals

Kovacevic, B.; Milosavljevic, M.M.; Veinović, M.; Marković, M.

2017, XII, 224 p. 54 illus., Hardcover

ISBN: 978-3-319-53611-8