EECS 225D, Lecture 25; 3/15/95
Ben Gold, Lecturer *Joseph Saracino, scribe*

Spectral Analysis / Vocoders Part II:

Linear Predictive Coding:

Linear Predictive Coding (LPC) was invented at Bell Labs in the 1960s as a new method of vocoding. The original work was geared towards a rate of 2400 bits/sec; more recent work has approached 10k bits/sec.

Concept behind LPC:
If we have a sequence of N values, the LPC vocoder works by predicting any value n based on previous values in the sequence. Note: LPC has always been performed with discrete sequences, though it is possible (and difficult) to build an analog LPC.

Example: If we have a set of numbers with some correlation: { . . . d, e, f, . . .}, we can express e as a function of d minus f.

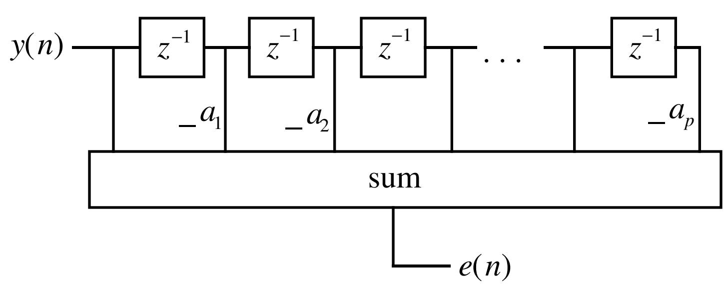
eqn 1: $$\hat{y}(n) = \sum_{j=1}^{p} y(n-j)a_j$$

eqn 2: $$e(n) = y(n) - \hat{y}(n)$$ (error signal)

eqn 2a: $$e(n) = y(n) - \sum_{j=1}^{p} y(n-j)a_j$$

Equation 2a can be expanded into a filter with p terms:

Filter 1:

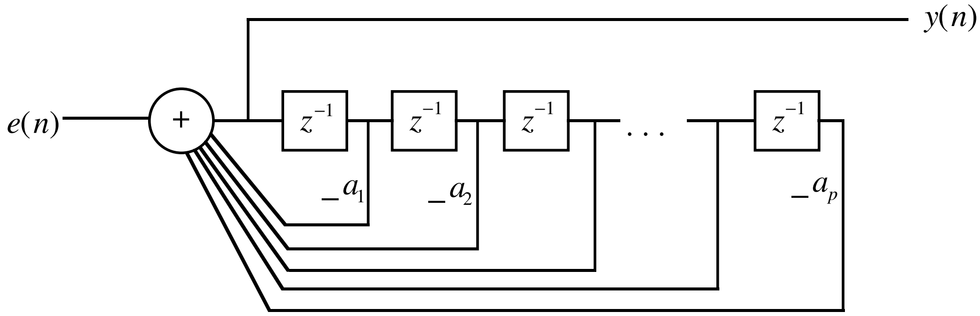


The Z-transform of Filter 1: $$E(z) = Y(z) - \sum_{j=1}^{p} Y(z)z^{-j}a_j$$ (polynomial in z)

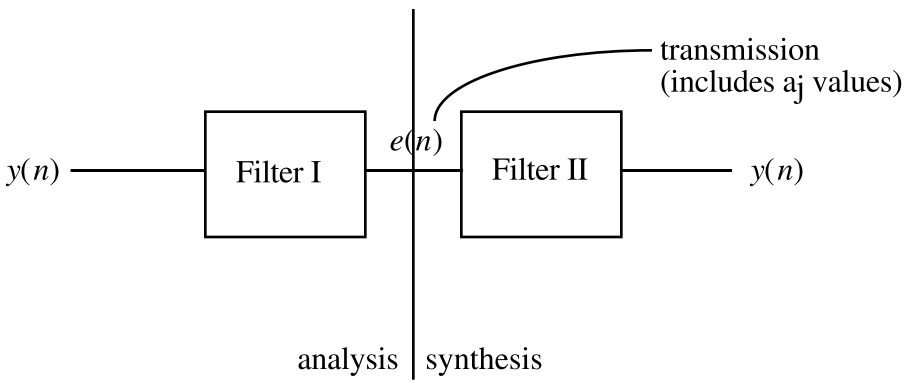
Note, this is an all-zero filter -- there are no terms in the denominator, therefore no poles in the Z-transform. Also, an all-zero is the inverse of an all-pole network. We assume that the coefficients, a_j, are either constant or slowly varying -- thus the receiver has knowledge of their value and can easily invert the above equation to get the original signal.

If we find the inverse we get: $$Y(z) = \frac{E(z)}{1 - \sum_{j=1}^{p} Y(z)z^{-j}a_j}$$

Which corresponds to Filter II:



Filters I and II can be combined to create an analysis/synthesis block with a transmission component between:



Analysis:
Ideally, we want to minimize the mean square output of Filter I. If the prediction method chosen is any good, most of the speech information should have been transferred to the a_j's and e(n) should contain substantially less information than the original signal -- achieving bandwidth reduction.

It is important now to compute the a_j coefficients. We try to find the minimum mean-squared value of the error signal over a relatively constant time interval (when the signal is sufficiently stationary).

eqn 3: $$\varepsilon = \sum_{n=0}^{N} e^2(n) = \sum_{n=0}^{N} [y(n) - \hat{y}(n)]^2$$

then compute $\frac{\partial \varepsilon}{\partial a_1} = 0$, $\frac{\partial \varepsilon}{\partial a_2} = 0$, . . ., $\frac{\partial \varepsilon}{\partial a_p} = 0$. We generate p equations with p unknowns and generate a solvable matrix.

Example, p = 5:

$$\begin{vmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \\ R_5 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \end{vmatrix} \begin{vmatrix} R_0 & R_1 & R_2 & R_3 & R_4 \\ R_1 & R_0 & R_1 & R_2 & R_3 \\ R_2 & R_1 & R_0 & R_1 & R_2 \\ R_3 & R_2 & R_1 & R_0 & R_1 \\ R_4 & R_3 & R_2 & R_1 & R_0 \end{vmatrix}$$

R is the correlation between adjacent samples (recall original concept).

$$R_j = \sum_{n=0}^{N-1} y(n)y(n-j);$$ From above: $R_1 = a_1 R_0 + a_2 R_1 + a_3 R_2 + a_4 R_3 + a_5 R_4$

This matrix can also be solved by using the **Levinson Recursion Method**. By recursing from any value of p to the next larger value p, successive values of a_j can be found.

If speech signals are not used, one gets strange results. Standard 2400bps LPC's use an excitation signal derived from pitch detection. Pitch estimation assumes a single voice (or instrument). Thus, for this input the pitch estimator is confounded. The tape plays the audible result of this confounding.