



语音信号处理

语音信号线性预测分析

李军锋

中国科学院声学研究所
中科院语言声学与内容理解重点实验室



大纲

1

概述

2

线性预测的基本原理

3

线性预测和语音信号模型的关系

4

线性预测方程组的解法

5

模型增益 G 的确定

6

线谱对分析法

7

线性预测的几种推演参数

8

线性预测分析的频域解释



概述

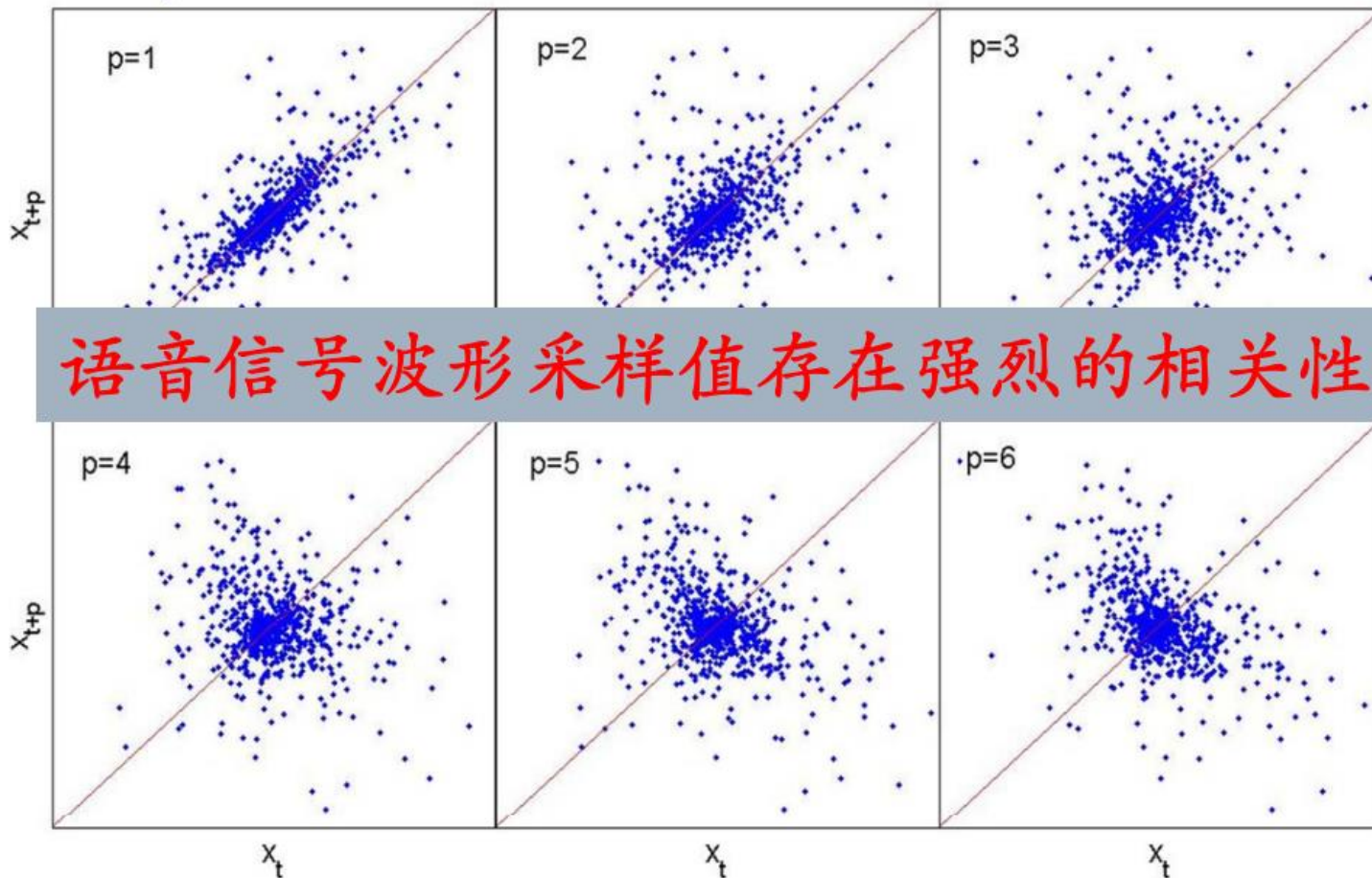
在各种语音处理技术中，线性预测是第一个真正得到实际应用的技术，可用于估计基本的语音参数如基音周期、共振峰频率、谱特征以及声道截面积函数等。

本讲主要介绍语音信号线性预测分析的基本原理，线性预测系数的求解方法以及线性预测的几种等价参数。



LPC分析的动因: 语音信号的时域相关性

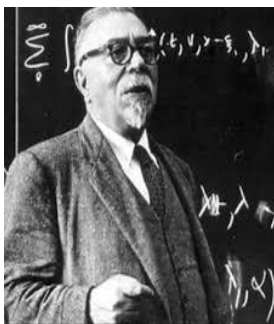
- 一段语音信号采样点 X_t , 构造散点图 (X_t, X_{t+p})



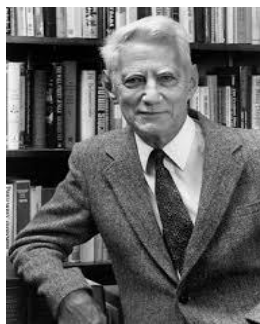


线性预测理论的发展史

- 1947年，美国科学家Norbert Wiener提出了线性预测理论，创立了控制论
- 1948年，美国科学家Claude Elwood Shannon提出了信息的测度理论，创立了信息论
- 1967年，日本科学家Fumitada Itakura（板仓文忠）将线性预测应用于了语音生成建模和语音合成



N. Wiener



C. E. Shannon



F. Itakura



LPC的应用

- 用作语音识别的特征参数
- 用作语音编码
- 进行共振峰的谱包络估计
- 进行语音合成（TTS）



常用特征

- 幅度、能量、过零率、基音周期
- 短时频谱

- ✓ 共振峰频率 (**formant**)
- ✓ 线谱对参数 (**LSP, linear spectrum pair**)
- ✓ 线性预测系数 (**LPC, linear prediction coefficients**)
- ✓ 线性预测倒谱系数 (**LPCC, linear prediction cepstral coefficients**)
- ✓ 梅尔频率倒谱系数 (**MFCC, Mel frequency cepstral coefficients**)

反应声道特性



反应人耳听觉特性





与LPC相关的语音编码算法

MPLPC: Multi-pulse LPC

RPE-LPC: Regular pulse excited LPC

CELP: Code excited linear prediction

MELP: Mix excited LP

VSELP: Vector sum excited LP

QCELP: Qualcomm CELP

ACELP: Algebraic CELP



LPC的基本思想

1. 一个语音的抽样能够用过去若干个语音抽样的线性组合来逼近
2. 这个线性预测的抽样和实际语音抽样之间存在着误差
3. 通过实现预测采样在最小均方误差意义上逼近实际采样，可以求取一组唯一的预测系数。

$$\hat{x}(n) = \sum_{i=1}^p a_i x(n-i)$$

$$e(n) = x(n) - \hat{x}(n)$$

$$E = \sum_n e^2(n) = \sum_n [x(n) - \hat{x}(n)]^2$$

预测系数：线性组合中所用的加权系数， $\{a_i\}$



设语音信号的样值序列为： $s(n)$ ， $n = 1, 2, \dots, n$

p 阶线性预测：根据信号过去 p 个取样值的加权和来预测信号当前取样值 $s(n)$ ，此时的预测器称为 p 阶预测器。

设 $\hat{s}(n)$ 为 $s(n)$ 的预测值，则有

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

线性预测系数： a_1, a_2, \dots, a_p

上式称为线性预测器，预测器的阶数为 p 阶。



p 阶线性预测器的传递函数为

$$P(z) = \sum_{i=1}^p a_i z^{-i}$$

线性预测误差 $e(n)$ ： 信号 $s(n)$ 与其线性预测 $\hat{s}(n)$ 值之差。 $e(n)$ 表示式为：

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

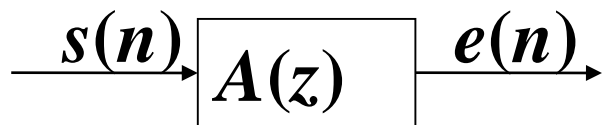


预测误差 $e(n)$ 是信号 $s(n)$ 通过如下系统的输出：

$$A(z) = 1 - \sum_{i=1}^p a_i z^{-i}$$

$A(z)$ 称为LPC误差滤波器。

LPC分析：即设计预测误差滤波器 $A(z)$ 的过程，也就是求解预测系数，使得预测器的误差 $e(n)$ 在某个预定的准则下最小。



LPC误差滤波器



线性预测的基本问题就是由语音信号直接求出一组线性预测系数 a_1 、 $a_2 \cdots a_p$ 使得在一短段语音波形中均方预测误差最小。将 $E[e^2(n)]$ 对各个系数求偏导，并令其结果为零，即

$$\frac{\partial E[e^2(n)]}{\partial a_k} = 2E[e(n) \frac{\partial e(n)}{\partial a_k}] = 0, \quad k = 1, 2, \cdots, p$$

由

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

得

$$\frac{\partial e(n)}{\partial a_k} = -s(n-k), \quad k = 1, 2, \cdots, p$$



$$\frac{\partial e(n)}{\partial a_k} = -s(n-k) , \quad k = 1, 2, \dots, p$$

$$-2E[e(n)s(n-k)] = 0 , \quad k = 1, 2, \dots, p$$

上式称为正交方程。

$$E[e(n)s(n-k)] = E[s(n)s(n-k) - \sum_{i=1}^p a_i s(n-i)s(n-k)] = 0, \quad k = 1, 2, \dots, p$$



令 $s(n)$ 的自相关序列为

$$R(k) = E[s(n)s(n-k)]$$

由于自相关序列为偶对称，因此

$$R(k) = R(-k) = E[s(n)s(n+k)]$$

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

上式称为标准方程式，它表明只要语音信号是已知的，则 p 个预测系数 通过求解该方程即可得到。



设

$$A_p = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}$$

$$R_p = \begin{bmatrix} R(0) & R(1) & \cdots & R(p-1) \\ R(1) & R(0) & \cdots & R(p-2) \\ \vdots & \vdots & & \vdots \\ R(p-1) & R(p-2) & \cdots & R(0) \end{bmatrix}$$

$$R_p^\alpha = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix}$$

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

矩阵形式为

$$R_p^\alpha - R_p A_p = 0$$

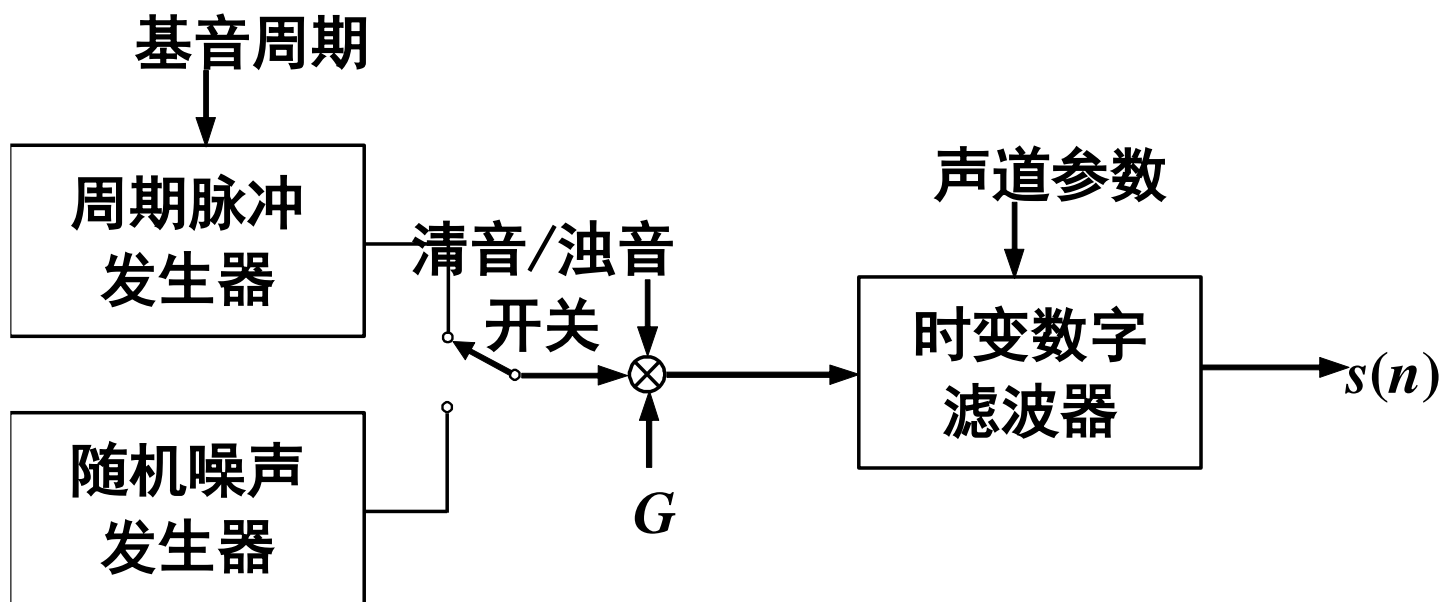
或者

$$A_p = R_p^{-1} R_p^\alpha$$

通过求解上式即可求得 p 个线性预测系数



LPC和语音信号模型的关系



语音产生的数字模型简化图



声门激励、声道调制和嘴唇辐射的合成贡献，可用如下数字时变滤波器表示

$$H(z) = \frac{S(z)}{U(z)} = \frac{G \left(1 - \sum_{l=1}^q b_l z^{-l}\right)}{1 - \sum_{i=1}^p a_i z^{-i}}$$

上式既有极点又有零点。按其有理式的不同，有如下三种信号模型：

- (1) 自回归滑动平均模型（ARMA模型）；
- (2) 自回归信号模型（AR模型）；
- (3) 滑动平均模型（MA模型）。



一般都用AR模型作为语音信号处理的常用模型。此时 $H(z)$ 写为

$$H(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - \sum_{i=1}^p a_i z^{-i}}$$

当 p 足够大时，上式几乎可以模拟所有语音信号的声道系统。

采用简化模型的主要优点：可以用线性预测分析法对增益 G 和滤波器系数进行直接而高效的计算。



在语音产生的数字模型中，语音抽样信号 $s(n)$ 和激励信号之间的关系可用下列差分方程来表示：

$$s(n) = \sum_{i=1}^p a_i s(n-i) + Gu(n)$$

可见，如果语音信号准确服从上式的模型，则 $e(n) = Gu(n)$ ，所以预测误差滤波器 $A(z)$ 是 $H(z)$ 的逆滤波器，故有下式成立：

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

$H(z)$ 称为合成滤波器。



线性预测误差滤波相当于一个逆滤波过程或逆逼近过程，当调整滤波器 $A(z)$ 的参数使输出 $e(n)$ 逼近一个白噪声序列 $u(n)$ 时， $A(z)$ 和 $H(z)$ 是等效的，而按最小均方误差准则求解线性预测系数正是使输出 $e(n)$ 白化的过程。



LPC阶数确定

◆预测误差与阶数的关系

◆选择**12阶**对绝大多数语音信号的声道模型可以取得足够的近似逼近

◆一般取**8~12阶**

◆阶数过高的副作用

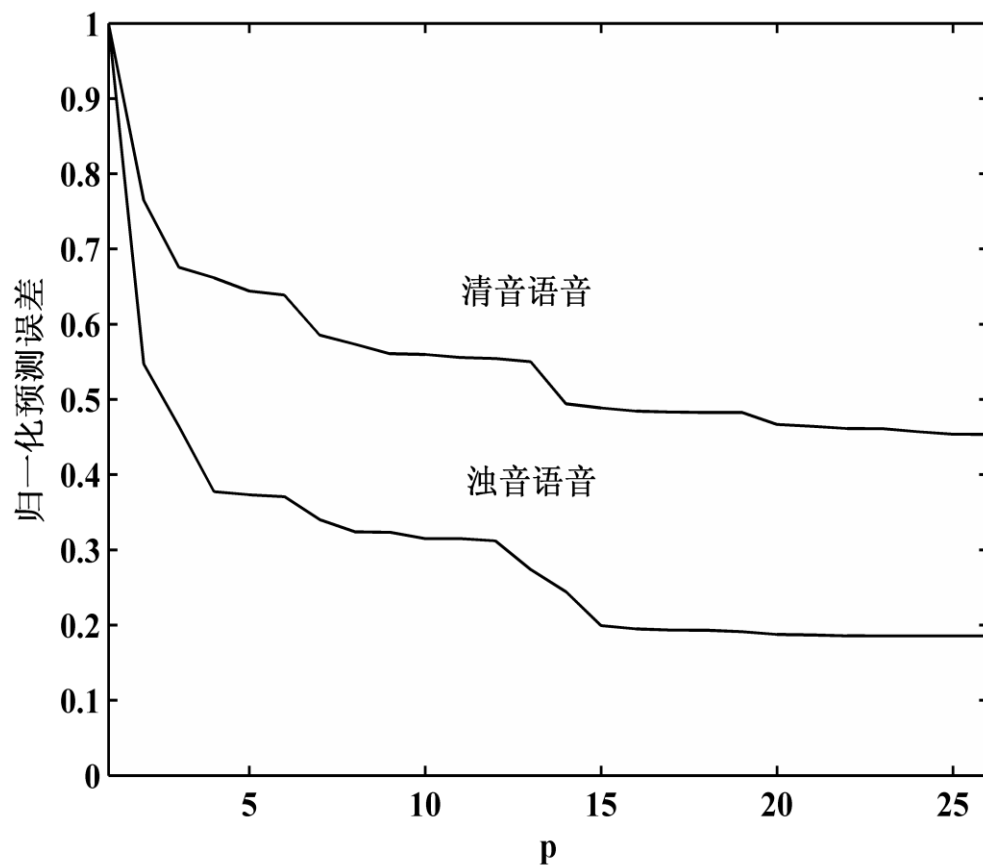
◆加大计算量

✓增加一些不必要的细节

✓清音与浊音的逼近效果不同：清音明显高



16kHz采样率时，清音和浊音语音相对应的归一化预测误差随阶数 p 的变化曲线如下图：





通过预加重对高频部分的提升

由于声门脉冲形状和口唇辐射的影响，语音信号的频谱在总趋势上会有高频衰落的现象。为抵消这一影响，通常在做**LPC**分析之前采用一阶**FIR**滤波器进行预加重，以进行高频提升。传递函数为

$$H(z) = 1 - \alpha z^{-1}$$

α 为预加重系数，通常取**0.9375**。



LPC求解

由于语音信号的时变特性，线性预测分析应该在短时的语音段上进行，即按帧进行。

对单个样本：

在每一帧信号内，当前的样本可以用该帧信号的前 p 个样本来预测，得到如下预测值：

$$\hat{x}(n) = \sum_{i=1}^p a_i x(n-i)$$

当前样本的预测误差 $e(n)$ 为：

$$e(n) = x(n) - \hat{x}(n) = x(n) - \sum_{i=1}^p a_i x(n-i)$$



LPC求解

对一帧样本:

为了在最小均方误差意义上计算一组最佳预测系数,定义短时预测均方误差为:

$$E_n = \sum_n e^2(n) = \sum_n [x(n) - \hat{x}(n)]^2 = \sum_n [x(n) - \sum_{i=1}^p a_i x(n-i)]^2$$

为了使 E_n 达到最小, $\{a_i\}$ 必须满足 $\partial E_n / \partial a_k = 0$ 式中 $k=1,2,\dots,p$.则得到:

$$\frac{\partial E_n}{\partial a_k} = -(2 \sum_n x(n)x(n-k) - 2 \sum_{i=1}^p a_i \sum_n x(n-k)x(n-i)) = 0$$



LPC方程的自相关解法

求解 p 个线性预测系数的依据，是预测误差滤波器的输出方均值或输出功率最小。称这一最小方均误差为正向预测误差功率，即

$$\begin{aligned} E_p &= E[e^2(n)]_{\min} = E\left\{e(n)\left[s(n) - \sum_{i=1}^p a_i s(n-i)\right]\right\} \\ &= E[e(n)s(n)] - \sum_{i=1}^p a_i E[e(n)s(n-i)] \end{aligned}$$

上式第二项为0。



$$E_p = E[e(n)s(n)] = E[s(n)s(n)] - \sum_{i=1}^p a_i E[s(n)s(n-i)] = R(0) - \sum_{i=1}^p a_i R(i)$$

$$R_p^{\alpha} - R_p A_p = 0$$

以上两式组合起来得

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(p) \\ R(1) & R(0) & \cdots & R(p-1) \\ R(2) & R(1) & \cdots & R(p-2) \\ \vdots & \vdots & & \vdots \\ R(p) & R(p-1) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} 1 \\ -a_1 \\ -a_2 \\ \vdots \\ -a_p \end{bmatrix} = \begin{bmatrix} E_p \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

称为尤勒-沃尔克（Yule-Walker）方程
方程的系数矩阵为托普利兹（Toeplitz）矩阵



可见，为了解得线性预测系数，必须首先计算出自相关序列 $R(k)$ ， $R(k)$ 可用下式估计

$$R(k) = E[s(n)s(n-k)] = \frac{1}{n} \sum_n s(n)s(n-k)$$

如果将预测误差功率 E_p 理解为预测误差的能量，则上式中的系数 $\frac{1}{n}$ 对线性预测方程的求解没有影响，因此可以忽略。但其中的求和范围 n 的不同定义，将会导致不同的线性预测解法。经典的方法有两种：一种是自相关法，另一种是协方差法。



自相关法的详细求解过程

利用对称托普利兹(Toeplitz)矩阵的性质，自相关法求解可用Levinson-Durbin（莱文逊-杜宾）递推算法求解。

该方法是目前广泛采用的一种方法。利用Levinson-Durbin算法递推时，从最低阶预测器开始，由低阶到高阶进行逐阶递推计算。



自相关法递推过程如下

$$k_i = \left[r(i) - \sum_{j=1}^{i-1} a_j^{(i-1)} r(i-j) \right] / E_{(i-1)} , \quad 1 \leq i \leq p$$

$$E_{(0)} = r(0)$$

$$E_i = (1 - k_i^2) E_{(i-1)}$$

$$a_i^{(i)} = k_i$$

$$a_j^{(i)} = a_j^{(i-1)} - k_i a_{i-j}^{(i-1)} , \quad 1 \leq j \leq i-1$$

联立左面5式可对 $i=1$ 、 $2\dots$ 、 p 进行递推求解，其最终解为

$$a_j = a_j^{(p)} , \quad 1 \leq j \leq p$$

对于 p 阶预测器，在上述求解预测器系数的过程中，阶数低于 p 的各阶预测器系数也同时得到。



协方差法

协方差法不规定语音信号序列 $s(n)$ 的长度范围，且需要确定的是信号序列之间的互相关函数，由此组成的协方差方程组系数矩阵已经不具有Toeplitz矩阵的性质，因此其方程的求解不同于自相关法。由于不需要加窗，协方差法计算精度较自相关法大大提高。



用协方差求解线性预测系数,不需要对语音信号进行加窗处理
 $\phi(k, i)$ 重新定义为:

$$\Phi(k, i) = \sum_{n=0}^{N-1} x(n-k)x(n-i), \quad k = 1, 2, \dots, p, \quad i = 0, 1, 2, \dots, p$$

设 $(n-i)=m$,则上式可写为:

$$\Phi(k, i) = \sum_{m=-i}^{N-i-1} x(m+(i-k))x(m), \quad k = 1, 2, \dots, p, \quad i = 0, 1, 2, \dots, p$$

可以看出,此处的 $\phi(k, i)$ 不仅取决于 k 和 i 的差值,还取决于 i 本身。因此,这里的 $\phi(k, i)$ 就不再是自相关函数。虽然仍有 $\phi(k, i) = \phi(i, k)$ 但是不能满足 $\phi(i+1, k+1) = \phi(i, k)$

这样,线性方程组可写成下面的形式:



$$\begin{bmatrix} \Phi(1,1) & \Phi(1,2) & \Phi(1,3) & \cdots & \Phi(1,p) \\ \Phi(2,1) & \Phi(2,2) & \Phi(2,3) & \cdots & \Phi(2,p) \\ \Phi(3,1) & \Phi(3,2) & \Phi(3,3) & \cdots & \Phi(3,p) \\ \vdots & \vdots & \vdots & & \vdots \\ \Phi(p,1) & \Phi(p,2) & \Phi(p,3) & \cdots & \Phi(p,p) \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \vdots \\ \hat{a}_p \end{bmatrix} = \begin{bmatrix} \Phi(1,0) \\ \Phi(2,0) \\ \Phi(3,0) \\ \vdots \\ \Phi(p,0) \end{bmatrix}$$

此方程组的系数不再是一个托布里兹矩阵,虽然它仍是对称阵,但主对角线和各个副对角线上的元素并不相等.

这种线性方程组也有多种解法,其中最常用的解法是乔里斯基(Choleskey)分解法,其基本思想是将系数矩阵采用消元法化成主对角线元素为1的上三角矩阵,然后再对变量逐个递推求解.

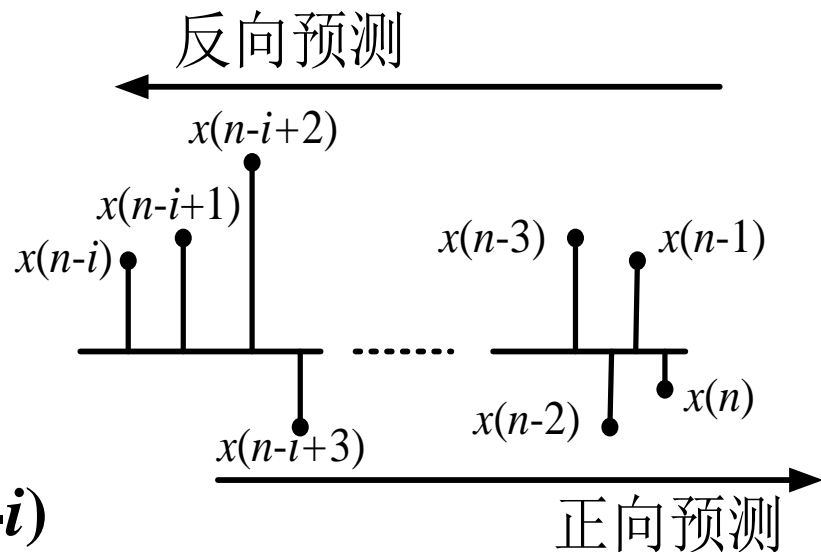


正向预测:

用 i 个过去的样本值 $x(n-1)$, $x(n-2),\dots,x(n-i)$ 来预测 $x(n)$

反向预测:

用时间上延迟时刻的样本值 $x(n-i+1),x(n-i+2),\dots,x(n)$ 来预测 $x(n-i)$



正向预测误差:用 i 个过去的样本值 $x(n-1),x(n-2),\dots,x(n-i)$ 来预测 $x(n)$ 时产生的误差

$$e^{(i)}(n) = x(n) - \sum_{j=1}^i a_j^{(i)} x(n-j)$$

反向预测误差:时间上延迟时刻的样本值 $x(n-i+1),x(n-i+2),\dots,x(n)$ 来预测 $x(n-i)$ 时的误差

$$b^{(i)}(n) = x(n-i) - \sum_{j=1}^i a_j^{(i)} x(n-i+j)$$



格型法

$$\begin{aligned}\text{令: } a_i^{(i)} &= k_i; \\ a_j^{(i)} &= a_j^{(i-1)} - k_i a_{i-j}^{(i-1)}; j = 1, \dots, i-1\end{aligned}$$

经过推导,可以得到: $e^{(i)}(n) = e^{(i-1)}(n) - k_i b^{(i-1)}(n-1)$

$$b^{(i)}(n) = b^{(i-1)}(n-1) - k_i e^{(i-1)}(n)$$

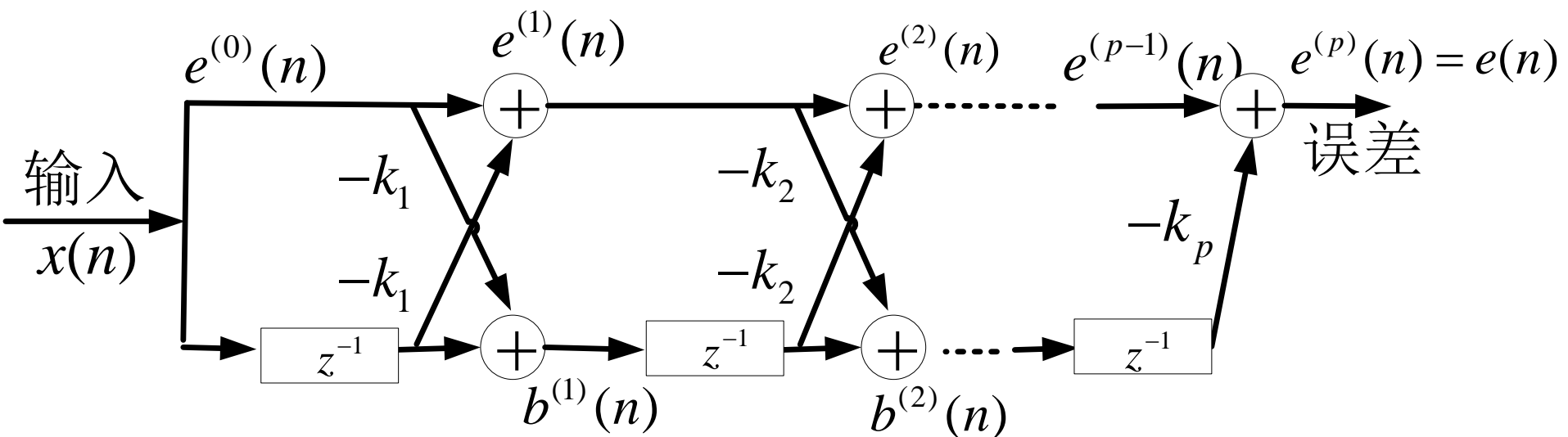
当 $i=0$ 时得到: $e^{(0)}(n) = b^{(0)}(n) = x(n)$

当 $i=p$ 时得到: $e^{(p)}(n) = e(n)$

式中 $e(n)$ 为 p 阶线性预测误差滤波器所输出的预测误差信号:通过推导可以得到适合于线性预测分析的格型滤波器结构形式



格型滤波器



其中滤波器输入为 $\mathbf{x}(n)$,输出为预测误差 $\mathbf{e}(n)$

接下来, 通过最小化 $e(n)$ 即可求出 k_n , 继而求出 \hat{a}_j



定义三种均方误差:

$$\left\{ \begin{array}{ll} \text{正向均方误差} & E^{(i)}(n) = E \left[\left(e^{(i)}(n) \right)^2 \right] \\ \text{反向均方误差} & B^{(i)}(n) = E \left[\left(b^{(i)}(n) \right)^2 \right] \\ \text{交叉均方误差} & C^{(i)}(n) = E \left[e^{(i)}(n) b^{(i)}(n-1) \right] \end{array} \right.$$

正向格型法逼近原则:使格型滤波器的第*i*节正向均方误差最小

即令
$$\frac{\partial E^{(i)}(n)}{\partial k_i} = 0$$

经过推导得到:
$$k_i^f = \frac{C^{(i-1)}(n)}{B^{(i-1)}(n-1)} = \frac{E[e^{(i-1)}(n)b^{(i-1)}(n-1)]}{E[(b^{(i-1)}(n-1))^2]}$$

式中*k*的上标*f*表示此反射系数是用正向误差最小准则求得的。它等于正反向预测误差的互相关和反向预测误差能量之比。



几种求解线性预测方法的比较

自相关法使用了窗函数来对语音信号进行人为的截断, 导致分辨率降低, 并且数据越短, 分辨率越低, 因而自相关法误差较大, 计算结果精度差。

协方差法由于不采用窗口函数, 精度高于自相关法和格型法, 但它的稳定性同样得不到保证并且计算量较大。

因此在语音信号处理中, 自相关法比协方差法用得少。

格型法无需加窗, 也不需计算自相关矩阵. 其计算结果精度很高, 对系统的稳定性也有保证. 但该方法运算量很大, 大致为自相关法或协方差法的4倍以上。改进算法“协方差格型法”将计算量恢复到自相关法的水平上。

因此格型法是一种很有生命力的线性预测算法。



模型增益G的确定

线性预测误差滤波相当于一个逆滤波过程或逆逼近过程，
当调整滤波器 $A(z)$ 的参数使输出 $e(n)$ 逼近一个白噪声序列 $u(n)$
时， $A(z)$ 和 $H(z)$ 是等效的，而按最小均方误差准则求解线性
预测系数正是使输出 $e(n)$ 白化的过程。



模型增益G的确定

由

$$H(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - \sum_{i=1}^p a_i z^{-i}}$$

得

$$Gu(n) = s(n) - \sum_{i=1}^p a_i s(n-i)$$

对上式两边乘以 $s(n)$ 并求平均值，等式右边为

$$E\left[\left(s(n) - \sum_{i=1}^p a_i s(n-i)\right)s(n)\right] = E[s^2(n)] - \sum_{i=1}^p a_i E[s(n-i)s(n)] = R(0) - \sum_{i=1}^p a_i R(i)$$



等式左边为

$$\begin{aligned} GE[u(n)s(n)] &= E \left[Gu(n) \left(Gu(n) + \sum_{i=1}^p a_i s(n-i) \right) \right] \\ &= G^2 E[u^2(n)] + G \sum_{i=1}^p a_i E[u(n)s(n-i)] \end{aligned}$$

因为
得到

$$E[u^2(n)] = 1 \quad \mathbf{E}[u(n)s(n-i)] = \mathbf{0}$$

$$G^2 = R(0) - \sum_{i=1}^p a_i R(i)$$

与

$$E_p = E[e(n)s(n)] = R(0) - \sum_{i=1}^p a_i R(i)$$

比较得出

$$G^2 = E_p$$



$$G = \sqrt{E_p}$$



线谱对LSP分析

线谱对LSP是与LPC系数等价的一种表示形式。由Itakura（板仓）引入的。由于LSP能够保证线性预测滤波器的稳定性，其小的系数偏差带来的谱误差也只是局部的，且LSP具有良好的量化特性和内插特性，因而已经在许多编码系统中得到成功的应用。LSP分析的主要缺点是运算量较大。



LSP的定义和特点

设线性预测逆滤波器 $A(z)$ 为

$$A(z) = 1 - \sum_{i=1}^p a_i z^{-i}$$

由 $A(z)$ 组成的 $p+1$ 阶对称和反对称多项式表示如下：

$$P(z) = A(z) + z^{-(p+1)} A(z^{-1})$$

$$Q(z) = A(z) - z^{-(p+1)} A(z^{-1})$$

其中

$$z^{-(p+1)} A(z^{-1}) = z^{-(p+1)} - a_1 z^{-p} - a_2 z^{-p+1} - \cdots - a_p z^{-1}$$



可以推出

$$P(z) = 1 - (a_1 + a_p)z^{-1} - (a_2 + a_{p-1})z^{-2} - \cdots - (a_p + a_1)z^{-p} + z^{-(p+1)}$$

$$Q(z) = 1 - (a_1 - a_p)z^{-1} - (a_2 - a_{p-1})z^{-2} - \cdots - (a_p - a_1)z^{-p} - z^{-(p+1)}$$

$P(z)$ 、 $Q(z)$ 分别为对称和反对称的实系数多项式，它们都有共轭复根。可以证明，当 $A(z)$ 的根位于单位圆内时， $P(z)$ 和 $Q(z)$ 的根都位于单位圆上，而且相互交替出现。如果阶数 p 是偶数，则 $P(z)$ 和 $Q(z)$ 各有一个实根，其中 $P(z)$ 有一个根 $z = -1$ ， $Q(z)$ 有一个根 $z = 1$ 。如果阶数 p 是奇数，则 $P(z)$ 有 $z = \pm 1$ 两个实根， $Q(z)$ 没有实根。



此处假定 p 是偶数，这样 $P(z)$ 和 $Q(z)$ 各有 $p/2$ 个共轭复根位于单位圆上，共轭复根的形式为 $z_i = e^{\pm j\omega_i}$

设 $P(z)$ 的零点为 $e^{\pm j\omega_i}$ $Q(z)$ 的零点为 $e^{\pm j\theta_i}$ ，则满足

$$0 < \omega_1 < \theta_1 < \cdots < \omega_{p/2} < \theta_{p/2} < \pi$$

线谱频率LSF: ω_i 和 θ_i ，分别为 $P(z)$ 和 $Q(z)$ 的第 i 个根

$$P(z) = (1 + z^{-1}) \prod_{i=1}^{p/2} (1 - z^{-1} e^{j\omega_i})(1 - z^{-1} e^{-j\omega_i}) = (1 + z^{-1}) \prod_{i=1}^{p/2} (1 - 2 \cos \omega_i z^{-1} + z^{-2})$$

$$Q(z) = (1 - z^{-1}) \prod_{i=1}^{p/2} (1 - z^{-1} e^{j\theta_i})(1 - z^{-1} e^{-j\theta_i}) = (1 - z^{-1}) \prod_{i=1}^{p/2} (1 - 2 \cos \theta_i z^{-1} + z^{-2})$$

$\cos \omega_i, \cos \theta_i$: 余弦域的 LSP系数，与LSF对应



由于LSP参数 ω_i 和 θ_i 成对出现，且反映信号的频谱特性，因此称为线谱对。它们就是线谱对分析所要求解的参数。

LSP参数的特性：

1. LSP参数都在单位圆上且降序排列。
2. 与LSP参数对应的LSF升序排列，且 $P(z)$ 和 $Q(z)$ 的根相互交替出现，这可使与LSP参数对应的LPC滤波器的稳定性得到保证。

原因：上述特性保证了在单位圆上，任何时候 $P(z)$ 和 $Q(z)$ 不可能同时为零。



3. LSP参数具有相对独立的性质。如果某个特定的LSP参数中只移动其中任意一个线谱频率的位置，那么它所对应的频谱只在附近与原始语音频谱有差异，而在其它LSP频率上则变化很小。

优点：有利于LSP参数的量化和内插。



4. LSP参数能够反映声道幅度谱的特点，在幅度大的地方分布较密，反之较疏。这样就相当于反映出了幅度谱中的共振峰特性。

原因：按照线性预测分析的原理，语音信号的谱特性可以由LPC模型谱来估计，将下面两式相加

$$P(z) = A(z) + z^{-(p+1)} A(z^{-1})$$

$$Q(z) = A(z) - z^{-(p+1)} A(z^{-1})$$

可得

$$A(z) = \frac{1}{2} [P(z) + Q(z)]$$



这样，功率谱可以表示为

$$\begin{aligned} |H(e^{j\omega})|^2 &= \frac{1}{|A(e^{j\omega})|^2} = 4|P(e^{j\omega}) + Q(e^{j\omega})|^{-2} \\ &= 2^{-p} [\sin^2(\omega/2) \prod_{i=1}^{p/2} (\cos \omega - \cos \theta_i)^2 + \cos^2(\omega/2) \prod_{i=1}^{p/2} (\cos \omega - \cos \omega_i)^2]^{-1} \end{aligned}$$

可见：LSP分析是用 p 个离散频率的分布密度来表示语音信号谱特性的一种方法。即在语音信号幅度谱较大的地方LSP分布较密，反之较疏。

5. 相邻帧LSP参数之间都具有较强的相关性，便于语音编码时帧间参数的内插。



LPC的几种推演参数

在语音编码算法中，通常将线性预测滤波器系数转换为与之等效的参数，再进行量化编码。这些参数一般是由线性预测滤波器系数推演出来的，称之为线性预测的推演参数。这些推演参数各有不同的物理意义和特性。下面分别进行介绍。

1. 反射系数
2. 对数面积比系数
3. 预测器多项式的根
4. LPC倒谱





反射系数

也称为部分相关系数，即PARCOR系数，用 k_i 表示。已知线性预测系数 a_i ， $i=1,2,\dots,p$ ，求反射系数 k_i 递推过程如下：

$$\begin{cases} a_j^{(p)} = a_j & 1 \leq j \leq p \\ k_i = a_i^{(i)} \\ a_j^{(i-1)} = [a_j^{(i)} + a_i^{(i)} a_{i-j}^{(i)}] / (1 - k_i^2) & 1 \leq j \leq i-1 \end{cases}$$



反过来，已知反射系数 k_i ，求相应的线性预测系数 a_i ， $i=1,2,\dots,p$ ，的递推过程如下：

$$\begin{aligned} a_i^{(i)} &= k_i \\ a_j^{(i)} &= a_j^{(i-1)} - k_i a_{i-j}^{(i-1)} \quad 1 \leq j \leq i-1 \\ a_j &= a_j^{(p)} \quad 1 \leq j \leq p \end{aligned}$$

为了保证相应的线性预测合成滤波器的稳定性，反射系数 k_i 通常取为

$$-1 \leq k_i \leq 1$$



对数面积比系数LAR

由反射系数可进一步推导出对数面积比系数，
其定义为

$$g_i = \log(A_{i+1}/A_i) = \log[(1 - k_i)/(1 + k_i)] \quad 1 \leq i \leq p$$

对上式两边取以e为底的指数整理可得

$$k_i = (1 - \exp(g_i))/(1 + \exp(g_i)) \quad 1 \leq i \leq p$$

在语音编码系统中LAR渐渐被LSF参数取代。



预测器多项式的根

通过求取预测器多项式的根，可以实现对共振峰的估计。

$A(z)$ 用它的一组根 $\{z_i, 1 \leq i \leq p\}$ 等效表示为

$$A(z) = 1 - \sum_{i=1}^p a_i z^{-i} = \prod_i (1 - z_i z^{-1})$$

若 p 为偶数， $A(z)$ 得到的是 $p/2$ 对复根

$$z_k = z_{kr} \pm j \cdot z_{ki}, \quad k = 1, 2, \dots, p/2$$

每一对根与信号谱中的一个共振峰相对应





预测器多项式的根

如果把Z平面的根转换到S平面，令 $z_k = e^{s_k T}$ ，其中 T 为采样间隔，设 $s_k = \sigma_k + j\Omega_k$ ，则有

决定共振峰的频率

$$\Omega_k = \frac{1}{T} \arctan\left(\frac{z_{ki}}{z_{kr}}\right)$$

决定共振峰的带宽

$$\sigma_k = \frac{1}{2T} \log(z_{kr}^2 + z_{ki}^2)$$





LPC倒谱

线性预测倒谱系数LPCC定义：

是LPC系数在倒谱域表示。指的是这个信号 z 变换的对数模函数的反 z 变换。通过对语音信号的傅里叶变换取模的对数再求反傅里叶变换可得到一个信号的倒谱。

优点： 计算量小，易于实现，对元音有较好描述能力。

缺点： 对辅音的描述能力较差，抗噪性能较差。



求解方法

设通过线性预测分析得到的声道模型系统函数为

$$H(z) = \frac{1}{1 - \sum_{i=1}^p a_i z^{-i}}$$

其冲激响应为 $h(n)$ ，倒谱为 $\hat{h}(n)$ ，则有

$$\hat{H}(z) = \ln H(z) = \sum_{n=1}^{\infty} \hat{h}(n) z^{-n}$$

将 $H(z)$ 代入倒谱表示式并将其两边对 z^{-1} 求导，整理可得

$$\left(1 - \sum_{i=1}^p a_i z^{-i}\right) \sum_{n=1}^{\infty} n \hat{h}(n) z^{-n+1} = \sum_{i=1}^p i a_i z^{-i}$$



$$(1 - \sum_{i=1}^p a_i z^{-i}) \sum_{n=1}^{\infty} n \hat{h}(n) z^{-n+1} = \sum_{i=1}^p i a_i z^{-i}$$

令上式两边的各次 z^{-1} 的系数分别相等，可得由LPC系数求倒谱系数的递推公式：

$$\hat{h}(n) = \begin{cases} a_n & n = 1 \\ a_n + \sum_{k=1}^{n-1} k \hat{h}(k) a_{n-k} / n & 1 < n \leq p+1 \\ \sum_{k=1}^{n-1} k \hat{h}(k) a_{n-k} / n & n > p+1 \end{cases}$$



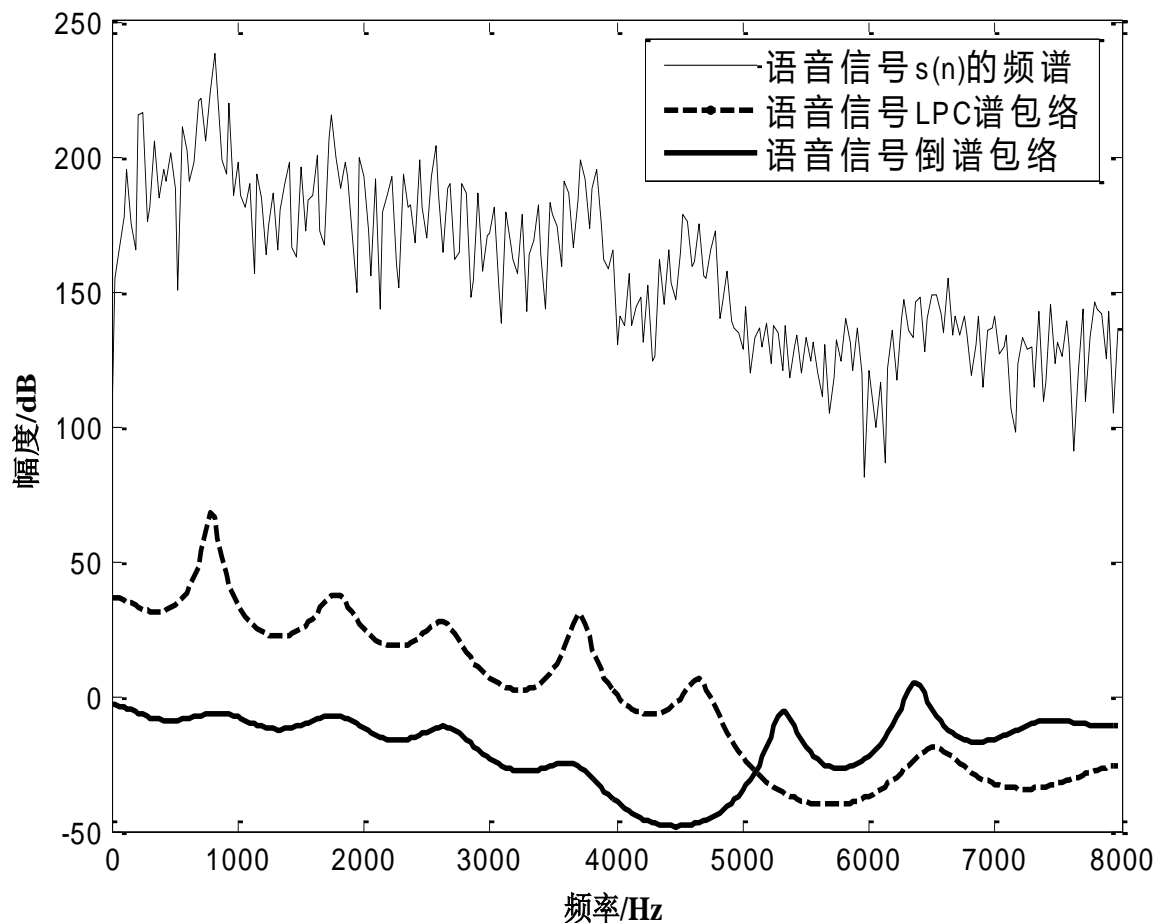
线性预测倒谱系数是一个右半序列。

语音信号的倒谱能较好地描述语音的共振峰特征，并较彻底地去掉了语音产生过程中的激励信息，因此在语音识别系统中得到了较好的应用效果。

实验表明，使用倒谱可以提高特征参数的稳定性。



下图给出了语音信号及其LPC谱包络与倒谱包络





预测误差滤波器的冲激响应及其自相关系数

全极点模型预测误差滤波器传递函数的单位冲激响应为

$$a(n) = \delta(n) - \sum_{i=1}^p a_i \delta(n-i) = \begin{cases} 1, & n = 0 \\ a_n, & 0 < n \leq P \\ 0, & \text{其他} \end{cases} \quad (5.63)$$

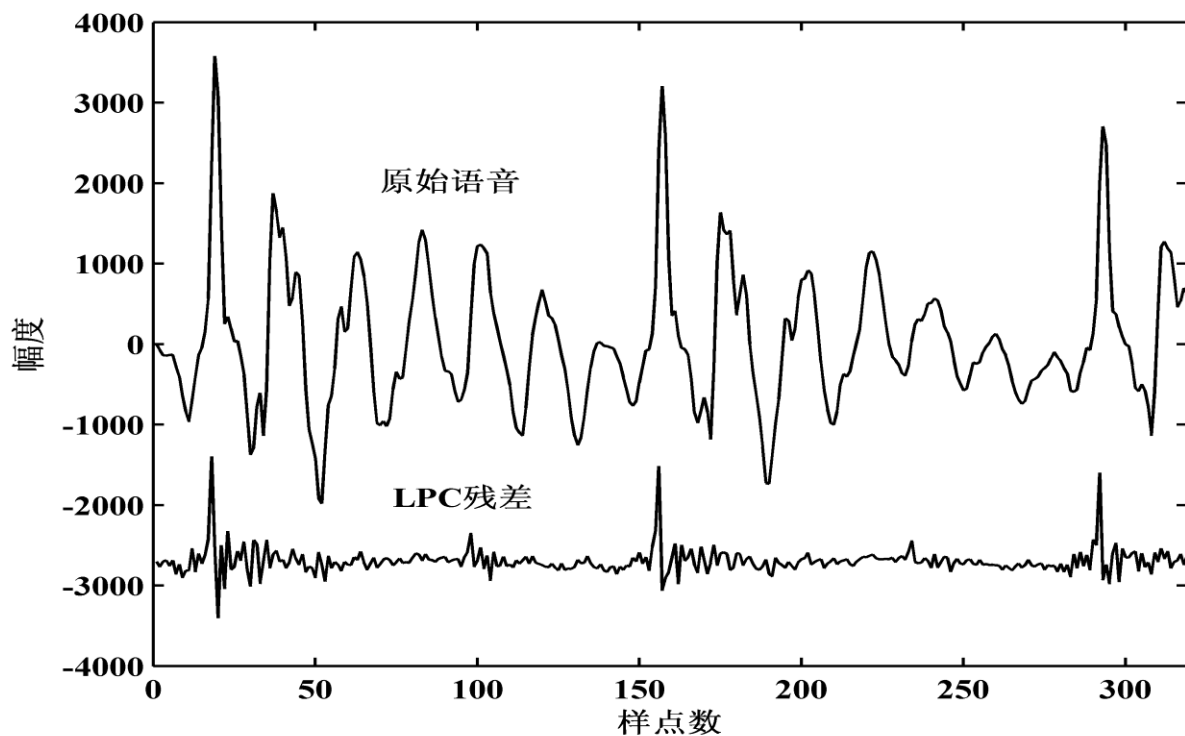
a_n 自相关函数为

$$R_a(j) = \sum_{n=0}^{p-j} a(n)a(n+j), \quad j = 1, 2, \dots, p \quad (5.64)$$



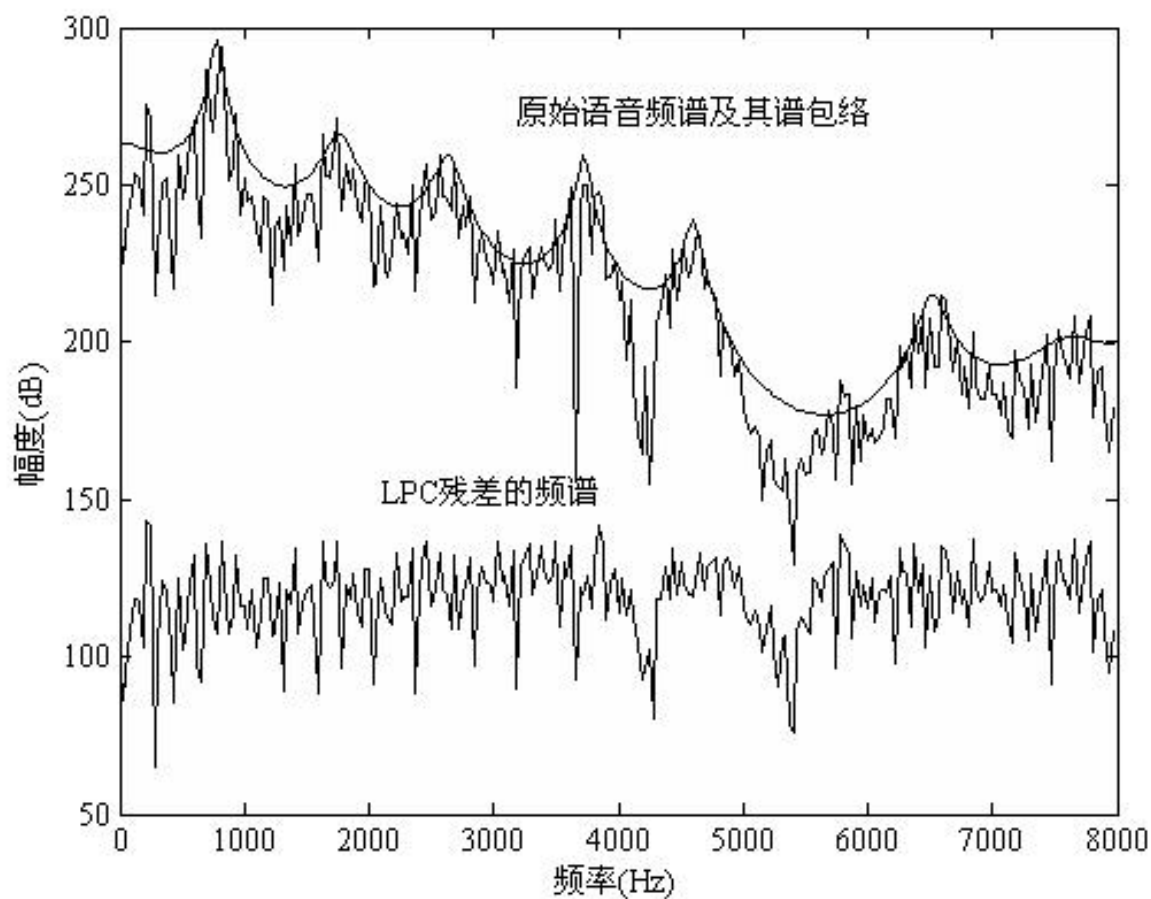


下图为原始语音信号和经过LPC逆滤波语音的时域波形图。



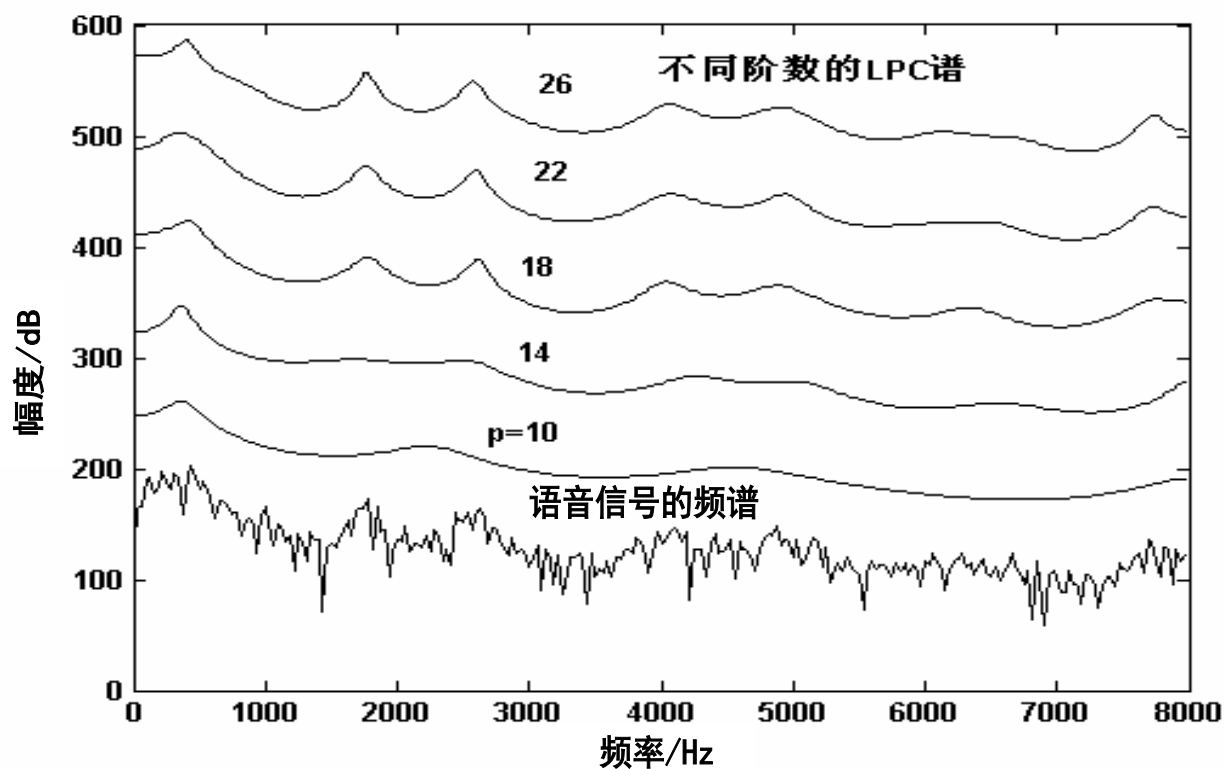


下图为原始语音信号和经过LPC逆滤波语音的频域波形图。





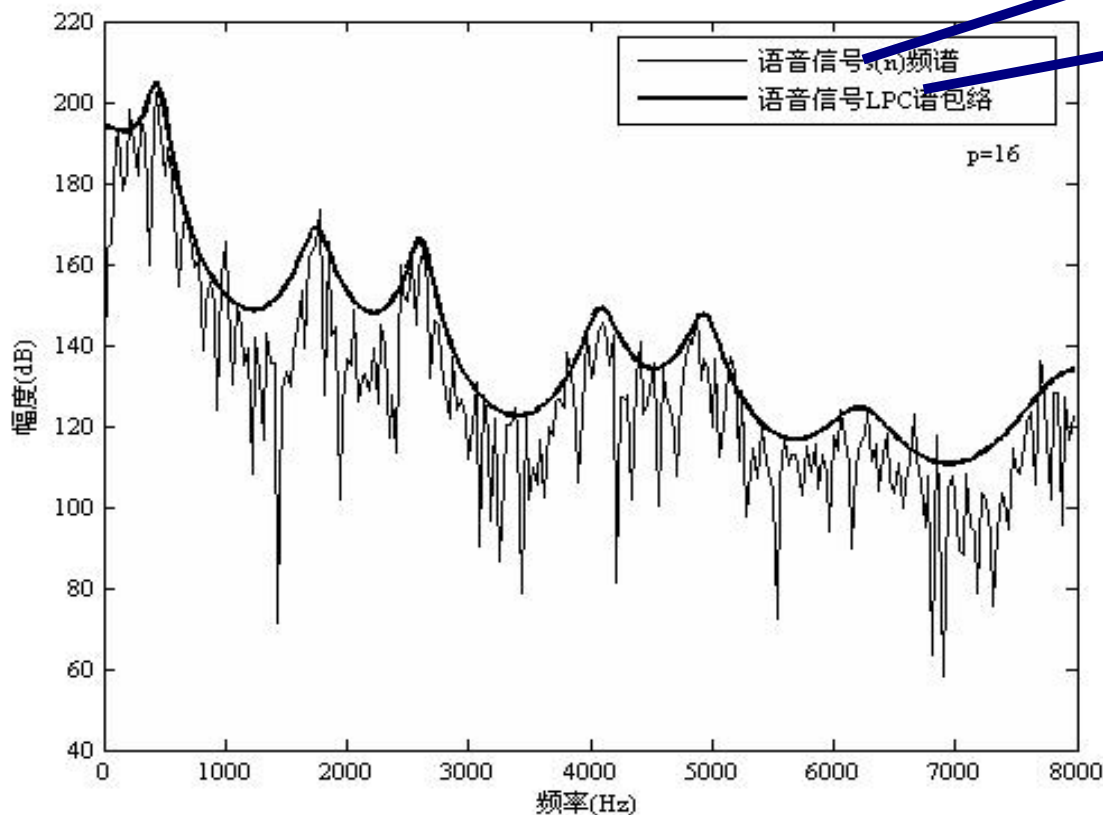
线性预测分析的阶数 p 可有效控制所得谱的平滑度：





LPC谱估计

用线性预测谱作为语音信号谱的解释



$$20\lg|S(e^{j\omega})|$$

$$20\lg|H(e^{j\omega})|$$

图6.8 LPC谱与实际谱的比较



谢谢





LPC分析的频域解释

由于语音产生模型中全极点滤波器的频率特性主要反映了声道的共振特性，而语音信号的LPC系数就是语音信号产生模型中全极点合成滤波器 $H(z)$ 的分母多项式的系数，因此当根据一帧语音的取样值计算出语音信号的LPC系数后，只要将 $z = e^{j\omega}$ 代入 $H(z)$ 进行计算，就意味着求得了这帧语音信号产生模型的频率特性。



$$H(e^{j\omega}) = \frac{G}{1 - \sum_{i=1}^p a_i e^{-j\omega i}} = \frac{G}{A(e^{j\omega})}$$

LPC分析可以看成是对语音信号短时谱进行估计的一种有效方法。

在语音产生模型中，语音的功率谱等于激励源功率谱与全极点合成滤波器频率特性模的平方的乘积，而激励源是准周期冲击序列或白噪声，其功率谱是平坦的。所以语音的功率谱主要由全极点滤波器的特性来决定。



最小预测误差的频域解释

由均方预测误差 $E[e^2(n)]$ 及Parseval定理知均方预测误差的频域表示式即功率谱为

$$\begin{aligned} E_p &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |E(e^{j\omega})|^2 d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |S(e^{j\omega})|^2 |A(e^{j\omega})|^2 d\omega = \frac{G^2}{2\pi} \int_{-\pi}^{\pi} \frac{|S(e^{j\omega})|^2}{|H(e^{j\omega})|^2} d\omega \end{aligned}$$

$S(e^{j\omega})$: 语音 $s(n)$ 的傅里叶变换

上式表明, 使 E_p 为最小等效于使语音的能量谱对比值的积分为最小。



LPC分析在频域上可理解为：给定语音信号的谱 $|S(e^{j\omega})|^2$ ，期望用一个 p 阶全极点滤波器作为其模型，该模型输出的谱 $|H(e^{j\omega})|^2$ 使比值 $|S(e^{j\omega})|^2 / |H(e^{j\omega})|^2$ 的积分最小。

当 $p \rightarrow \infty$ 时， $\lim_{p \rightarrow \infty} |H(e^{j\omega})|^2 = |S(e^{j\omega})|^2$

上式表明，如果 p 足够大，则我们就能以任意小的误差用全极点模型来逼近信号谱。



注意：

即使 $p \rightarrow \infty$ 时，

$$H(e^{j\omega}) = S(e^{j\omega})$$

不一定成立，即模型谱的频率响应不一定等于信号的傅里叶变换，因为 $S(e^{j\omega})$ 不一定是最小相位的，而 $H(e^{j\omega})$ 必定是最小相位的。

原因： $H(e^{j\omega})$ 是一个全极点滤波器的转移函数，其极点应全部位于单位圆内。



$$\begin{aligned} E_p &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |E(e^{j\omega})|^2 d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |S(e^{j\omega})|^2 |A(e^{j\omega})|^2 d\omega = \frac{G^2}{2\pi} \int_{-\pi}^{\pi} \frac{|S(e^{j\omega})|^2}{|H(e^{j\omega})|^2} d\omega \end{aligned}$$

波形匹配规律解释：由上式知，按最小均方误差求解时， $|S(e^{j\omega})| > |H(e^{j\omega})|$ 的区域在总误差中所起的作用比 $|S(e^{j\omega})| < |H(e^{j\omega})|$ 的区域大。因此LPC谱误差准则有利于在谱峰附近良好匹配，而在谱谷附近匹配较差。