

第6章：功率谱估计的现代方法

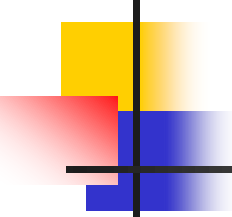
预备知识：经典谱估计

功率谱的一个等价定义是

$$P(f) = \lim_{M \rightarrow \infty} E \left\{ \frac{1}{2M+1} \left| \sum_{n=-M}^M x(n) e^{-j2\pi f n} \right|^2 \right\}$$

在只有观测数据集 $\{x(0), x(1), \dots, x(N-1)\}$ 的情况下，
周期图的定义

$$\hat{P}_{per}(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n) e^{-j2\pi f n} \right|^2$$



周期图谱估计的性能要点:

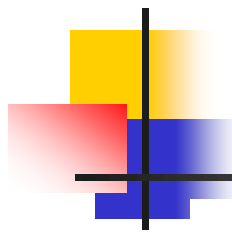
(1) 估计器均值

周期图谱估计器的均值表示为

$$E\{\hat{P}_{per}(f)\} = \int_{-1/2}^{1/2} W_B(f - \lambda) P(\lambda) d\lambda$$

这里, $W_B(f)$ 是三角窗 (也称 Bartlett 窗) 的 DTFT,

由窗函数的性质得到 $\lim_{N \rightarrow \infty} E\{\hat{P}_{per}(f)\} = P(f)$ 渐进无偏



(2) 估计器的方差

周期图谱估计器的方差近似为

$$\text{Var}\{\hat{P}_{per}(f)\} \approx P^2(f) \left[1 + \left(\frac{\sin(2\pi Nf)}{N \cdot \sin(2\pi f)} \right)^2 \right]$$

对于频率 $f \neq 0, \pm 1/2$ ，和较大的 N 值，方差近似为

$$\text{Var}\{\hat{P}_{per}(f)\} \approx P^2(f)$$

周期图谱估计器的方差不随记录数据长度 N 而减小，而是趋于常数，因此它不是功率谱的一致估计。



加窗：

加窗，降低谱估计的旁瓣的能量，不改善方差特性。

加窗的周期图估计器为

$$\hat{P}_{per}(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n)w(n)e^{-j2\pi fn} \right|^2$$

不管是否采用显式加窗的周期图估计器，实质上都已隐含了加窗，如果没有特殊的窗函数被采用，实际上是加了矩形窗，时域加窗限制了周期图估计器的分辨率。周期图估计器的能够分辨出的频率间隔为 $O(1/N)$ 。



平均周期图估计器

$$\hat{P}_{per}^{(m)}(f) = \frac{1}{L} \left| \sum_{n=0}^{L-1} x_m(n) e^{-j2\pi f n} \right|^2$$

$$\hat{P}_{avper}(f) = \frac{1}{K} \sum_{m=0}^{K-1} \hat{P}_{per}^{(m)}(f)$$

如果各段数据近似是不相关的，估计器的方差为

$$\text{Var}\{\hat{P}_{avper}(f)\} = \frac{1}{K} \text{Var}(\hat{P}_{avr}^{(m)}(f)) \approx \frac{1}{K} P^2(f)$$

平均周期图估计器的能够分辨出的频率间隔变成 $O(1/L)$
频率分辨率下降了。



周期图的快速计算

画出在离散频率点 $f_k = k / N, (\omega_k = 2\pi k / N)$ 上的谱图

$$\hat{P}_{per}(f_k) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn / N} \right|^2$$

通过 FFT 进行高效计算。如果希望得到在更密集的

$f_k = k / N', (\omega_k = 2\pi k / N'), N' > N$ 频率点上的谱图，

可以通过给观测数据尾部补零得到更细致的图谱，令

$$x'(n) = \begin{cases} x(n), & n = 0, 1, \dots, N-1 \\ 0, & n = N, N+1, \dots, N'-1 \end{cases}$$



Blackman-Tukey (BT) 方法

对维纳-欣钦定理的加窗实现，通过观测的 N 各数据，可以估计自相关序列的如下值。

$$\hat{r}_x(l) = \begin{cases} \frac{1}{N} \sum_{n=0}^{N-l-1} x(n+l)x^*(n) & 0 \leq l \leq N-1 \\ \hat{r}_x^*(-l) & -(N-1) \leq l < 0 \\ 0 & |l| \geq N \end{cases}$$

由于估计的自相关序列在序号较大时是不可靠的，通过加窗只取其中一部分用于谱估计，以改善谱估计的可靠性。



窗函数的定义满足如下 3 条：

$$0 \leq w(k) \leq w(0) = 1$$

$$w(-k) = w(k)$$

$$|k| > M, w(k) = 0$$

这个窗实际只保留 $|k| \leq M$ 的那些自相关值，一般常取 $M \leq N/5$

定义 BT 功率谱估计器为

$$\hat{P}_{BT}(f) = \sum_{k=-M}^M w(k) \hat{r}_x(k) e^{-j2\pi f k}$$



Blackman—Tukey (BT) 方法的特点

- (1) 在 $M = N - 1$ ，窗函数为矩形窗时，BT 谱估计等于周期图谱估计。
- (2) BT 估计器能够分辨出的频率间隔为 $O(1/M)$ 估计器的均值为

$$E\{\hat{P}_{BT}(f)\} \approx \int_{-1/2}^{1/2} W(f - \lambda) P(\lambda) d\lambda$$

- (3) BT 估计器的方差为

$$Var\{\hat{P}_{BT}(f)\} \approx \frac{1}{N} P^2(f) \sum_{k=-M}^M w^2(k)$$



例：如果取 Bartlett 窗和 $M = N/5$ ，得到

$$\text{Var}\{\hat{P}_{BT}(f)\} \approx \frac{2M}{3N} P^2(f) = \frac{1}{7.5} P^2(f)$$

与周期图相比，降低了谱分辨率情况，也降低了估计器的方差。



AR模型法谱估计

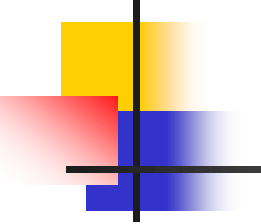
假设一个随机过程可以由AR(p)刻画

$$x(n) = -\sum_{k=1}^p a^*(k)x(n-k) + v(n)$$

它的功率谱为

$$P_{AR}(f) = \frac{\sigma^2}{\left| 1 + a^*(1)e^{-j2\pi f} + \dots + a^*(p)e^{-j2\pi fp} \right|^2}$$

$$\sigma^2 = E[|v(n)|^2]$$



给出一组观测数据 $\{x(0), x(1), \dots, x(N-1)\}$,

得到估计的参数集 $\{\hat{a}(1), \hat{a}(2), \dots, \hat{a}(p), \hat{\sigma}^2\}$,

得到一个估计的功率谱密度 PSD。 ,

$$\hat{P}_{AR}(f) = \frac{\hat{\sigma}^2}{\left| 1 + \sum_{k=1}^p \hat{a}^*(k) e^{-j2\pi f k} \right|^2}$$

最大熵谱估计 (MESE)

假设已知 $\{r(0), r(1), \dots, r(p)\}$ ，为了确定 PSD，
外推 $r(p+1), r(p+2), \dots$ ，一种原则是使信号熵最大，
即有最大随机性。

对于高斯过程，熵可以表示成：

$$C \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln P_{xx}(f) df$$

由已知 $p+1$ 个自相关值构成如下约束方程：

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} P_{xx}(f) e^{j2\pi fk} df = r(k) \quad k = 0, 1, \dots, p$$

且知：

$$P_{xx}(f) = \sum_{k=-\infty}^{+\infty} r(k) e^{-j2\pi fk}$$

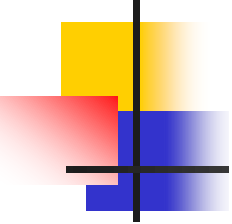
用 Lagrangian 乘积法构成目标函数。

$$S = \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln P_{xx}(f) df + \sum_{k=0}^p \lambda_i \left(\int_{-\frac{1}{2}}^{\frac{1}{2}} P_{xx}(f) e^{j2\pi f k} df - r(k) \right)$$

并且求： $\frac{\partial S}{\partial r(k)} = 0 \quad |k| = p+1, p+2, \dots$

推导得： $\int_{-1/2}^{1/2} \frac{e^{-j2\pi f k}}{P_{xx}(f)} df = 0 \quad |k| \geq p+1$

这隐含着： $\frac{1}{P_{xx}(f)} = \sum_{k=-p}^p c_k e^{-j2\pi f k}$



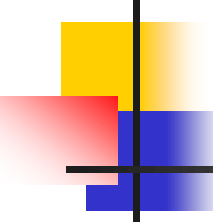
并且 $c_k^* = c_{-k}$ 以确保 $p_{xx}(f)$ 是实的。即求得:

$$P_{xx}(f) = \frac{1}{\sum_{k=-p}^p c_k e^{-j2\pi f k}}$$

上式带回 $p+1$ 个约束方程, 经过整理, 最后求得:

$$P_{xx}(f) = \frac{\sigma^2}{\left| 1 + \sum_{k=1}^p a^*(k) e^{-j2\pi f k} \right|^2}$$

这里 σ^2 和 $a(k)$ 满足方程：↵

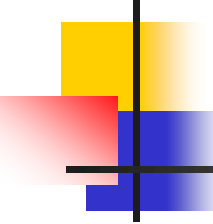

$$R \begin{pmatrix} a(1) \\ a(2) \\ \vdots \\ a(p) \end{pmatrix} = - \begin{pmatrix} r^*(1) \\ r^*(2) \\ \vdots \\ r^*(p) \end{pmatrix}$$

和：↵

$$\sigma^2 = r(0) + \sum_{k=1}^p a(k)r(k)$$

这正是 Yule-Walker 方程，由此得到结论：
最大熵估计和 AR 谱估计是一致的。↵

AR模型谱估计的协方差方法



★ 观察数据窗是 $[0, N-1]$ ，将这组数据分成初始数据。

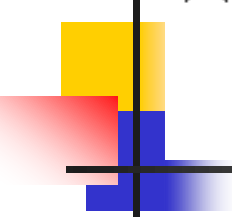
$$\mathbf{x}_0 = [x(0), x(1), \cdots, x(p-1)]^T$$

数据 $\mathbf{x} = [x(p), x(p+1), \cdots, x(N-1)]^T$ 。

★ 前向预测系数与 $AR(p)$ 参数是一致的，用求解前向预测的方法求 $AR(p)$ 的参数。预测误差为

$$f_p(n) = x(n) - \hat{x}(n | X_{n-1})$$

$$= x(n) + \sum_{k=1}^p a^*(k)x(n-k)$$



★ 对观测窗外数据不作假设，相当于没有加窗概念。
的引入，预测误差只存在 $n = p$ 至 $n = N - 1$ 之间。

★ 预测误差和

$$\xi = \sum_{n=p}^{N-1} |f_p(n)|^2$$

求解系数集 $\{a(1), a(2), \dots, a(p)\}$ ，使预测误差和最小，即

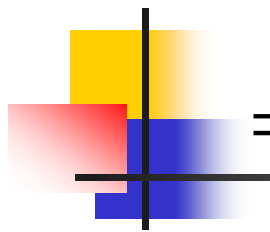
$$\xi = \sum_{n=p}^{N-1} \left| x(n) + \sum_{k=1}^p a^*(k)x(n-k) \right|^2$$

最小，这是典型的 LS 问题。



★确定数据矩阵


$$\mathbf{f}^* = \begin{bmatrix} f_p^*(p) \\ f_p^*(p+1) \\ \vdots \\ f_p^*(N-1) \end{bmatrix} = \begin{bmatrix} x^*(p) \\ x^*(p+1) \\ \vdots \\ x^*(N-1) \end{bmatrix} + \begin{bmatrix} \sum_{k=1}^p a(k)x^*(p-k) \\ \sum_{k=1}^p a(k)x^*(p+1-k) \\ \vdots \\ \sum_{k=1}^p a(k)x^*(N-1-k) \end{bmatrix}$$



$$= \begin{bmatrix} x^*(p) \\ x^*(p+1) \\ \vdots \\ x^*(N-1) \end{bmatrix} + \begin{bmatrix} \mathbf{x}^H(p-1) \\ \mathbf{x}^H(p) \\ \vdots \\ \mathbf{x}^H(N-2) \end{bmatrix} \mathbf{a} = \mathbf{x}_f + A\mathbf{a}$$

$$A^H = [\mathbf{x}(p-1), \mathbf{x}(p), \dots, \mathbf{x}(N-2)]$$

$$= \begin{bmatrix} x(p-1) & x(p) & \dots & x(N-2) \\ x(p-2) & x(p-1) & \dots & x(N-3) \\ \vdots & \vdots & \vdots & \vdots \\ x(0) & x(1) & \dots & x(N-p-1) \end{bmatrix}$$

 $AR(p)$ 模型系数的 LS 解为

$$\hat{\mathbf{a}} = -(A^H A)^{-1} A^H \mathbf{x}_f = -\left(\frac{1}{N-p} A^H A\right)^{-1} \left(\frac{1}{N-p} A^H \mathbf{x}_f\right)$$

LS 的最小误差和为

$$\xi_{\min} = \xi_x + \hat{\mathbf{a}}^H A^H \mathbf{x}_f = \mathbf{x}_f^H \mathbf{x}_f + \hat{\mathbf{a}}^H A^H \mathbf{x}_f$$

$AR(p)$ 模型激励白噪声的方差估计为

$$\hat{\sigma}_v^2 = \frac{1}{N-p} \xi_{\min} = \frac{1}{N-p} (\mathbf{x}_f^H \mathbf{x}_f + \hat{\mathbf{a}}^H A^H \mathbf{x}_f)$$

算法 1:

(1) 确定模型阶 p ，由观察数据 $\{x(0), x(1), \dots, x(N-1)\}$ 构成数据矩阵。

(2) 对 $(A^H A)$ 进行特征分解，得到 W 个不为零的特征值 $\sigma_1^2, \sigma_2^2, \dots, \sigma_W^2$ ，和部分特征矩阵 $V_1 = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_W]$ 。

$AR(p)$ 的系数表示为： $\hat{\mathbf{a}} = V_1 \Sigma^{-2} V_1^H A^H \mathbf{x}_f$ ，并求出 $\hat{\sigma}_v^2$

(3) 带入公式

$$\hat{P}_{AR}(f) = \frac{\hat{\sigma}_v^2}{\left| 1 + \sum_{k=1}^p \hat{a}(k) e^{-j2\pi f k} \right|^2}$$

由要求的显示精度和频率范围，计算 $f_i = i\Delta f$ 处的 PSD 值，并画出图形。

AR 模型谱估计的协方差方法的参数估计性能

假设： 1) 数据长度很长。 2) \mathbf{x} 是高斯的

AR 模型协方差方法是参数的 MLE 估计。

当数据长度 $N \gg p$ 时，估计 $\hat{\mathbf{a}}$ 和 $\hat{\sigma}^2$ 的 CR 下界为

$$C_{\mathbf{a}, \sigma^2} = E \left\{ \begin{pmatrix} \hat{\mathbf{a}} - E(\hat{\mathbf{a}}) \\ \hat{\sigma}^2 - E(\hat{\sigma}^2) \end{pmatrix} \begin{pmatrix} \hat{\mathbf{a}} - E(\hat{\mathbf{a}}) \\ \hat{\sigma}^2 - E(\hat{\sigma}^2) \end{pmatrix}^T \right\} = \begin{pmatrix} \frac{\sigma^2}{N} R_x^{-1} & 0 \\ 0 & \frac{2\sigma^4}{N} \end{pmatrix}.$$

且是渐进无偏的，估计的参数符合高斯分布。

AR 模型谱估计的协方差方法的功率谱密度估计性能

功率谱估计也是一致估计，当 $N \rightarrow \infty, p \rightarrow \infty$ 时，PSD 估计的均值、方差、不同频率点的互协方差如下：

$$E[\hat{P}_{AR}(f)] = P_{AR}(f)$$

$$\text{Var}[\hat{P}_{AR}(f)] = \begin{cases} \frac{4p}{N} P_{AR}^2(f), & f = 0, \pm 1/2 \\ \frac{2p}{N} P_{AR}^2(f), & \text{Otherwise} \end{cases}$$

$$\text{Cov}[\hat{P}_{AR}(f_1), \hat{P}_{AR}(f_2)] = 0, \quad f_1 \neq f_2$$

AR 功率谱估计随 N / p 增加得到改善



改进协方差方法

利用前向、后向预测误差平均最小，得到 $AR(p)$ 模型参数的解。
前向预测误差和写为↵

$$\xi^f = \sum_{n=p}^{N-1} |f(n)|^2 = \sum_{n=p}^{N-1} \left| x(n) + \sum_{i=1}^p a(i)x(n-i) \right|^2 \quad \leftarrow$$

后向预测误差和为↵

$$\xi^b = \sum_{n=p}^{N-1} |b(n)|^2 = \left[\sum_{n=p}^{N-1} \left| x(n-p) + \sum_{i=1}^p a(i)x(n-p+i) \right|^2 \right]$$

问题的解归结为使前向和后向预测误差和的平均最小↵



自相关方法

最大时间范围 $[0, N + p - 1]$, 计算该范围的预测误差和,

$$\xi = \sum_{n=0}^{N+p-1} \left| x(n) + \sum_{k=1}^p a(k)x(n-k) \right|^2$$

假设 $x(n) = 0$, 除非 $0 \leq n \leq N - 1$ 。

Burg 算法（间接参数估计，直接反射系数估计）

Levinson 递推算法:若已知初始 $P_0 = r_{xx}(0)$ ，如果得到反射系数 $\{k_1, k_2, \dots, k_p\}$ ，则可以递推地得到最终 P 阶的 AR 参数。

$$\{a_p(1), a_p(2), \dots, a_p(p), P_p = \sigma^2\}$$

假设 $m-1$ 阶各量均已求得，估计反射系数 k_m ，使得

$$\rho_m = \frac{1}{2}(\rho_m^f + \rho_m^b)$$

$$= \frac{1}{2(N-m)} \left[\sum_{n=m}^{N-1} |\hat{f}_m(n)|^2 + \sum_{n=m}^{N-1} |\hat{b}_m(n)|^2 \right]$$

最小。

利用格型预测误差滤波器的递推关系，

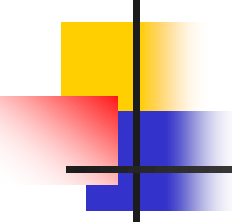
$$\hat{f}_m(n) = \hat{f}_{m-1}(n) + k_m^* \hat{b}_{m-1}(n-1)$$

$$\hat{b}_m(n) = \hat{b}_{m-1}(n-1) + k_m \hat{f}_{m-1}(n)$$

带入上式，注意到未知数只有 k_m 并令

$$\frac{\partial \rho_m}{\partial k_m} = 0, \text{ 求得}$$

$$\hat{k}_m = \frac{-2 \sum_{n=m}^{N-1} \hat{f}_{m-1}^*(n) \hat{b}_{m-1}(n-1)}{\sum_{n=m}^{N-1} (|\hat{f}_{m-1}(n)|^2 + |\hat{b}_{m-1}(n-1)|^2)}$$



结合 Levinson 递推算法，和新的反射系数公式，总结 Burg 算法如下

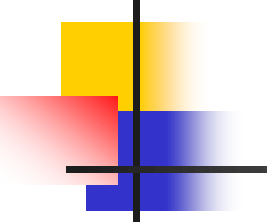
(1) 初始值: \leftarrow

$$\hat{r}_{xx}(0) = \frac{1}{N} \sum_{n=0}^{N-1} |x(n)|^2 \quad \hat{P}_0 = \hat{r}_{xx}(0) \leftarrow$$

$$\hat{f}_0(n) = x(n) \quad n = 1, 2, \dots, N-1 \leftarrow$$

$$\hat{b}_0(n) = x(n) \quad n = 0, 1, 2, \dots, N-2$$

(2) 对 $m = 1, 2, \cdots p$, 递推



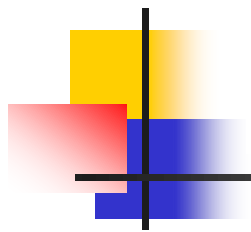
$$k_m = \frac{-2 \sum_{n=m}^{N-1} \hat{f}_{m-1}^*(n) \hat{b}_{m-1}(n-1)}{\sum_{n=m}^{N-1} (|\hat{f}_{m-1}(n)|^2 + |\hat{b}_{m-1}(n-1)|^2)}$$

$$\hat{P}_m = (1 - |k_m|^2) \hat{P}_{m-1}$$

$$\hat{a}_m(i) = \begin{cases} \hat{a}_{m-1}(i) + \hat{k}_m \hat{a}_{m-1}^*(m-i) & i = 1, 2, \cdots, m-1 \\ \hat{k}_m & i = m \end{cases}$$

$$\hat{f}_m(n) = \hat{f}_{m-1}(n) + k_m^* \hat{b}_{m-1}(n-1), \quad n = m+1, \cdots, N-1$$

$$\hat{b}_m(n) = \hat{b}_{m-1}(n-1) + k_m \hat{f}_{m-1}(n), \quad n = m, \cdots, N-2$$




(3) 得到 $AR(p)$ 模型系数 $\{\hat{a}(i), i = 1, 2, \dots, p\}$ 和 $\hat{\sigma}_v^2 = \hat{P}_p$

将估计参数带入公式。

$$\hat{P}_{AR}(f) = \frac{\hat{\sigma}_v^2}{\left| 1 + \sum_{k=1}^p \hat{a}(k) e^{-j2\pi f k} \right|^2}$$

根据所要求的显示精度和频率范围，计算 $f_i = i\Delta f$ 处的 PSD 值，并画出图形。



几种方法比较

对如上几种方法的主要优缺点列于下。

- ◆ 自相关方法可用 Levinson 快速算法，运算量小，
但分辨率受窗长度限制。
- ◆ 自协方程方法，分辨率高，运算量较大。
- ◆ 改进自协方差方法，分辨率高，无谱线分裂和偏移，运算量大。
- ◆ Burg 算法，可用改进的 Levinson 递推算法，分辨率高，
但有谱线分裂和偏移（对正弦信号）。

自相关方法总有稳定的解

自协方差方法在极少情况下有不稳定解

Burg 算法在数据很短情况下不及自协方差方法

在数据很长时，这些方法趋于一致



模型阶选择

最终预测误差准则：↵

$$FPE(k) = \frac{N+k}{N-k} \hat{P}_k$$

Akaike 信息准则：↵

$$AIC(k) = N \ln \hat{P}_k + 2k \quad \leftarrow$$

当 $k=p$ 时，以上准则为最小，则确定为 p 阶



其他模型谱估计方法

MA模型谱估计：Durbin方法（自学）

ARMA模型谱估计：改进Yule-Walker方程（自学）