信号处理的数学方法
aka 数字信号与图像处理

Mathematical Methods in Signal Processing

张思容
zhangsirong@buaa.edu.cn

数学与系统科学学院，北京航空航天大学
School of Mathematics and Systems Science, Beihang University

May 1, 2012

Chapter 4: 随机信号处理
Random signal processing

随机信号滤波
- 随机信号系统
- 点估计
- 最佳线性波形估计：Wiener 滤波器

谱估计
- 谱分解
- 经典谱估计
- 现代谱估计:ARMA模型

参考书：
胡广书：数字信号处理：理论，算法与实现(第二版)，清华大学出版社。
G.Strang: Computational Science and engineering:

真实世界的信号处理

真实信号=理想信号+噪音信号$=x(t)+e(t)$

- 理想信号：函数空间表示；L^2
  噪音信号: 最简单是白噪音随机过程(完全不可预期)；经过LTI系统得到平稳随机过程;
  （其他噪音模型？：电线交流电噪音60赫兹，图像的椒盐噪音;）
- 统计信号处理: 信号的平均的统计特征是确定的,可以利用数据统计量估计信号的统计特征。
  主要内容： 信号分析(谱估计) $\to$ 自相关函数，功率谱
  信号处理(线性滤波Wiener滤波器，最小二乘，自适应滤波，Kalman滤波). $\to$ 信号估计，最小二乘法

随机变量与离散随机过程

Remark (随机变量)

- CDF分布函数：$F_x(a)=Pr(x(\xi)<a)$ PDF密度函数 $f_x(a)=F'(a)$
- 数字特征： $EX, Var(X)=DX, EX^2, EX^3, \dots$
- 大样本与大数定律,中心极限定理;
- 估计：$\mu=1/N\sum_i X_i \to EX$,
  $s=1/(N-1)\sum_i (X_i-\mu)^2 \to DX$.

Remark (离散随机过程)

- 联合概率分布：$F(x_1,\dots,x_M;n_1,\dots,n_m)=Pr(x(n_i)\le x_i, 1\le i\le M$
- 统计特征:$\mu(n)=E(x(n))$, $s(n)$;
  自相关矩阵：$r(n_1,n_2)=E(x(n_1)x(n_2)$
  自协方差矩阵：$\gamma(n_1,n_2)=r(n,n_2)-\mu(n_1)\mu(n_2)$
- 宽平稳过程：$\mu(n)=\mu, r(n_1,n_2)=r(|n_1-n_2|)$.

平稳信号自相关序列

Proposition

- $r_x(0) = \sigma_x^2 + |\mu_x|^2 \geq r_x(k)$
- $r_x(k) = r_x^{(} - k)$
- 非负定 $\sum_k \sum_m a_k r(k-m) a_m^* \geq 0$

证明：构造利用$E(x(n) \pm X(m)^2 \geq 0$,详略。

Remark (实用的统计估计条件)

- 遍历性:有限时间的平均统计的极限等于期望值
  相关遍历性定理：依赖与期望和二阶矩的有限性条件；
- 遍历定理：
  $E(x(n)) = \lim_N \frac{1}{2N+1} \sum_{i=-N}^{N} x(i)$
  $E(x(n)x(n-k)) = \lim_N \frac{1}{2N+1} \sum_{i=-N}^{N} x(i)x(i-k)$

功率谱的计算

- 直接法：P等于能量(谱的平方和)的时间平均;
- 间接法：P等于自相关函数的DTFT；
- 特别：假定时间遍历的功率谱有：
  $P(e^{iw}) = \lim_M E(\frac{1}{2M+1}|\sum_{n=-M}^{n=M} x(n)e^{-iwn}|^2)$

Theorem (Wiener-Khintchine)

$P(e^{iw}) = DTFT(r_x(k)) = \sum_k r(k)e^{-iwk}$.

证明： $P(e^{iw}) = \lim_M E(\frac{1}{2M+1}|\sum_{n=-M}^{n=M} x(n)e^{-iwn}|^2)$，展开求期望得到$r(k)$,再求和，求极限即可。
详略。

LTI系统作用于平稳信号

Theorem (存在性)

给定$x(n,\xi)$是有限期望的平稳随机过程，通过$BIBO$稳定LTI系统$h[k]$,得到 $y(n,\xi) = \sum_k h[k]x[n-k;\xi]$
$y(n,\xi)$按概率*1*收敛，且如果$x(n,\xi)$方差有限，则输出也是方差有限，且它是平稳随机过程。

LTI系统的系统响应:时间域

- 输出期望$\mu(y[n]) = \mu(x[n])H(e^{i0})$
- 互相关系数$r_{xy}(k) = h(-k) * r_{xx}(k)$
- 自相关系数$r_y(k) = r_h(k) * r_x(k)$,
  其中$r_h(k) = h(n) * h(n-k)$，为系统的相关序列。 特别输出功率$P_y = r_y(0) = \sum_k r_h(k) r_x(k)$.

LTI系统的系统响应: 频率域

- z变换 $S_y(z) = H(z)H(1/z)S_x(z)$
- 自功率谱 $S_y(e^{i\omega}) = |H(e^{i\omega})|^2 S_x(e^{i\omega})$
- 自功率谱只能得到系统的幅度响应，不能得到相位响应。

估计问题

给定数据$x(0), x(1), \ldots, x(n-1)$，利用随机变量估计确定参数$\theta := \hat{\theta} = g(x(0), \ldots, x(n-1))$.

- 估计器的性能？与真实值得差距？
- 是否最优？怎样找到最优估计？

Definition (估计分类)

- 估计器的性能:
  - 无偏估计：$E(\hat{\theta}) = \theta$
  - 渐进无偏：$n \to \infty$, $\lim E(\hat{\theta}) = \theta$
  - 有偏估计：$b(\theta) = E(\hat{\theta}) - \theta$
- 最优准则:
  均方误差准则*(MSE)*: $mse(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$
  一致估计：$n \to \infty, mse(\hat{\theta}) \to 0$.

理想估计器：最小方差无偏估计器(MVU)

矩估计

- 无偏均值估计: $\hat{\mu}_x = \frac{1}{N}\sum x(n)$
  白噪声过程方差: $Var(\hat{\mu}_x) = \sigma_x^2/N$
  一般过程:$Var(\hat{\mu}_x) = \frac{1}{N}\sum(1-\frac{|k|}{N})c_x(k)$,$c_x$为$x$的协方差函数。
- 有偏方差估计: $\hat{\sigma}_x^2 = \frac{1}{N}\sum(x(k)-\hat{\mu}_x)^2$
  白噪声过程方$E(\hat{\sigma}_x^2) = \frac{N}{N-1}\sigma_x^2$, 渐近无偏。
  一般过程:$Var(\hat{\sigma}_x^2) \approx \frac{c_x^4}{N}$, c_x^4是x的四阶中心矩。
- 自相关估计
  $$\hat{r}_x(k) = \begin{cases} \frac{1}{N}\sum_{n=1}^{N-1} x(n)x^*(n-k) & 0 \le k \le N-1 \\ \hat{r}_x(-k) & -(N-1) \le k < 0 \\ 0 & \text{otherwise.} \end{cases}$$
  渐进无偏; 自相关矩阵是半正定; 方差趋于零(一致估计);

EXAMPLE (简单随机过程)

$x(n) = A + w(n)$,$w(n)$是$WGN(0,\sigma^2)$白噪声。估计A.

线性波形估计

EXAMPLE

已知信号$x_1(n), x_2(n), \ldots, x_k(n)$,估计信号$y(n)$.
设$\hat{y(n)} = H(x_1, x_2, \ldots, x_k)$为一个估计信号。
误差信号$e(n) = y(n) - \hat{y}(n)$.
特别有线性估计, H是线性函数。

Remark (最佳估计器)

误差$e(n)$满足一定准则的估计器。
一般正负同样重要, $|e(n)|$, $|e(n)|^\alpha$.
重要:$MMSE$最小均方平方误差$L(n) = E[|e(n)|]^2$.

Gauss 19世纪发展;
Wiener和Kolmogorov 最佳滤波; Kalman等进一步发展;

例子

EXAMPLE (array signals)

阵列信号处理: $x_k(n)$是空间信号, $y(n)$是某个方向待估计信号。

EXAMPLE (信道均衡)

$y(n) = s(n) + e(n)$,构造滤波器$\hat{y}(n) = y(n) * h(n) = s(n)$.

EXAMPLE (linear predictor)

线性预测: $y(n) = x(n)$,由$x(n-1), x(n-2), \ldots, x(n-k)$估计。
向前预测及向后预测。

EXAMPLE (inverse system)

逆系统: LTI系统 $\hat{y}(n) = h(n) * x(n)$,误差$e(n) = y(n) - \hat{y}(n)$。

线性均方误差估计

Definition (LMMSE)

设$\hat{y}[n] = \sum c_k[n]x_k[n]$,即$\hat{Y} = C^\tau X$,即$X$是数据向量$x_1, x_2, \ldots, x_k$,$C$是系数向量。
误差曲面: 误差准则$P = E(|e|^2) = E(|\hat{y}-y|^2)$,
$P(C) = E((C^\tau X - y)^2) = E|y|^2 - C^\tau d - d^\tau C + C^\tau RC$, 其中$d = E(xy)$, $R = E(xx^\tau$,$P(C)$称为估计的误差性能曲面。

当R正定时有唯一最小解。 如果使用有理系统, 误差性能曲面可能非二次曲面。

Remark (最小解的求法)

以上曲面方程可以直接用偏导数, 但变量太多。 常见有矩阵等对向量和矩阵的求导简化写法。设x, y, b是向量, A是矩阵比如
$\frac{\partial x^\tau b}{\partial x} = b$, $\frac{\partial x^\tau x}{\partial x} = 2x$, $\frac{\partial x^\tau Ay}{\partial x} = Ay$
$\frac{\partial x^\tau Ay}{\partial A} = xy^\tau$, $\frac{\partial x^\tau A^\tau Ax}{\partial A} = 2Axx^\tau$, $\frac{\partial trA}{\partial A} = I$

法方程

Remark (求解过程)

配方法$P(C) = E|y|^2 - C^\tau d - d^\tau C + C^\tau RC$
$P(C) = E(y^2) - d^\tau R^{-1} d + (RC-d)^\tau R^{-1}(RC-d)$, 如果$R$正定，
取得最小值的充要条件$RC_0 = d$; $P_{\min} = P_y - d^\tau C_0$.

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1k} \\ r_{21} & r_{22} & \cdots & r_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ r_{k1} & r_{k2} & \cdots & r_{kk} \end{bmatrix} * \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{pmatrix}; r_{ij} = E(x_i x_j); d_i = E(x_i y)$$

以上称为法方程。

Corollary (正交性)

法方程的解满足：估计误差e_0和x正交。$E(xe_0) = 0$. 特别$E(\hat{y}e_0) = 0$.

Wiener-Hopf 方程

Remark

最优FIR滤波器:利用输入信号$x[n]$的样本，估计期望信号$y[n]$，
特别有线性估计$\hat{y}[n] = \sum_{m=1}^{M} h[n,k]x[n-m] = C^\tau[n]X[n]$
对应LTI系统， $C[n]$与时间无关，是常数!

Proposition (平稳过程最优FIR滤波器)

记R为自相关矩阵（$Toeplitz$），$r_k = E(x[n]x[n-k])$, 互相关矢量$d = [d_i] = (r_{yx}(i))$;设LTI系统系数$H[k]$,
$Wiener$-$Hopf$方程：$\sum_{k=0}^{k=M} h[k]r(n-k) = r_{yx}(n) = d_n$.
$MMSEP(C_0) = P_y - \sum h[k]r_{yx}(k)$

Remark

一般不用最优IIR滤波器, IIR可以由FIR逼近；FIR的方程是线性方程!

最优线性预测滤波器***

- 线性预测: 利用信号其他样本(时间)来估计某一时刻的信号值。
  前向预测：令$y[n] = x[n]$,$x_k[n] = x[n-k], 1 \le k \le M$，
  记$\hat{y} = x(n|X_{M-1})$,
  后向预测:
  令$y[n] = x[n-M]$,$x_k[n] = x[n-M+k], 1 \le k \le M$，
  记$\hat{y} = x(n-M|X_n)$,
- 前向预测方程: R自相关矩阵，$d_k = r_{x_{n-k}y} = r(-k)$;
  Wiener-Hopf方程: $Rh_0 = \vec{r}$
  MMSE $P_0 = r(0) - d^\tau h_0$.
- 后向预测方程: R自相关矩阵，$d_k = r_{x_{n+1-k}y} = r(M+1-k)$;
  Wiener-Hopf方程: $Rh_0 = \vec{r}^B$
  MMSE $P_0 = r(0) - \vec{r}^B h_0$.

滤波器的设计实现

Remark (设计)

1. 求解法方程。得到系数C；
2. 计算MSE $P(C)$,验证满足设计要求；
3. 计算估计值$\hat{y}$；

- 法方程求解可用任何算法，但实用中可用更有效算法；
- 阶数固定算法 vs 阶数递归算法
- 已知信息是二阶矩；特别有时不变系统，二阶矩为常数!

Remark (算法效率)

- $Gauss$消去法:$O(M^3)$
- LDU分解：$O(M^2)$
- $Levinson$递归算法: $O(M^2)$包含所有低阶估计；我们记m阶估计$\hat{y}_m$；对应系数$\vec{C}_m$,矩$\vec{d}_m$.

最小二乘与实用滤波器

- Wiener滤波器是理想滤波器。已知条件：信号的二阶矩R以及信号与目标的相关系数r_{dx}.
- 最小二乘滤波器: 宽平稳过程，已知$x(n), d(n)$，估计R, r_{dx}.
  LS滤波器的法方程为$\hat{R}\vec{C} = \vec{r_d}$，最小能量$E = E_y - \vec{r_d}^{\tau} C$.
  其中$\hat{R} = \vec{X}^{\tau}\vec{X}, \vec{r_d} = \vec{X}^{\tau} d$.
- 自适应滤波器: 非宽平稳过程，已知$x(n), d(n)$,
  时变线性方程$A_t w_t = D_t$，能量(最小二乘)$E = |Aw - D|^2$,
  一般可用迭代算法计算 $w_{t+1} = w_t - \Delta E$，比如Newton方法。
  自适应滤波器是学习过程，收敛足够快，从而可改变滤波器系数，适应不同环境。
- Kalman滤波器：非宽平稳过程，时变线性方程$D_t = A_t w_t + b$， 且有状态方程$w_{t+1} = F w_t + c$,其中b, c是噪音或误差。
  一般可用迭代算法: Kalman滤波器算法(比较繁琐，略去).与自适应滤波器存在对应。

以上算法的实现依赖于应用线性代数中的基本结果。

谱估计概述

Newton提出谱；傅立叶给出傅立叶变换；

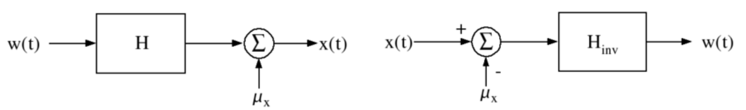
- 傅立叶系数计算A_k, B_k： 谐波分析仪(Thomson),物理机器。
- 经典谱估计：
  直接法： Schuster 周期图： 功率$=A_k^2 + B_k^2$
  改进：平均周期谱；但是： 太慢。
  间接法：Wiener—Khintchine定理：计算自相关函数的傅立叶变换。
  改进：Blackman-Turkey： 自相关谱估计方法
- 现代谱估计：
  参数法模型：ARMA，Yule-Walker 求解
  非参数法：MUSIC声音分解；
  平稳信号与时变信号：短时傅立叶变换，Wigner分布等。

谱分解定理**

Definition (正则随机信号)
如果平稳随机信号满足*Paley-Wiener*条件称为正则的；
即$\int_{-\pi}^{\pi} |\ln S(\omega)| d\omega < \infty$

Theorem (谱分解)**
如果平稳随机信号是正则的，必有分解
$S(z) = \sigma^2 Q(z) Q^*(1/z^*)$, 即$S(\omega) = |Q(e^{i\omega})|^2 \sigma^2$
其中$Q(z)$是最小相位系统。

任一平稳信号和白噪音信号可以通过一个可逆最小相位系统互相得到。



混合过程

Remark
平稳信号含有连续的功率谱和离散的功率谱(谐波信号),称为混合信号。

Theorem (Wold分解定理)
任一平稳过程$x(n)$可以写成$x(n) = x_r(n) + x_p(n)$,其中$x_r(n)$是正则过程，$x_p(n)$是离散谱的可预测过程。 且$E(x_r x_p) = 0$(正交).
$r_x(k) = r_r(k) + r_p(k)$,可以写成 $x_r(n) = \sum_k b_k v(n-k)$($MA(\infty)$)过程。
$x_p(n) = -\sum_k a(k) x(n-k)$.

Theorem (Kolmogorov定理)
任一$ARMA$过程可以用无穷阶AR过程表示。

ARMA模型

Definition (ARMA方程)

如果平稳随机信号满足 $x(n) = -\sum_k^p a(k)x(n-k) + \sum_k^q b(k)w(n-k)$, 其中$w(n)$是白噪音输入;称为*ARMA(p,q)*过程。自回归移动平均模型。

设系统函数 $H(z) = \frac{B(z)}{A(z)}$

例子:

- MA模型 即ARMA(0,q)模型，$x(n) = \sum b(k)v(n-k)$,全零点模型。
  正则的平稳随机过程：清辅音(不用声带振动的语音),连续的功率谱和
- AR模型: 即ARMA(p,0)模型，$x(n) = \sum a(k)x(n-k) + w(n)$,全极点模型。
  可预测过程，元音及含声带振动的辅音，拟周期性；离散的功率谱

经典谱估计概述

- 自相关函数的估计：
  $r(k) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x(n)x(n+k)$
  验证：估计是有偏，渐进无偏估计；**方差性能好（可以构造无偏估计但方差不好)
  说明：有限信号是矩形窗截断；等价与自相关加三角窗（延迟窗)。
- 直接法： $P(k) = \frac{1}{N}|X_n(k)|^2$
  改进(Bartlett，Welch等):分段求傅立叶系数，再平均。 减小误差，降低分辨率。
- 间接法：有限信号的自相关函数的有限傅立叶变换。
  $P(w) = \sum_{n=-M}^{M} r(n)e^{-iwn}, |M| \le N-1$
  分析：有数据窗的截断；延迟窗的泄漏，有限变换的分辨率降低等问题；

Yule-Walker方程

AR(p)模型 $\sum_{k=0}^{p} a_k x(n-k) = w(n)$
两边同时乘以$x^*(n-m)$,取期望
$E\left[\sum_k a_k^* x(n-k)x^*(n-m)\right] = E\left[w(n)x(n-m)\right]$
由因果性$w(n)$与$x(n-m)$不相关； $m=0; \sigma_w^2 = \sum a_k r(k)$
$m>0; \sum_{k=0}^{p} a_k^* r(m-k) = 0$
$R_x\vec{a} = -\vec{r}$为Yule-Walker方程。
$\vec{a} = (a_1,\ldots,a_p), \vec{r} = (r^*(1),\ldots,r^*(p))$

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(p-1) \\ r^*(1) & r(0) & \cdots & r(p-2) \\ \vdots & \vdots & \vdots & \vdots \\ r^*(p-1) & r^*(p-2) & \cdots & r(0) \end{bmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} r^*(1) \\ r^*(2) \\ \vdots \\ r^*(p) \end{pmatrix}$$

模型参数计算

增广方程：$R_{p+1}a_p = \sigma_w^2 u$

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(p) \\ r^*(1) & r(0) & \cdots & r(p-1) \\ \vdots & \vdots & \vdots & \vdots \\ r^*(p) & r^*(p-1) & \cdots & r(0) \end{bmatrix} \begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} \sigma_w^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

利用$r(0),\ldots,r(p)$可以求出a_p和σ_w^2.
实际中应用观测值估计R,求参数。
注意：ARMA,MA模型对应的求参数方程是非线性的!

Wiener-Hopf 方程和Yule-Walker 方程

- 前向预测误差滤波器$e(n) = x(n) - \sum h[k]x[n-k]$
  Wiener-Hopf方程：$Rh_0 = \vec{r}$
- AR(p)模型系数计算 $\sum_{k=0}^{p} a_k x(n-k) = w(n)$；
  $R_x\vec{a} = -\vec{r}$为Yule-Walker方程。
- Wiener-Hopf 方程和Yule-Walker 方程的一致性。
  增广方程：$R_{p+1}a_p = \sigma_w^2 u$
  $$\begin{bmatrix} r(0) & r(1) & \cdots & r(p) \\ r^*(1) & r(0) & \cdots & r(p-1) \\ \vdots & \vdots & \vdots & \vdots \\ r^*(p) & r^*(p-1) & \cdots & r(0) \end{bmatrix} \begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} \sigma_w^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$
- 滤波器算法的实现关键是求解以上方程!