

一般来说,确定性信号的滤波可以认为是对信号不同频谱分量的取舍,例如低通滤波器保留信号的低频分量,丢弃信号的高频分量。对于随机信号,滤波可以认为是对其功率谱不同分量的取舍,在时间域对应的操作就是对自相关函数的修改。随机信号滤波的本质是对其统计特征和概率特征的修改,而不是对波形的修改。因此,滤波器的设计准则也将基于相关统计特征,这和确定性信号滤波有显著的不同。本章在介绍滤波器基本概念的基础上,重点分析基于平方误差准则的滤波器及预测器设计,即 Wiener 滤波、卡尔曼滤波、最小二乘滤波和线性预测;然后介绍基于输出信噪比准则的滤波器(匹配滤波器)设计;最后介绍自适应滤波。本章只考虑数字随机信号的滤波。

5.1 数字滤波器的基本概念

一个典型的数字滤波器框图如图 5-1 所示。

设输入信号为 $x(n)$, 输出信号为 $y(n)$, 这里只考虑线性滤波, 滤波过程可用以下差分方程来表示:

$$y(n) = - \sum_{i=1}^{M-1} a_i y(n-i) + \sum_{i=0}^{N-1} b_i x(n-i) \quad (5-1)$$

由式(5-1)可知该数字滤波器的传递函数为

$$H(z) = \frac{\sum_{i=0}^{N-1} b_i z^{-i}}{1 + \sum_{i=1}^{M-1} a_i z^{-i}} \quad (5-2)$$

其单位冲激响应函数为

$$h(n) = z^{-1}(H(z)) \quad (5-3)$$

$$y(n) = h(n) \otimes x(n) = \sum_{i=0}^{\infty} h(i)x(n-i) \quad (5-4)$$

式中, $n < 0$ 时, 有 $h(n) = 0$, 这样的滤波器系统称为因果系统。

如果冲激响应函数是有限长的, 即

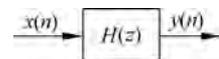


图 5-1 数字滤波器框图

$$h(n) = \begin{cases} h_n, & 0 \leq n \leq N \\ 0, & \text{其他} \end{cases} \quad (5-5)$$

则称此滤波器为有限冲激响应(Finite Impulse Response, FIR)滤波器, 否则, 称为无限冲激响应(Infinite Impulse Response, IIR)滤波器。

如果 $h(n)$ 满足如下条件:

$$\begin{cases} h(n) = 0, & n < 0 \\ \sum_{n=0}^{\infty} |h(n)| < C \end{cases} \quad (5-6)$$

则称此滤波器是因果的, 并且是稳定的。在这一章讨论稳定、因果的 FIR 或 IIR 滤波器。

随机信号滤波的一般问题模型如图 5-2 所示, 其中 $x(n)$ 是观测信号, $d(n)$ 是期望信号。目标是通过滤波器参数的合理设置, 使得滤波器输出 $y(n)$ 与期望信号尽可能接近, 即使得误差信号 $e(n) = d(n) - y(n)$ 尽量小。

图 5-3 给出了产生观测信号的两模型, 其中, 图 5-3(a) 表示观测数据是由目标信号 $s(n)$ 和噪声干扰 $w(n)$ 组成的, 图 5-3(b) 表示信号不仅受到加性噪声的干扰, 还存在未知系统引起的畸变。假设 $s(n)$ 与 $w(n)$ 是广义平稳的随机信号, 希望通过数据处理后, 使噪声受到抑制, 同时增强并恢复目标信号。根据期望信号的不同, 一般可分为如下三种情况:

- (1) $d(n) = s(n)$, 此时希望从 $x(n)$ 中估计出 $s(n)$, 这是一个滤波问题;
- (2) $d(n) = s(n+k)$, $k > 0$, 此时希望估计 $x(n)$ 将来的值, 这是一个预测问题;
- (3) $d(n) = s(n-k)$, $k > 0$, 此时希望估计 $x(n)$ 以前的值, 这是一个平滑问题。

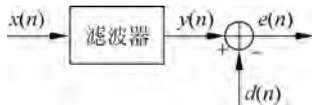


图 5-2 随机信号滤波的一般问题模型

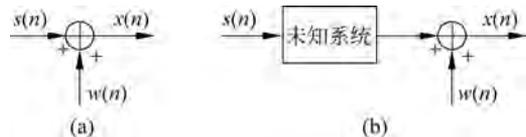


图 5-3 产生观测信号的模型

5.2 Wiener 滤波

Wiener(维纳)滤波器的设计准则是: 考虑线性时不变滤波器, 以使误差信号达到最小均方误差, 它是目前应用最广泛的滤波器。

5.2.1 最小均方误差准则与正交性原理

设线性滤波器的冲激响应函数为 $h(n)$, $n \geq 0$, 即

$$y(n) = \sum_{k=0}^{\infty} h(k)x(n-k) \quad (5-7)$$

误差信号为

$$e(n) = d(n) - y(n) = d(n) - \sum_k h(k)x(n-k) \quad (5-8)$$

最小均方误差(Minimum Mean-Squared Error, MMSE)准则是使均方误差 $E[|e(n)|^2]$ 最小,

以此来设计滤波器。

构造如下目标函数：

$$J = E[|e(n)|^2] \quad (5-9)$$

要使目标函数最小,应有 $\frac{\partial J}{\partial h(k)} = 0, k=0, 1, 2, \dots$ 。

由于

$$\begin{aligned} \frac{\partial J}{\partial h(k)} &= \frac{\partial E[|e(n)|^2]}{\partial h(k)} \\ &= E \left[e^*(n) \frac{\partial [e(n)]}{\partial h(k)} \right] \\ &= E \left[e^*(n) \frac{\partial \left(d(n) - \sum_{k=0}^{\infty} h(k)x(n-k) \right)}{\partial h(k)} \right] \\ &= -E[e^*(n)x(n-k)] \end{aligned} \quad (5-10)$$

$\frac{\partial J}{\partial h(k)} = 0$, 即

$$E[e^*(n)x(n-k)] = 0, \quad k=0, 1, 2, \dots \quad (5-11)$$

这说明在 MMSE 准则下,误差 $e(n)$ 与每个输入样本 $x(n-k)$ 都是正交的,这就是所谓的正交性原理。式(5-11)和第3章中提到的正交性原理是一致的。

当滤波器的选择满足式(5-11)的正交性准则时,滤波器最优输出为

$$y(n) = \sum_{k=0}^{\infty} h(k)x(n-k) \quad (5-12)$$

不难证明: $E[e^*(n)y(n)] = 0$, 即输出信号与误差信号是正交的。

满足正交性准则的 Wiener 滤波器的均方误差为

$$\begin{aligned} E_{\min} &= E(|e(n)|^2) \\ &= E \left[\left(d(n) - \sum_{k=0}^{+\infty} h(k)x(n-k) \right) e^*(n) \right] \\ &= E[d(n)e^*(n)] \\ &= E \left[|d(n)|^2 - \sum_{k=0}^{+\infty} h^*(k)x^*(n-k)d(n) \right] \\ &= \sigma_d^2 - \sum_{k=0}^{+\infty} h^*(k)R_{dx}(k) \end{aligned} \quad (5-13)$$

5.2.2 Wiener-Hopf 正则方程

由正交性原理 $E[e^*(n)x(n-m)] = 0$ 可知

$$E \left[\left(d(n) - \sum_{k=0}^{\infty} h(k)x(n-k) \right) x^*(n-m) \right] = 0 \quad (5-14)$$

经简单计算,可得

$$R_{dx}(m) - \sum_{k=0}^{\infty} h(k)R_x(m-k) = 0 \quad (5-15)$$

上式等价于

$$\sum_{k=0}^{\infty} h(k)R_x(m-k) = R_{dx}(m) \quad (5-16)$$

此方程称为 Wiener-Hopf 方程,又称为 Wiener 滤波器的正则方程。

对于 FIR Wiener 滤波器,有

$$\sum_{k=0}^N h(k)R_x(m-k) = R_{dx}(m) \quad (5-17)$$

取 $m=0,1,2,\dots,N$,可得如下方程组:

$$\begin{bmatrix} R_x(0) & R_x(-1) & \cdots & R_x(-N) \\ R_x(1) & R_x(0) & \cdots & R_x(-N+1) \\ \vdots & \vdots & \ddots & \vdots \\ R_x(N) & R_x(N-1) & \cdots & R_x(0) \end{bmatrix} \begin{bmatrix} h(0) \\ h(1) \\ \vdots \\ h(N) \end{bmatrix} = \begin{bmatrix} R_{dx}(0) \\ R_{dx}(1) \\ \vdots \\ R_{dx}(N) \end{bmatrix} \quad (5-18)$$

根据式(2-48)自相关矩阵定义,上式可简写为

$$\mathbf{R}_x^T \mathbf{h} = \mathbf{r}_{dx}$$

如果通过观测数据可得到自相关函数 $\hat{R}_x(m)$ 及互相关函数 $\hat{R}_{dx}(m)$ 的估值,则根据方程组(5-18)可求解出滤波器系数 $\mathbf{h} = \mathbf{R}_x^{-T} \mathbf{r}_{dx}$ 。

下面考虑一种特殊情况,假设信号传输过程受到噪声干扰,观察信号为 $x(n) = s(n) + \omega(n)$,滤波的目的是降低噪声的影响,考虑两种特殊情况:

(1) 若 $d(n) = s(n)$,且 $s(n)$ 与 $\omega(n)$ 互不相关,则

$$\begin{aligned} R_{dx}(m) &= E[s(n)x^*(n-m)] \\ &= E[s(n)(s^*(n-m) + \omega^*(n-m))] \\ &= R_s(m) \end{aligned} \quad (5-19)$$

同时有

$$R_x(m) = R_s(m) + R_u(m) = R_s(m) + \delta(m)\sigma^2 \quad (5-20)$$

此时 Wiener-Hopf 方程为

$$\sum_{k=0}^N h(k)R_x(m-k) = R_s(m), \quad m=0,1,2,\dots,N \quad (5-21)$$

(2) 若 $d(n) = s(n+\alpha)$, $\alpha > 0$,则 $R_{dx}(m) = R_s(m+\alpha)$ 。

此时 Wiener-Hopf 方程为

$$\sum_{k=0}^N h(k)R_x(m-k) = R_s(m+\alpha), \quad m=0,1,2,\dots,N \quad (5-22)$$

例 5-1 已知 $x(n) = s(n) + \omega(n)$,其中信号 $s(n)$ 是 AR(1)过程, $s(n) = 0.6s(n-1) + u(n)$, $u(n)$ 是 $N(0, 0.64)$ 的白噪声, $\omega(n)$ 是 $N(0, 1)$ 标准高斯过程。试设计一个长度为 2 的 Wiener 滤波器估计 $s(n)$ 。

解 $s(n)$ 的功率谱为

$$S(\omega) = \frac{0.64}{|1 - 0.6e^{-j\omega}|^2} = \frac{0.64}{1.36 - 1.2\cos\omega} \quad (5-23)$$

对功率谱做傅里叶反变换,得到 $s(n)$ 的相关函数 $R_s(m) = 0.6^{|m|}$ 。根据信号模型可得

$$R_x(m) = R_s(m) + 1 \cdot \delta(m) = 0.6^{|m|} + \delta(m) \quad (5-24)$$

由 Wiener-Hopf 方程可得

$$\begin{aligned} 2h(0) + 0.6h(1) &= 1 \\ 0.6h(0) + 2h(1) &= 0.6 \end{aligned} \quad (5-25)$$

解得滤波器的系数为

$$\begin{aligned} h(0) &= 0.451 \\ h(1) &= 0.165 \end{aligned} \quad (5-26)$$

图 5-4 给出了例 5-1 中各信号的功率谱,可以看到式(5-26)所给出的滤波器能显著降低噪声的影响。

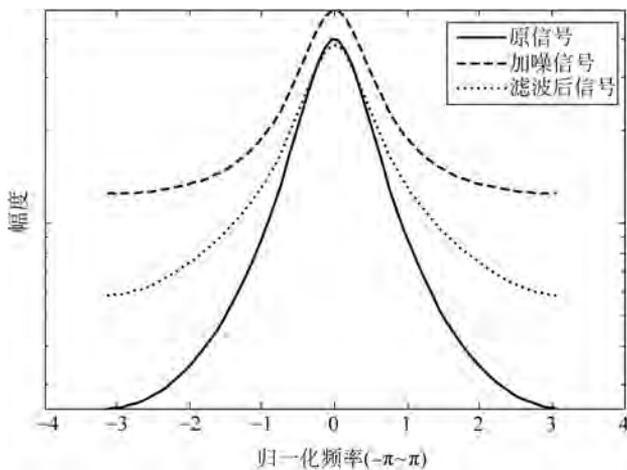


图 5-4 例 5-1 中滤波器的滤波效果

5.2.3 Wiener 滤波器的求解

求解 Wiener 滤波器的方法大致有如下 3 种:

- (1) 对于 FIR 滤波器,可直接用 Wiener-Hopf 方程组求解。
- (2) 对 IIR 滤波器,由 Wiener-Hopf 方程可知

$$\sum_{k=0}^{+\infty} h(k)R_x(m-k) = R_{dx}(m) \quad (5-27)$$

对式(5-27)进行 z 变换可得

$$H(z)P_x(z) = P_{dx}(z) \quad (5-28)$$

$$H(z) = \frac{P_{dx}(z)}{P_x(z)} \quad (5-29)$$

因为 $R_x(m), R_{dx}(m)$ 是双边序列,这里需采用双边 z 变换。对式(5-29)采用双边逆 z 变换可求出滤波器系数 $h(k)$,但由于采用双边变换, $h(k)$ 不是因果的,因此实用性较差,下面给出一种 IIR 滤波器的因果实现。

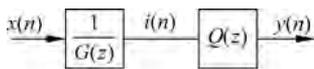


图 5-5 Wiener 滤波器

(3) 将 Wiener 滤波器看成两个滤波器的级联,其中第一个滤波器为因果 IIR 系统,第二个滤波器为因果 FIR 系统,如图 5-5 所示。

图 5-5 中,第一个系统称为白化滤波器,输出 $i(n)$ 是功率谱为 σ^2 的白噪声,传递函数为 $G^{-1}(z)$;第二个系统的传递函数是 $Q(z)$ 。

图 5-5 中所示的 Wiener 滤波器的传递函数为 $H(z) = \frac{Q(z)}{G(z)}$,所以只需求出 $G^{-1}(z)$ 和 $Q(z)$ 即可。

设 $x(n)$ 的功率谱密度函数为 $S_x(z)|_{z=e^{j\omega}}$, $G(z)$ 是可逆的最小相位系统,由 $X(z) = G(z)I(z)$,可知

$$S_x(z) = \sigma^2 G(z) \cdot G(z^{-1}) \quad (5-30)$$

假设 $S_x(z)$ 能分解为 $S_x(z) = S_x^+(z) \cdot S_x^-(z^{-1})$,其中一部分零极点在单位圆内,即 $|z| < 1$ [记为 $S_x^+(z)$],而另一部分在单位圆外,即 $|z| > 1$ [记为 $S_x^-(z)$],则

$$G(z) = S_x^+(z) \quad (5-31)$$

实际上由于 $S_x(z)$ 是有理谱密度,总可以分解为 $S_x(z) = S_x^+(z) \cdot S_x^-(z^{-1})$ 的形式,这就是所谓的谱因式分解定理。

考查图 5-5 中第二个滤波器,有

$$y(n) = \sum_{k=0}^{\infty} q(k) i(n-k) \quad (5-32)$$

式(5-32)中,利用正交性原理,第二个滤波器的 Wiener-Hopf 方程(参阅式(5-17))为

$$\sum_{k=0}^{\infty} q(k) R_i(m-k) = R_{di}(m), \quad m=0,1,2,\dots \quad (5-33)$$

由于 $i(n)$ 是白噪声,所以 $R_i(m-k) = \delta(m-k)\sigma^2$,代入式(5-32),可得

$$q(m)\sigma^2 = R_{di}(m) \quad (5-34)$$

$$q(m) = \frac{1}{\sigma^2} R_{di}(m), \quad m=0,1,2,\dots \quad (5-35)$$

设 $R_{di}(m)$ 的 z 变换为 $\Gamma(z) = \sum_{k=-\infty}^{\infty} R_{di}(k)z^{-k}$,由于 $Q(z) = \sum_{k=0}^{\infty} q(k)z^{-k}$,所以

$$Q(z) = \frac{1}{\sigma^2} [\Gamma(z)]_+ \quad (5-36)$$

式中, $[\Gamma(z)]_+ = \sum_{k=0}^{\infty} R_{di}(k)z^{-k}$ 是 $R_{di}(m)$ 的单边 z 变换。

求解 $R_{di}(m)$ 过程如下:

$$\begin{aligned} R_{di}(m) &= E [d(n) i^*(n-m)] \\ &= E \left[\sum_{k=0}^{\infty} v^*(k) x^*(n-m-k) \cdot d(n) \right] \\ &= \sum_{k=0}^{\infty} v^*(k) R_{dx}(m+k) \end{aligned} \quad (5-37)$$

其中 $v(k)$ 是第一个滤波器的冲激响应函数, 即

$$\frac{1}{G(z)} = V(z) = \sum_{k=0}^{\infty} v(k) z^{-k} \quad (5-38)$$

对式(5-37)进行 z 变换

$$\begin{aligned} \Gamma_{di}(z) &= \sum_{m=-\infty}^{\infty} \left(\sum_{k=0}^{\infty} v^*(k) R_{dx}(m+k) \right) \cdot z^{-m} \\ &= \sum_{k=0}^{\infty} v^*(k) \sum_{m=-\infty}^{\infty} R_{dx}(m+k) z^{-(m+k)} z^k \\ &= \sum_{k=0}^{\infty} v^*(k) z^k \sum_{m=-\infty}^{\infty} R_{dx}(m+k) z^{-(m+k)} \\ &= \sum_{k=0}^{\infty} v^*(k) z^k \cdot \Gamma_{dx}(z) \\ &= v^* ((z^*)^{-1}) \Gamma_{dx}(z) \\ &= \frac{1}{G^* ((z^*)^{-1})} \Gamma_{dx}(z) \end{aligned} \quad (5-39)$$

所以

$$Q(z) = \frac{[\Gamma_{di}(z)]_+}{\sigma^2} = \frac{1}{\sigma^2} \left[\frac{\Gamma_{dx}(z)}{G^* ((z^*)^{-1})} \right]_+ \quad (5-40)$$

所以最终的 IIR Wiener 滤波器的传递函数为

$$H(z) = \frac{Q(z)}{G(z)} = \frac{1}{\sigma^2 G(z)} \left[\frac{\Gamma_{dx}(z)}{G^* ((z^*)^{-1})} \right]_+ \quad (5-41)$$

下面举例说明该方法的求解过程。

例 5-2 已知 $x(n) = s(n) + w(n)$, 其中信号 $s(n)$ 是 AR(1) 过程, $s(n) = 0.8s(n-1) + u(n)$, $u(n)$ 是 $N(0, 0.36)$ 的白噪声, $w(n)$ 是 $N(0, 1)$ 标准高斯过程。试设计一个因果 IIR Wiener 滤波器估计 $s(n)$ 。

解 $s(n)$ 是 AR(1) 过程, 其自相关函数对应的 z 变换为

$$S_s(z) = \frac{0.36}{(1 - 0.8z^{-1})(1 - 0.8z)} \quad (5-42)$$

由题意可知

$$S_x(z) = S_s(z) + 1 = \frac{1.6(1 - 0.5z^{-1})(1 - 0.5z)}{(1 - 0.8z^{-1})(1 - 0.8z)} \quad (5-43)$$

可知 $\sigma^2 = 1.6$, 且

$$G(z) = \frac{1 - 0.5z^{-1}}{1 - 0.8z^{-1}} \quad (5-44)$$

互相关函数 $R_{dx}(m) = R_{ds}(m) = R_{ss}(m)$, 所以, $\Gamma_{dx}(z) = S_s(z)$ 。

$$\left[\frac{\Gamma_{dx}(z)}{G(z^{-1})} \right]_+ = \left[\frac{\frac{0.36}{(1 - 0.8z^{-1})(1 - 0.8z)}}{\frac{1 - 0.5z}{1 - 0.8z}} \right]_+$$

$$\begin{aligned}
 &= \left[\frac{0.36}{(1 - 0.8z^{-1})(1 - 0.5z)} \right]_+ \\
 &= \frac{0.6}{1 - 0.8z^{-1}}
 \end{aligned} \tag{5-45}$$

所以得到 Wiener 滤波器的传递函数为

$$\begin{aligned}
 H(z) &= \frac{1}{1.6} \frac{1}{G(z)} \left[\frac{\Gamma_{xd}(z)}{G(z^{-1})} \right]_+ \\
 &= \frac{1}{1.6} \frac{1 - 0.8z^{-1}}{1 - 0.5z^{-1}} \frac{0.6}{1 - 0.8z^{-1}} \\
 &= \frac{3}{8} \frac{1}{1 - 0.5z^{-1}}
 \end{aligned} \tag{5-46}$$

求 z 逆变换, 得到滤波器的系数(单位冲激响应函数)为

$$h(n) = \frac{3}{8} \cdot \left(\frac{1}{2}\right)^n, \quad n \geq 0 \tag{5-47}$$

图 5-6 给出了例 5-2 中滤波器的降噪滤波效果, 可以看到式(5-47)所给出的一阶 IIR 滤波器性能优于图 5-4 中的 FIR 滤波器。

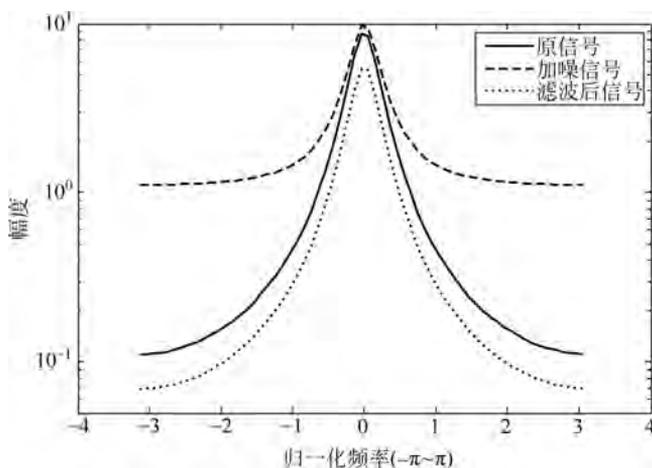


图 5-6 例 5-2 中滤波器的滤波效果

5.3 线性预测

Wiener 滤波是一个有广泛通用性的滤波理论。本节将用 Wiener 滤波理论来分析随机信号的可预测性问题, 这里只考虑线性预测。线性预测的基本含义是用随机信号过去时刻的线性组合预测当前时刻的取值, 线性预测是考察随机信号时间相关特性的重要手段, 广泛应用于多个领域。

设 $x(n)$ 在 n 时刻之前 p 个数据为 $x(n-p), x(n-p+1), \dots, x(n-1)$, 利用这 p 个数据预测 n 时刻的值 $x(n)$ (如图 5-7(a) 所示), 称为“前向预测”, 线性预测表示为

$$\hat{x}(n) = - \sum_{k=1}^p a_k x(n-k) \quad (5-48)$$

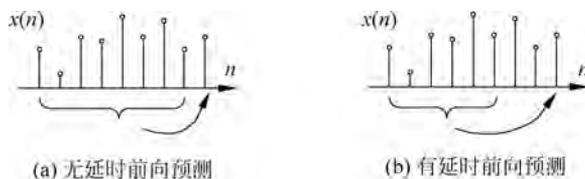


图 5-7 前向预测示意图

式(5-48)中, $\hat{x}(n)$ 为预测值, a_k 为预测系数。与 5.1 节所提到的预测问题相比, 上述预测问题忽略了观测噪声的影响, 可更集中考虑信号本身的特性。后续的推导也可拓展到噪声下的预测问题。式(5-48)称为无延时预测, 当存在延时 D 时, 线性预测表示为(如图 5-7(b) 所示)

$$\hat{x}(n) = - \sum_{k=1}^p a_k x(n-D-k) \quad (5-49)$$

有延时预测与无延时预测的预测器求解采用相同的方法, 这里只讨论无延时预测。定义预测误差为

$$e(n) = x(n) - \hat{x}(n) \quad (5-50)$$

下面基于最小均方误差准则推导预测器系数求解过程, 预测均方误差为

$$\rho = E(|e(n)|^2) = E \left[\left| x(n) + \sum_{k=1}^p a_k x(n-k) \right|^2 \right] \quad (5-51)$$

欲使 ρ 最小, 令 $\frac{\partial \rho}{\partial a_k} = 0$, 可得

$$E[x(n-k)e^*(n)] = 0, \quad k = 1, 2, \dots, p \quad (5-52)$$

式(5-52)和式(5-11)一致, 是正交性原理在线性预测中的体现。要得到最佳预测器的求解, 先对式(5-52)求共轭, 并用 m 代替 k , 可得

$$E \left[x^*(n-m) \left(x(n) + \sum_{k=1}^p a_k x(n-k) \right) \right] = 0, \quad m = 1, 2, \dots, p \quad (5-53)$$

可以得到如下包含预测器系数的方程:

$$R_x(m) = - \sum_{k=1}^p a_k R_x(m-k), \quad m = 1, 2, \dots, p \quad (5-54)$$

或以下矩阵形式:

$$\begin{bmatrix} R_x(0) & R_x(-1) & \cdots & R_x(-p+1) \\ R_x(1) & R_x(0) & \cdots & R_x(-p+2) \\ \vdots & \vdots & \ddots & \vdots \\ R_x(p-1) & R_x(p-2) & \cdots & R_x(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} R_x(1) \\ R_x(2) \\ \vdots \\ R_x(p) \end{bmatrix} \quad (5-55)$$

此时最小的 ρ 值(最小预测均方误差)为

$$\rho_{\min} = E[x(n)(x(n) - \hat{x}(n))^*] = R_x(0) + \sum_{k=1}^p a_k^* R_x(k) \quad (5-56)$$

方程(5-54)与方程(5-55)称为线性预测 Wiener-Hopf 方程,解此方程可以求解出线性预测模型的预测系数 a_k 及最小预测均方误差 ρ 。

前面叙述的模型使用了 p 个 n 时刻之前的数据预测 $x(n)$ 为“前向预测”。对平稳信号,时间相关性与时间起点无关。因此,亦可用 $n, n-1, \dots, n-p+1$ 的数据向后预测 $x(n-p)$,这称为“后向预测”。

为了表示清楚,把前向预测(forward prediction)模型记为

$$\begin{cases} \hat{x}^f(n) = -\sum_{k=1}^p a_k^f x(n-k) \\ e^f(n) = x(n) - \hat{x}^f(n) \\ \rho^f = E[|e^f(n)|^2] \end{cases} \quad (5-57)$$

利用 $x(n), x(n-1), \dots, x(n-p+1)$ 预测 $x(n-p)$,称为后向预测(backward prediction),如图 5-8 所示,其数学模型为

$$\hat{x}^b(n-p) = -\sum_{k=1}^p a_k^b x(n-p+k) \quad (5-58)$$

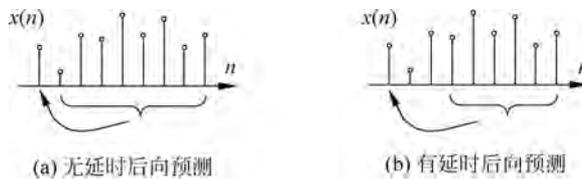


图 5-8 后向预测示意图

误差记为: $e^b(n) = x(n-p) - \hat{x}^b(n-p)$ 。

均方误差记为: $\rho^b = E[|e^b(n)|^2]$ 。

利用正交性原理,令 ρ^b 最小,同样可得后向预测的 Wiener-Hopf 方程

$$E[x(n-p+m)e^*(n)] = 0, \quad m=1,2,\dots,p \quad (5-59)$$

$$E\left[x^*(n-p+m)\left(x(n-p) + \sum_{k=1}^p a_k^b x(n-p+k)\right)\right] = 0, \quad m=1,2,\dots,p \quad (5-60)$$

$$R_x(-m) = -\sum_{k=1}^p a_k^b R_x(k-m), \quad m=1,2,\dots,p \quad (5-61)$$

式(5-61)的矩阵表达式为

$$\begin{bmatrix} R_x(0) & R_x(1) & \cdots & R_x(p-1) \\ R_x(-1) & R_x(0) & \cdots & R_x(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_x(-p+1) & R_x(-p+2) & \cdots & R_x(0) \end{bmatrix} \begin{bmatrix} a_1^b \\ a_2^b \\ \vdots \\ a_p^b \end{bmatrix} = - \begin{bmatrix} R_x(-1) \\ R_x(-2) \\ \vdots \\ R_x(-p) \end{bmatrix} \quad (5-62)$$

式(5-62)和式(5-55)形式类似,利用自相关函数的共轭对称特性,对式(5-62)两边取共轭,对照式(5-55),得到 $a_k^f = [a_k^b]^*$,即前向预测系数和后向预测系数互为共轭。利用和前向预测类似的方法,可求得预测误差为

$$\rho_{\min}^b = R_x(0) + \sum_{k=1}^p [a^b(k)]^* R_x(-k) \quad (5-63)$$

注意到预测误差肯定是实数,对式(5-63)取共轭并利用预测系数关系和自相关函数特性,不难得出式(5-63)和式(5-56)是等价的,即当预测器长度相同时,前向预测和后向预测的误差相等。前向预测和后向预测的关系总结如下:

- (1) $\rho_{\min}^b = \rho_{\min}^f$;
- (2) 若 a_k^f, a_k^b 是实数,则 $a_k^f = a_k^b$ 。
- 若 a_k^f, a_k^b 是复数,则 $a_k^f = [a_k^b]^*$ (共轭关系)。

前向预测、后向预测在变长度预测器的迭代求解中有重要应用,将在后面结合参数化功率谱估计深入分析。接下来考虑当随机信号可以用第4章线性模型描述时,其最佳线性预测的结果。首先考虑 AR 模型,一个 p 阶 AR 模型如下:

$$x(n) = - \sum_{k=1}^p \varphi_k x(n-k) + u(n) \quad (5-64)$$

其中, $u(n)$ 是零均值白噪声, φ_k 是模型系数。利用模型参数构造如下预测模型:

$$\hat{x}(n) = - \sum_{k=1}^p \varphi_k x(n-k) \quad (5-65)$$

这时预测误差刚好为 $u(n)$, 因为 $u(n)$ 和过去时刻的模型输出不相关, 不难验证式(5-65)满足正交性原理, 因此是最佳线性预测。

对比式(5-64)和式(5-65), 可看到 p 阶 AR 模型和 p 阶线性预测器是一致的, 定义

$$A(z) = 1 + \sum_{k=1}^p \varphi_k z^{-k} \quad (5-66)$$

AR 模型和线性预测的关系如图 5-9 所示, 由于滤波器 $A(z)$ 能将 $x(n)$ 变成一个白噪声 $u(n)$ [预测误差信号 $e(n)$], 所以称 $A(z)$ 为白化滤波器, 或反滤波器。AR 模型、白化滤波器及线性预测器分别示于图 5-9(a)、图 5-9(b)和图 5-9(c)。

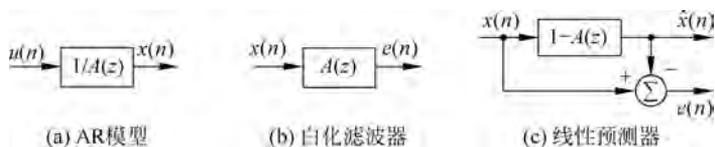


图 5-9 AR 模型、白化滤波器和线性预测器

5.4 卡尔曼滤波

在应用中, Wiener 滤波理论必须把用到的全部数据存储起来, 估计接收信号的自相关矩阵, 然后再进行计算。按照这种滤波方法设置的专用计算机的存储量与计算量比较大, 并且很难进行实时处理。在解决非平稳过程的滤波问题时, 难以得到高效方法。到 20 世纪 50 年代中期, 随着空间技术的发展, 这种方法越来越不能满足实际应用的需要, 面临着新的挑战。1960 年和 1961 年, 卡尔曼(R. E. Kalman)和布西(R. S. Bucy)提出了递推滤波算法, 成功地将状态变量法引入滤波理论中, 用消息与干扰的状态空间模型代替了通常用来描述

它们的协方差函数,将状态空间描述与离散时间更新联系起来,适于计算机直接进行运算,而不用去寻求滤波器冲激响应的明确公式。这种方法得出的是表征状态估计值及其均方误差的微分方程,给出的是递推算法。这就是著名的卡尔曼理论,或称卡尔曼-布西滤波。

卡尔曼滤波不要求保存过去的测量数据,当新的数据到来时,根据新的数据和前一时刻诸量的估值,借助于系统本身的状态转移方程,按照一套递推公式,即可算出新诸量的估值。与 Wiener 滤波器不同,卡尔曼滤波器能够利用先前的运算结果,再从当前数据提供的最新信息,即可得到当前的估值。卡尔曼递推算法大大减少了滤波装置的存储量和计算量,并且突破了平稳随机过程的限制,使卡尔曼滤波器适用于对时变信号的实时处理。

卡尔曼滤波基于信号的状态空间模型,如下式所示:

$$\mathbf{x}_{n+1} = \mathbf{F}\mathbf{x}_n + \mathbf{G}\mathbf{u}_n + \mathbf{w}_n \quad (5-67)$$

$$\mathbf{y}_{n+1} = \mathbf{H}\mathbf{x}_{n+1} + \mathbf{v}_{n+1} \quad (5-68)$$

其中, \mathbf{x}_n 为 $p \times 1$ 维系统状态向量, \mathbf{F} 为 $p \times p$ 维系统矩阵, \mathbf{u}_n 为 $k \times 1$ 维输入控制向量, \mathbf{w}_n 为系统动态噪声向量, \mathbf{G} 为系统控制矩阵,大小为 $p \times k$ 。式(5-67)表示系统特性可由若干参数形成的状态向量描述。当前的状态与前一个时刻的状态、当前输入和系统噪声有关,系统可能存在多路输入。

\mathbf{y}_{n+1} 为观测向量,大小为 $r \times 1$ 。 \mathbf{H} 为观测矩阵,大小为 $r \times p$,描述系统状态对输出的影响。 \mathbf{v}_{n+1} 为观测噪声向量,大小为 $r \times 1$ 。

式(5-67)中系统状态随时间变化,因此式(5-67)称为状态方程,式(5-68)称为观测方程。

对于给定的系统,状态空间模型为已知,状态空间模型具有良好的一般性,对许多实际问题可用状态空间建模,实际上第4章的线性模型可以认为是状态空间模型的特例,例如例5-2中所描述的噪声中的信号 AR(p)模型: $y(n) = x(n) + v(n)$ 。若定义系统的状态为 $\mathbf{x}_n = [x_n, x_{n-1}, \dots, x_{n-p+1}]^T$ (为描述方便,这里将 $x(n)$ 简记为 x_n),可用状态空间模型表示为

$$\begin{bmatrix} x_{n+1} \\ x_n \\ \vdots \\ x_{n-p+2} \end{bmatrix} = \begin{bmatrix} -\varphi_1 & -\varphi_2 & \cdots & -\varphi_p \\ 1 & & & 0 \\ & \ddots & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} x_n \\ x_{n-1} \\ \vdots \\ x_{n-p+1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \mathbf{u}_{n+1} \quad (5-69)$$

$$\begin{bmatrix} y_{n+1} \\ y_n \\ \vdots \\ y_{n-p+2} \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \begin{bmatrix} x_{n+1} \\ x_n \\ \vdots \\ x_{n-p+2} \end{bmatrix} + \begin{bmatrix} v_{n+1} \\ v_n \\ \vdots \\ v_{n-p+2} \end{bmatrix} \quad (5-70)$$

下面讨论卡尔曼滤波的具体流程。卡尔曼滤波关注在系统矩阵、控制矩阵、观测矩阵已知的情况下根据观测数据对状态空间的估计。假设在时刻 n ,基于 n 时刻以前所获得的全部知识,对状态变量 \mathbf{x}_n 做出一个预测估计,记为 $\hat{\mathbf{x}}_{n|n-1}$,则预测估计的误差为

$$\mathbf{e}_{n|n-1} = \mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1} \quad (5-71)$$

称为预测误差。预测误差是零均值的,其协方差矩阵为

$$\mathbf{C}_{n|n-1} = E[\mathbf{e}_{n|n-1} \mathbf{e}_{n|n-1}^H] = E[(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1})(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1})^H] \quad (5-72)$$

在预测估计 $\hat{\mathbf{x}}_{n|n-1}$ 的基础上,利用 n 时刻所获取的新观测数据 \mathbf{y}_n 进一步改善对 \mathbf{x}_n 的估

计,记为 $\hat{\mathbf{x}}_n$,称为更新估计,更新估计通过下式完成

$$\hat{\mathbf{x}}_n = \hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1}) \quad (5-73)$$

其中, \mathbf{K}_n 为待定的增益矩阵,称为卡尔曼增益, $(\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1})$ 称为新息,代表由新观测数据 \mathbf{y}_n 获得的新信息。

更新估计的误差记作 \mathbf{e}_n ,则有

$$\mathbf{e}_n = \mathbf{x}_n - \hat{\mathbf{x}}_n = \mathbf{x}_n - [\hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1})] \quad (5-74)$$

卡尔曼滤波的实质是寻找适当的增益矩阵 \mathbf{K}_n ,使更新估计的均方误差达到最小。更新估计的均方误差为

$$E[|\mathbf{e}_n|^2] = E[(\mathbf{x}_n - \hat{\mathbf{x}}_n)^H (\mathbf{x}_n - \hat{\mathbf{x}}_n)] \quad (5-75)$$

而更新估计的协方差矩阵为

$$\mathbf{C}_n = E[\mathbf{e}_n \mathbf{e}_n^H] = E[(\mathbf{x}_n - \hat{\mathbf{x}}_n)(\mathbf{x}_n - \hat{\mathbf{x}}_n)^H] \quad (5-76)$$

比较以上两式可知,更新估计的均方误差 $E[|\mathbf{e}_n|^2]$ 为协方差矩阵 \mathbf{C}_n 的对角元素之和,因此,要使均方误差最小,就等效于使协方差矩阵对角元素之和最小。

将式(5-73)代入式(5-74),可得

$$\begin{aligned} \mathbf{e}_n &= \mathbf{x}_n - \hat{\mathbf{x}}_n \\ &= \mathbf{x}_n - [\hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n (\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1})] \\ &= \mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1} - \mathbf{K}_n (\mathbf{H}\mathbf{x}_n + \mathbf{v}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1}) \\ &= \mathbf{e}_{n|n-1} - \mathbf{K}_n \mathbf{H} \mathbf{e}_{n|n-1} - \mathbf{K}_n \mathbf{v}_n \\ &= (\mathbf{I} - \mathbf{K}_n \mathbf{H}) \mathbf{e}_{n|n-1} - \mathbf{K}_n \mathbf{v}_n \end{aligned} \quad (5-77)$$

由式(5-76)得到

$$\mathbf{C}_n = (\mathbf{I} - \mathbf{K}_n \mathbf{H}) \mathbf{C}_{n|n-1} (\mathbf{I} - \mathbf{K}_n \mathbf{H})^H - \mathbf{K}_n \mathbf{R}_v \mathbf{K}_n^H \quad (5-78)$$

其中, \mathbf{R}_v 是观测噪声的自相关矩阵。注意式(5-78)用到了信号与噪声的不相关特性。令

$$\mathbf{D}_n \mathbf{D}_n^H = \mathbf{H} \mathbf{C}_{n|n-1} \mathbf{H}^H - \mathbf{R}_v \quad (5-79)$$

将式(5-78)展开,并将 $\mathbf{D}_n \mathbf{D}_n^H$ 代入,得

$$\mathbf{C}_n = \mathbf{C}_{n|n-1} - \mathbf{K}_n \mathbf{H} \mathbf{C}_{n|n-1} - \mathbf{C}_{n|n-1} \mathbf{H}^H \mathbf{K}_n^H + \mathbf{K}_n \mathbf{D}_n \mathbf{D}_n^H \mathbf{K}_n^H \quad (5-80)$$

可见, \mathbf{C}_n 中含有 \mathbf{K}_n 的二次项和一次项,对其进行配方,得到

$$\mathbf{C}_n = \mathbf{C}_{n|n-1} + (\mathbf{K}_n \mathbf{D}_n - \mathbf{A})(\mathbf{K}_n \mathbf{D}_n - \mathbf{A})^H - \mathbf{A} \mathbf{A}^H \quad (5-81)$$

其中, \mathbf{A} 为未知矩阵,式(5-81)展开为

$$\begin{aligned} \mathbf{C}_n &= \mathbf{C}_{n|n-1} + \mathbf{K}_n \mathbf{D}_n \mathbf{D}_n^H \mathbf{K}_n^H - \mathbf{K}_n \mathbf{D}_n \mathbf{A}^H - \mathbf{A} \mathbf{D}_n^H \mathbf{K}_n^H + \mathbf{A} \mathbf{A}^H - \mathbf{A} \mathbf{A}^H \\ &= \mathbf{C}_{n|n-1} + \mathbf{K}_n \mathbf{D}_n \mathbf{D}_n^H \mathbf{K}_n^H - \mathbf{K}_n \mathbf{D}_n \mathbf{A}^H - \mathbf{A} \mathbf{D}_n^H \mathbf{K}_n^H \end{aligned} \quad (5-82)$$

比较式(5-80)和式(5-82),可得

$$\mathbf{A} = \mathbf{C}_{n|n-1} \mathbf{H}^H (\mathbf{D}_n^H)^{-1} \quad (5-83)$$

观察式(5-81),其第一项和第三项均与增益矩阵 \mathbf{K}_n 无关,只有第二项包含 \mathbf{K}_n 。考虑到 \mathbf{C}_n 及第二项为对称矩阵,故其主对角线元素是非负的,若要使 \mathbf{C}_n 的主对角线元素之和达到最小,须要求第二项的主对角线元素之和达到最小,即为零。因此,有

$$\mathbf{K}_n \mathbf{D}_n - \mathbf{A} = 0 \quad (5-84)$$

则

$$\begin{aligned} \mathbf{K}_n &= \mathbf{A}\mathbf{D}_n^{-1} = \mathbf{C}_{n|n-1}\mathbf{H}^{\text{H}}(\mathbf{D}_n^{\text{H}})^{-1}\mathbf{D}_n^{-1} \\ &= \mathbf{C}_{n|n-1}\mathbf{H}^{\text{H}}(\mathbf{H}\mathbf{C}_{n|n-1}\mathbf{H}^{\text{H}} - \mathbf{R}_v)^{-1} \end{aligned} \quad (5-85)$$

\mathbf{K}_n 即为使更新估计均方误差达到最小的卡尔曼增益矩阵。

将式(5-85)代入式(5-78),可得更新估计的协方差矩阵

$$\mathbf{C}_n = (\mathbf{I} - \mathbf{K}_n\mathbf{H})\mathbf{C}_{n|n-1} \quad (5-86)$$

由上述分析可知,从预测估计 $\hat{\mathbf{x}}_{n|n-1}$ 和预测误差协方差矩阵 $\mathbf{C}_{n|n-1}$ 出发,即可求得卡尔曼增益矩阵 \mathbf{K}_n ,然后由新的观测向量 \mathbf{y}_n 计算出新的信息 $(\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1})$,再加上 $\hat{\mathbf{x}}_{n|n-1}$,就完成了对状态变量 \mathbf{x}_n 的更新估计 $\hat{\mathbf{x}}_n$ 。

为了使上述估计过程能递推地进行,还需要由 n 时刻的估计更新 $\hat{\mathbf{x}}_n$ 及其协方差矩阵 \mathbf{C}_n ,来计算 $n+1$ 时刻的预测估计 $\hat{\mathbf{x}}_{n+1|n}$ 及其协方差矩阵 $\mathbf{C}_{n+1|n}$ 。有了 $\hat{\mathbf{x}}_{n+1|n}$ 和 $\mathbf{C}_{n+1|n}$,就可以重复上述估计过程,进而得到 $\hat{\mathbf{x}}_{n+1}$ 和 \mathbf{C}_{n+1} ,完成 $n+1$ 时刻的更新估计。

由状态方程

$$\mathbf{x}_{n+1} = \mathbf{F}\mathbf{x}_n + \mathbf{G}\mathbf{u}_n + \mathbf{w}_n \quad (5-87)$$

若动态噪声 \mathbf{w}_n 是零均值的,并且不同 n 时刻的 \mathbf{w}_n 是不相关的,作为对 \mathbf{x}_{n+1} 的预测估计 $\hat{\mathbf{x}}_{n+1|n}$,暂时忽略 \mathbf{w}_n 的作用,于是取

$$\mathbf{x}_{n+1|n} = \mathbf{F}\hat{\mathbf{x}}_n + \mathbf{G}\mathbf{u}_n \quad (5-88)$$

其中,输入控制向量 \mathbf{u}_n 为已知,那么 $n+1$ 时刻的预测误差为

$$\mathbf{e}_{n+1|n} = \mathbf{x}_{n+1} - \hat{\mathbf{x}}_{n+1|n} = \mathbf{F}\mathbf{x}_n + \mathbf{G}\mathbf{u}_n + \mathbf{w}_n - (\mathbf{F}\hat{\mathbf{x}}_n + \mathbf{G}\mathbf{u}_n) = \mathbf{F}\mathbf{e}_n + \mathbf{w}_n \quad (5-89)$$

注意到 \mathbf{w}_n 与 \mathbf{e}_n 是互不相关的,那么 $n+1$ 时刻预测误差协方差矩阵为

$$\begin{aligned} \mathbf{C}_{n+1|n} &= E[\mathbf{e}_{n+1|n}\mathbf{e}_{n+1|n}^{\text{H}}] \\ &= E[(\mathbf{F}\mathbf{e}_n + \mathbf{w}_n)(\mathbf{F}\mathbf{e}_n + \mathbf{w}_n)^{\text{H}}] \\ &= \mathbf{F}\mathbf{C}_n\mathbf{F}^{\text{H}} + \mathbf{R}_w \end{aligned} \quad (5-90)$$

其中, \mathbf{R}_w 是 \mathbf{w}_n 的自相关矩阵。式(5-90)可进一步用于 $n+1$ 时刻卡尔曼增益矩阵的计算,从而形成迭代计算。根据最小均方误差准则,可以确定卡尔曼滤波器的初始值

$$\begin{aligned} \hat{\mathbf{x}}_{-1} &= E[\mathbf{x}_{-1}] \\ \mathbf{C}_{-1} &= E[(\mathbf{x}_{-1} - \hat{\mathbf{x}}_{-1})(\mathbf{x}_{-1} - \hat{\mathbf{x}}_{-1})^{\text{H}}] \end{aligned} \quad (5-91)$$

式(5-91)需要已知模型状态的统计量信息。实际应用中当无法得到这些先验信息时,卡尔曼滤波可以采用更灵活的初始化手段,后续的迭代更新会优化估计结果。例如 $\hat{\mathbf{x}}_{-1}$ 可以用全零或全 1 向量初始化, \mathbf{C}_{-1} 可以用单位矩阵初始化。

至此,完成了向量卡尔曼滤波算法的推导,总结其递推过程如下:

步骤 1: 建立状态空间模型,如式(5-67)和式(5-68)所示。

步骤 2: 设置初始化条件,即 $n=-1$ 时,采用如式(5-91)所示的初始化,即

$$\begin{cases} \hat{\mathbf{x}}_n = E[\mathbf{x}_n] \\ \mathbf{C}_n = E[(\mathbf{x}_n - \hat{\mathbf{x}}_n)(\mathbf{x}_n - \hat{\mathbf{x}}_n)^{\text{H}}] \end{cases}$$

步骤 3: 预测,如式(5-88)所示

$$\mathbf{x}_{n+1|n} = \mathbf{F}\hat{\mathbf{x}}_n + \mathbf{G}\mathbf{u}_n$$

步骤 4: 计算预测误差的协方差矩阵,如式(5-90)所示

观测方程:

$$\begin{bmatrix} y_{n+1} \\ y_n \\ \vdots \\ y_{n-\rho+2} \end{bmatrix} = \begin{bmatrix} x_{n+1} & x_n & \cdots & x_{n-L+2} \\ x_n & x_{n-1} & \cdots & x_{n-L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-\rho+2} & x_{n-\rho+1} & \cdots & x_{n-\rho-L+3} \end{bmatrix} \begin{bmatrix} h_{n+1}(0) \\ h_{n+1}(1) \\ \vdots \\ h_{n+1}(L-1) \end{bmatrix} + \begin{bmatrix} \omega_{n+1} \\ \omega_n \\ \vdots \\ \omega_{n-\rho+2} \end{bmatrix} \quad (5-97)$$

与前述状态空间模型的最大不同是式(5-68)中的观测矩阵是时变的,但在所考虑的信道估计问题中,训练序列,即观测矩阵是已知的,其时变性不会影响卡尔曼滤波的处理流程。

卡尔曼滤波和 Wiener 滤波都基于最小均方误差准则,但两者有很大不同,主要体现在如下。

(1) 卡尔曼滤波和 Wiener 滤波需要不同的统计量信息。由式(5-18)、式(5-29)所确定的 Wiener 滤波器需要计算目标信号和观测信号的互相关函数及观测信号的自相关函数,卡尔曼滤波只需系统噪声和观测噪声的统计量信息。卡尔曼滤波还需要初始化状态空间模型的系统状态,但如前面所提到,系统状态初始化可采用灵活的初始化策略。

(2) 卡尔曼滤波更适用于时变系统,这主要是因为 Wiener 滤波求解一个固定的滤波器并应用于所有情况,而卡尔曼滤波则采用每个符号更新的迭代计算。例如由式(5-95)、式(5-97)所描述的时变信道参数估计问题,难以用 Wiener 滤波直接求解。

5.5 最小二乘滤波

前述 Wiener 滤波器需要接收信号相关二阶矩信息,卡尔曼滤波需预先知道输入控制信号和噪声的二阶矩。实际应用中这些统计信息往往无法预先得到,仅有部分观测信号及其对应的期望信号。一种解决方案是从可用的数据估计出需要的二阶矩,再采用前述的 MMSE 滤波器;另一种是采用新的最小化性能标准,直接基于可用数据设计出最佳滤波器。本节将介绍第二种方法。

考虑式(5-7)所示的滤波问题,这里只考虑 FIR 滤波器,设线性滤波器的冲激响应函数为 $h(k)$, $0 \leq k \leq K-1$,则滤波器输出信号可表示为

$$y(n) = \sum_{k=0}^{K-1} h(k)x(n-k) \quad (5-98)$$

令 $d(n)$ 表示期望信号,误差信号为

$$e(n) = d(n) - y(n) = d(n) - \sum_{k=0}^{K-1} h(k)x(n-k) \quad (5-99)$$

采用基于 3.7 节的最小二乘准则求解滤波器系数,其目标函数是最小化每个时隙点误差模的平方和,如下所示:

$$\min_{h(k)} \sum_{n=0}^{N-1} |e(n)|^2 \quad (5-100)$$

其中, N 代表观测信号及期望信号的长度。下面推导滤波器参数的求解算法。定义观测信号向量为

$$\mathbf{y} = [y(N-1), \dots, y(1), y(0)]^T \quad (5-101)$$

观测信号整体可用以下矩阵相乘的方式描述

$$\mathbf{y} = \begin{bmatrix} y(N-1) \\ y(N-2) \\ \vdots \\ y(1) \\ y(0) \end{bmatrix} = \begin{bmatrix} x(N-1) & x(N-2) & \cdots & x(N-K) \\ x(N-2) & x(N-3) & \cdots & x(N-K-1) \\ \vdots & \vdots & \ddots & \vdots \\ x(1) & x(0) & \cdots & x(-K+2) \\ x(0) & x(-1) & \cdots & x(-K+1) \end{bmatrix} \begin{bmatrix} h(0) \\ h(1) \\ \vdots \\ h(K-2) \\ h(K-1) \end{bmatrix} \quad (5-102)$$

上式简记为

$$\mathbf{y} = \mathbf{X}\mathbf{h} \quad (5-103)$$

定义如下误差信号向量

$$\mathbf{e} = [e(N-1), \dots, e(1), e(0)]^T \quad (5-104)$$

误差信号能量可写为

$$\begin{aligned} E &= \mathbf{e}^H \mathbf{e} = (\mathbf{d}^H - \mathbf{h}^H \mathbf{X}^H)(\mathbf{d} - \mathbf{X}\mathbf{h}) \\ &= \mathbf{d}^H \mathbf{d} - \mathbf{h}^H \mathbf{X}^H \mathbf{d} - \mathbf{d}^H \mathbf{X}\mathbf{h} + \mathbf{h}^H \mathbf{X}^H \mathbf{X}\mathbf{h} \end{aligned} \quad (5-105)$$

根据第3章的求导公式,可得

$$\begin{aligned} \frac{\partial(\mathbf{d}^H \mathbf{d})}{\partial \mathbf{h}} &= 0, \quad \frac{\partial(\mathbf{h}^H \mathbf{X}^H \mathbf{d})}{\partial \mathbf{h}} = 0 \\ \frac{\partial(\mathbf{d}^H \mathbf{X}\mathbf{h})}{\partial \mathbf{h}} &= (\mathbf{X}^H \mathbf{d})^*, \quad \frac{\partial(\mathbf{h}^H \mathbf{X}^H \mathbf{X}\mathbf{h})}{\partial \mathbf{h}} = (\mathbf{X}^H \mathbf{X})^T \mathbf{h}^* \end{aligned}$$

由 $\partial E / \partial \mathbf{h} = 0$,可得

$$(\mathbf{X}^H \mathbf{X})\mathbf{h} = \mathbf{X}^H \mathbf{d} \quad (5-106)$$

$$\hat{\mathbf{h}} = (\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H \mathbf{d} \quad (5-107)$$

最小二乘解和式(5-18)中的 Wiener 解有相同的形式。定义如下基于时间平均的自相关函数估计和基于时间平均的互相关估计

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{x}(n) \mathbf{x}^H(n) \quad (5-108)$$

$$\hat{\mathbf{c}} = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{x}^*(n) d(n) \quad (5-109)$$

易知 $\hat{\mathbf{R}} = (\mathbf{X}^H \mathbf{X})^* / N$, $\hat{\mathbf{c}} = \mathbf{X}^H \mathbf{d} / N$, 利用共轭对称矩阵的特性,得到最小二乘解的另一种形式

$$\mathbf{h} = \hat{\mathbf{R}}^{-T} \hat{\mathbf{c}} \quad (5-110)$$

因此,可以认为最小二乘解是利用时间平均代替统计平均的 Wiener 解。将式(5-107)代入式(5-105)可以得到最小二乘解的最优误差为

$$E_{\min} = E_d - \hat{\mathbf{c}}^H \hat{\mathbf{R}}^{-1} \hat{\mathbf{c}} \quad (5-111)$$

可以证明,这和 Wiener 解的形式也是相同的。当采用加权最小二乘法时,误差重写为

$$E = \mathbf{e}^H \mathbf{e} = (\mathbf{d}^H - \mathbf{h}^H \mathbf{X}^H) \mathbf{W} (\mathbf{d} - \mathbf{X}\mathbf{h}) \quad (5-112)$$

其中, \mathbf{W} 是加权矩阵,其作用是体现不同误差分离重要性的差异性。加权最小二乘解的推导可采用相同的方法。

5.6 匹配滤波器

信噪比是雷达、声呐探测系统的基本指标,在雷达发展的初期,一般采用信噪比作为衡量雷达接收机抗干扰性能的指标。对于含有噪声的观测信号,接收滤波器的首要任务是降低噪声功率,即提高信噪比。1943年, North 从最大信噪比准则出发,建立了匹配滤波器理论。简单而言,匹配滤波器就是最大化滤波器输出的信号噪声功率比,并且这里的信号功率一般指信号的瞬时功率。

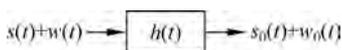


图 5-11 连续时间线性系统

首先以连续时间系统阐述匹配滤波器的含义。这里只考虑实信号,如图 5-11 所示用于降噪的线性系统,输入为

$$x(t) = s(t) + w(t) \quad (5-113)$$

其中, $s(t)$ 为确定性目标信号, $w(t)$ 为零均值高斯白噪声, 方差记为 σ^2 。滤波器是线性系统, 它满足迭加原理。因而, 可得到滤波器的输出为

$$y(t) = s_0(t) + w_0(t) \quad (5-114)$$

注意到 $s_0(t)$ 是确定性信号, 根据线性时不变系统特性, 可表示为

$$s_0(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\Omega) H(\Omega) e^{j\Omega t} d\Omega \quad (5-115)$$

$w_0(t)$ 是随机信号, 根据第 4 章所推导的线性时不变系统特性, $w_0(t)$ 的功率谱为

$$S_{w_0}(\Omega) = S_w(\Omega) |H(\Omega)|^2 = \sigma^2 |H(\Omega)|^2 \quad (5-116)$$

得到噪声的平均功率为

$$E[w_0^2(t)] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \sigma^2 |H(\Omega)|^2 d\Omega \quad (5-117)$$

定义在某个时刻滤波器输出端信号的瞬时功率与噪声的平均功率比值为

$$D_{t_0} = \frac{s_0^2(t_0)}{E[w_0^2(t)]} \quad (5-118)$$

把式(5-115)、式(5-117)代入, 得到

$$D_{t_0} = \frac{1}{2\pi} \frac{\left| \int_{-\infty}^{+\infty} S(\Omega) H(\Omega) e^{j\Omega t_0} d\Omega \right|^2}{\int_{-\infty}^{+\infty} \sigma^2 |H(\Omega)|^2 d\Omega} \quad (5-119)$$

要求 D_{t_0} 的最大值, 相当于求上确界, 可利用如下许瓦兹不等式

$$\left| \int_{-\infty}^{+\infty} a(\Omega) b^*(\Omega) d\Omega \right|^2 \leq \int_{-\infty}^{+\infty} |a(\Omega)|^2 d\Omega \int_{-\infty}^{+\infty} |b(\Omega)|^2 d\Omega \quad (5-120)$$

其中, $a(\Omega)$, $b(\Omega)$ 是积分有限的任意函数, 上式中等号成立的充分必要条件是

$$a(\Omega) = cb^*(\Omega) \quad (5-121)$$

其中, c 为常数。在式(5-120)中, 令

$$a(\Omega) = \sigma e^{j\Omega t_0} H(\Omega), \quad b(\Omega) = S(\Omega)/\sigma \quad (5-122)$$

则

$$\left| \int_{-\infty}^{+\infty} S(\Omega) H(\Omega) e^{j\Omega t_0} d\Omega \right|^2 \leq \int_{-\infty}^{+\infty} \sigma^2 |H(\Omega)|^2 d\Omega \int_{-\infty}^{+\infty} (|S(\Omega)|^2/\sigma^2) d\Omega \quad (5-123)$$

代入式(5-119),得到

$$D_{t_0} \leq \frac{1}{2\pi\sigma^2} \int_{-\infty}^{+\infty} |S(\Omega)|^2 d\Omega \quad (5-124)$$

要使上式成立,只需令

$$H(\Omega) = c e^{-j\Omega t_0} S^*(\Omega) / \sigma^2 \quad (5-125)$$

可以看到,这时滤波器是由目标信号完全确定的,因此 $h(t)$ 称为匹配滤波器。接下来考虑离散时间线性系统,见图 5-12。以从时间域推导匹配滤波器表达式的方式进行讨论。假设输入信号为有限长度序列 $s(n)$ ($n = 0, 1, \dots, N-1$), 离散时间线性滤波器的单位样值响应为 $h(n)$, $h(n)$ 在区间 $s(n)$ ($n = 0, 1, \dots, N-1$) 上取值,在该区间之外为零,噪声 $w(n)$ 为零均值高斯白噪声,方差为 σ^2 。

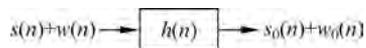


图 5-12 离散时间线性系统示意图

根据离散时间线性系统理论,输出信号 $s_0(n)$ 为

$$s_0(n) = \sum_{k=0}^{N-1} s(k)h(n-k) \quad (5-126)$$

输出噪声信号的平均功率为(参见习题 4.1)

$$E(|w_0[n]|^2) = \sigma^2 \sum_{k=0}^{N-1} |h(k)|^2 \quad (5-127)$$

任意取一时刻,不失一般性,取时刻 $N-1$,瞬时功率为 $|s_0(N-1)|^2$ 。定义离散时间线性滤波器的输出信噪比为

$$d_0 = \frac{|s_0(N-1)|^2}{E[w_0^2(n)]} \quad (5-128)$$

将式(5-126)和式(5-127)代入式(5-128),得

$$d_0 = \frac{\left| \sum_{k=0}^{N-1} s(k)h(N-1-k) \right|^2}{\sigma^2 \sum_{k=0}^{N-1} h^2(k)} \quad (5-129)$$

令 $\mathbf{s} = [s(0) \ s(1) \ \dots \ s(N-1)]^T$, $\mathbf{h} = [h(N-1) \ h(N-2) \ \dots \ h(0)]^T$, 则

$$d_0 = \frac{1}{\sigma^2} \frac{|\mathbf{h}^T \mathbf{s}|^2}{\mathbf{h}^H \mathbf{h}} = \frac{1}{\sigma^2} \frac{|(\mathbf{h}^*)^H \mathbf{s}|^2}{\mathbf{h}^H \mathbf{h}} \quad (5-130)$$

根据许瓦兹不等式可得

$$|(\mathbf{h}^*)^H \mathbf{s}|^2 \leq ((\mathbf{h}^*)^H (\mathbf{h}^*)) (\mathbf{s}^H \mathbf{s}) = (\mathbf{h}^H \mathbf{h}) (\mathbf{s}^H \mathbf{s}) \quad (5-131)$$

当且仅当 $\mathbf{h}^* = c\mathbf{s}$ 时,式(5-131)等号成立,其中 c 为常数,因此

$$d_0 \leq \frac{1}{\sigma^2} \frac{(\mathbf{h}^H \mathbf{h}) (\mathbf{s}^H \mathbf{s})}{\mathbf{h}^H \mathbf{h}} = \frac{1}{\sigma^2} (\mathbf{s}^H \mathbf{s}) \quad (5-132)$$

令 $E = \mathbf{s}^H \mathbf{s} = \sum_{n=0}^{N-1} |s[n]|^2$ 代表信号的能量,所以

$$d_0 \leq \frac{1}{\sigma^2} (\mathbf{s}^H \mathbf{s}) = \frac{E}{\sigma^2} \quad (5-133)$$

当且仅当 $\mathbf{h}^* = c\mathbf{s}$ 时式(5-133)等号成立, c 只影响滤波器的放大倍数,可令 $c=1$,所以,当

$$h^*(N-1-n) = s(n), \quad n = 0, 1, \dots, N-1 \quad (5-134)$$

时,输出信噪比达到最大,或等价于

$$h(n) = s^*(N-1-n), \quad n = 0, 1, \dots, N-1 \quad (5-135)$$

输出的最大信噪比为

$$d_{0\max} = E/\sigma^2 \quad (5-136)$$

例 5-3 假定输入信号为

$$s(n) = \begin{cases} 1, & 0 \leq n \leq 3 \\ 0, & \text{其他} \end{cases} \quad (5-137)$$

求匹配滤波器的单位样值响应和匹配滤波器的输出信号。

解 信号长度 $N=4$, 则匹配滤波器的单位样值响应为

$$h(n) = s(N-1-n) = \begin{cases} 1, & 0 \leq n \leq 3 \\ 0, & \text{其他} \end{cases} \quad (5-138)$$

输出信号为

$$s_0(n) = \sum_{k=0}^{N-1} s(k)h(n-k) = \begin{cases} n+1, & 0 \leq n \leq 3 \\ 7-n, & 4 \leq n \leq 6 \\ 0, & \text{其他} \end{cases} \quad (5-139)$$

输入信号、匹配滤波器的单位样值响应和输出信号的波形如图 5-13 所示。

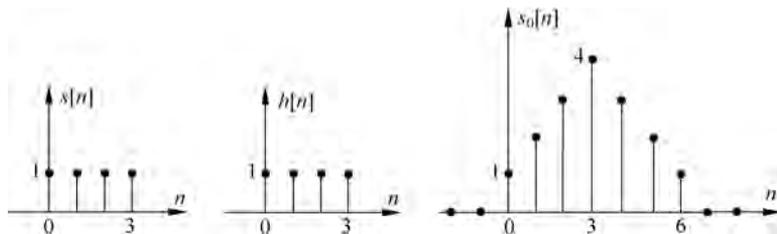


图 5-13 匹配滤波器输出

5.7 自适应滤波

5.7.1 自适应滤波器的基本概念

Wiener 滤波器、卡尔曼滤波器都是以已知信号和噪声的统计特征为基础,具有固定的滤波器系数。因此,仅当实际输入信号的统计特征与设计滤波器所依据的先验信息一致时,这类滤波器才是最佳的。否则,这类滤波器不能提供最佳性能。在实际中,往往难以预知这些统计特性,故实现不了真正的最佳滤波。在没有任何关于信号和噪声的先验知识的条件下,自适应滤波器利用前一时间已获得的滤波器参数自动调节现时刻的滤波器参数,以适应信号和噪声未知或随机变化的统计特性,从而实现最优滤波。

自适应信号处理的研究工作始于 20 世纪中叶。美国通用电气公司研究出了简单的自适应滤波器,用以消除混杂在有用信号中的噪声和干扰。Widrow B. 等于 1967 年提出的自适应滤波理论,可使自适应滤波系统的参数自动地调整达到最佳状况,而且在设计时,只需

要很少的或根本不需要任何关于信号与噪声的先验统计知识,这种滤波器的实现差不多像 Wiener 滤波器那样简单,而滤波性能几乎与卡尔曼滤波器一样好。20 世纪 70 年代中期, Widrow B. 等提出自适应滤波器及其算法,发展了最佳滤波设计理论。

自适应信号处理的应用领域包括通信、雷达、声呐、地震学、导航系统、生物医学和工业控制等。自适应滤波器是相对固定滤波器而言的,固定滤波器属于经典滤波器,它滤波的频率是固定的,自适应滤波器滤波的频率则是自动适应输入信号而变化的,所以其适用范围更广。自适应滤波器的一般结构框图如图 5-14 所示,其中 $d(n)$ 为滤波器输出期望信号, $x(n)$ 为滤波器输入信号, $e(n)$ 为滤波器输出和期望信号之间的误差信号。自适应滤波器以当前时刻的滤波器输出为基础,以误差最小化为准则,更新当前时刻的滤波器系数。

自适应滤波器根据应用环境的不同有不同的结构。下面结合实际应用介绍几种自适应滤波器的实现结构。图 5-15 给出了自适应噪声消除的原理图。 $S(n)$ 为噪声干扰下的目标信号, n_0, n_1 为互不相关的白噪声。自适应噪声消除的目的是将信号 $S(n)$ 从噪声中分离出来,其基本原理是将参考噪声 n_1 经过滤波器处理后接近实际噪声 n_0 。

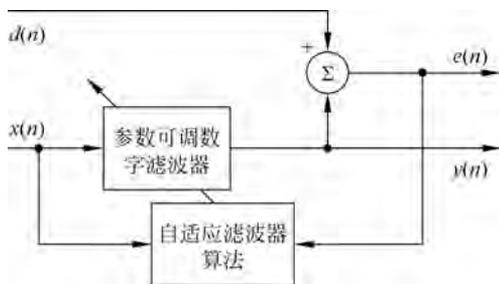


图 5-14 自适应滤波器的结构示意图

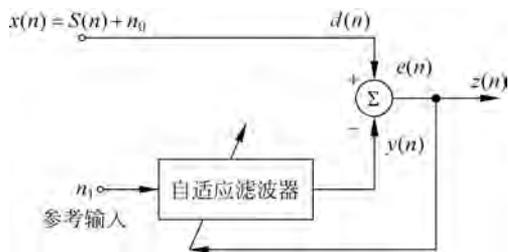


图 5-15 自适应噪声消除原理图

自适应滤波在生物医学信号处理中的一个典型应用是“胎儿心电图自适应干扰清除”,如图 5-16 所示。其中目标信号——胎儿心电信号受到母亲心电信号的干扰,消除这种干扰的一种有效处理方式是将母亲胸部探头得到的心电信号(基本不包含胎儿心电信号)作为参考信号,经过自适应滤波器的滤波后接近胎儿心电信号中的干扰信号。

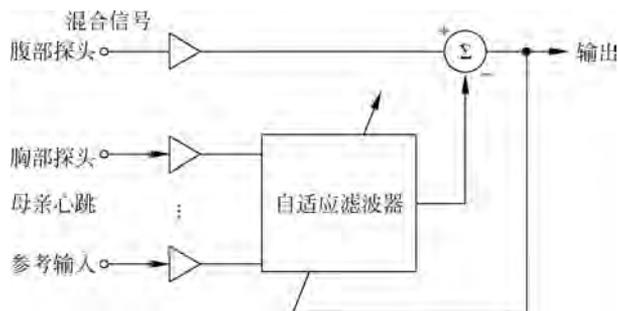


图 5-16 胎儿心电图自适应干扰清除

下面两小节介绍两种经典的自适应滤波算法,为了简化讨论,只考虑实信号。

5.7.2 LMS 自适应滤波器

本节介绍 LMS(Least Mean Square, 最小均方)自适应滤波器。

考虑阶数为 N 的 FIR 自适应滤波器

$$y(n) = \sum_{k=0}^N h(k)x(n-k) \quad (5-140)$$

为了下面推导方便,引入信号的向量描述如下:

$$\begin{aligned} \mathbf{h} &= [h(0), h(1), \dots, h(N)]^T \\ \mathbf{x}(n) &= [x(n), x(n-1), \dots, x(n-N)]^T \end{aligned} \quad (5-141)$$

误差信号为

$$\begin{aligned} e(n) &= d(n) - y(n) \\ &= d(n) - \mathbf{h}^T \mathbf{x}(n) \end{aligned} \quad (5-142)$$

均方误差为

$$\begin{aligned} \xi &= E(|e(k)|^2) \\ &= E[d^2(k) - 2d(k)\mathbf{x}^T(k)\mathbf{h} + \mathbf{h}^T \mathbf{x}(k)\mathbf{x}^T(k)\mathbf{h}] \\ &= E[d^2(k)] + \mathbf{h}^T E[\mathbf{x}(k)\mathbf{x}^T(k)]\mathbf{h} - 2E[d(k)\mathbf{x}^T(k)]\mathbf{h} \end{aligned} \quad (5-143)$$

记 $\mathbf{R}_x = E(\mathbf{x}(k)\mathbf{x}^T(k))$ 为滤波器输入信号的自相关矩阵, $\mathbf{r}_{dx} = E[d(k)\mathbf{x}^T(k)]$ 为输入信号和期望信号的互相关向量, 得到均方误差的简洁表达式为

$$\xi = E(|e(k)|^2) = \mathbf{h}^T \mathbf{R}_x \mathbf{h} - 2\mathbf{r}_{dx}^T \mathbf{h} + D_d \quad (5-144)$$

当输入信号是广义平稳的随机信号时, 由上式可见 ξ 是滤波器系数 \mathbf{h} 的二次函数。我们称 $\xi(\mathbf{h})$ 为误差性能函数, 因为 \mathbf{R}_x 是半正定矩阵, 误差性能函数的曲面是一个向下凹的抛物面, 典型的二维均方误差性能函数如图 5-17 所示。

从图 5-17 可以看到误差性能函数有唯一最小点, 通过对滤波器系数求偏导数并令偏导数为零可求出滤波器系数, 这也是 Wiener 滤波所采用的方法。在自适应滤波中, 我们希望得到一种迭代求解的形式, 使得滤波器系数可实时跟踪信号统计特性的变化。由于误差性能函数是二次函数, 连续可导并且有唯一全局最小点, 因此可以用降低的梯度下降法进行迭代。梯度下降法的示意图如图 5-18 所示, 其基本思想是: 要达到最优解, 不管初始权值如何选择, 只要在调整过程中, 权值的调整沿 $\xi(\mathbf{h})$ 的负梯度方向进行, 就可以保证最终收敛到最优解。

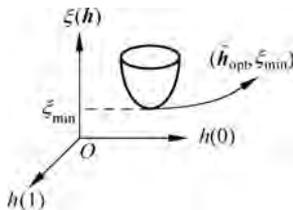


图 5-17 二维均方误差性能函数

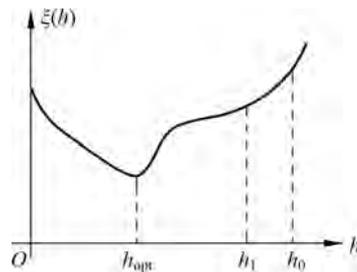


图 5-18 梯度下降法示意图

可见, 梯度法可表示为如下的迭代公式:

$$\mathbf{h}_{n+1} = \mathbf{h}_n + \mu(-\nabla \xi(\mathbf{h}) |_{\mathbf{h}=\mathbf{h}_n}) \quad (5-145)$$

这里引入时间 n , \mathbf{h}_n 表示时刻 n 的滤波器系数(第 n 次迭代求解的结果), μ 为迭代步长, 一般是小于 1 的常数。式(5-145)的梯度为

$$\nabla \xi(\mathbf{h}) |_{\mathbf{h}=\mathbf{h}_n} = 2\mathbf{R}_x \mathbf{h}_n - 2\mathbf{r}_{dx} \quad (5-146)$$

式(5-146)需计算自相关矩阵, 这种梯度计算方法不能应用于非平稳信号, 和自适应滤波的初衷不一致。Widrow 等提出了一种近似的计算梯度的有效方法, 其原理是用当前时刻的误差 $|e(n)|^2$ 代替均方误差 $E(|e(n)|^2)$ 。此时有

$$\hat{\xi}(\mathbf{h}) = |e(n)|^2 \quad (5-147)$$

$$e(n) = d(n) - \mathbf{h}^T \mathbf{x}(n) \quad (5-148)$$

$$\nabla \hat{\xi}(\mathbf{h}) |_{\mathbf{h}=\mathbf{h}_n} = -2e(n)\mathbf{x}(n) \quad (5-149)$$

容易证明 $\nabla \hat{\xi}(\mathbf{h}) |_{\mathbf{h}=\mathbf{h}_n}$ 是 $\nabla \xi(\mathbf{h}) |_{\mathbf{h}=\mathbf{h}_n}$ 的无偏估计。此时梯度法迭代公式为

$$\mathbf{h}_{n+1} = \mathbf{h}_n + 2\mu e(n)\mathbf{x}(n) \quad (5-150)$$

此算法称为 Widrow-Hopf LMS 算法, 计算非常简单, 并且不涉及信号的统计量, 每个时刻的更新迭代只需当前接收信号及期望信号, 符合自适应滤波实时更新信号特征的初衷。

LMS 算法的实现步骤如下:

步骤 1: 初始化, 选择合适的步长 μ , 初始化 \mathbf{h}_0 的值(一般为零或随机初始化), 设 $n=0$;

步骤 2: 迭代更新, 如式(5-148)和式(5-150)所示:

$$e(n) = d(n) - \mathbf{h}_n^T \mathbf{x}(n)$$

$$\mathbf{h}_{n+1} = \mathbf{h}_n + 2\mu e(n)\mathbf{x}(n)$$

步骤 3: 判断是否收敛, 如果不收敛, 令 $n=n+1$, 返回步骤 2, 一般如果滤波器在连续两次的更新小于某个门限值则认为算法收敛。

LMS 算法简单, 易于软硬件实现, 应用非常广泛。从算法流程可以预见 μ 值的选择对算法有重要影响。 μ 越大, 收敛越快, 但容易振荡, μ 值小, 收敛慢, 但比较平稳。

下面基于式(5-146)分析 μ 值的选择对算法稳定性的影响。根据式(5-146)把更新算法重写为

$$\mathbf{h}_{n+1} = \mathbf{h}_n - 2\mu(\mathbf{R}_x \mathbf{h}_n - \mathbf{r}_{dx}) \quad (5-151)$$

注意到最优滤波器系数(记为 \mathbf{h}_{opt})满足 Wiener-Hopf 方程 $\mathbf{R}_x \mathbf{h}_{\text{opt}} = \mathbf{r}_{dx}$, 式(5-151)可改写为

$$\mathbf{h}_{n+1} = (\mathbf{I} - 2\mu\mathbf{R}_x)\mathbf{h}_n + 2\mu\mathbf{R}_x\mathbf{h}_{\text{opt}} \quad (5-152)$$

其中, \mathbf{I} 是大小和 \mathbf{R}_x 相等的单位矩阵。令 $\mathbf{v}_n = \mathbf{h}_n - \mathbf{h}_{\text{opt}}$, 式(5-152)可写为

$$\mathbf{v}_{n+1} = (\mathbf{I} - 2\mu\mathbf{R}_x)\mathbf{v}_n \quad (5-153)$$

记 \mathbf{R}_x 的特征向量矩阵为 \mathbf{Q} , 得到的特征值分解为

$$\mathbf{R}_x = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1} \quad (5-154)$$

其中, $\mathbf{\Lambda}$ 是包含 \mathbf{R}_x 特征值的对角矩阵, 由于实随机信号的自相关矩阵是对称矩阵, 即 $\mathbf{R}_x = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1} = \mathbf{R}_x^T = \mathbf{Q}^{-T}\mathbf{\Lambda}\mathbf{Q}^T$, 所以

$$\mathbf{Q}^{-1} = \mathbf{Q}^T \quad (5-155)$$

式(5-153)可重写为

$$\begin{aligned} \mathbf{v}_{n+1} &= (\mathbf{Q}\mathbf{Q}^{-1} - 2\mu\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1})\mathbf{v}_n \\ &= \mathbf{Q}(\mathbf{I} - 2\mu\mathbf{\Lambda})\mathbf{Q}^{-1}\mathbf{v}_n \end{aligned} \quad (5-156)$$

式(5-156)左乘 \mathbf{Q}^{-1} ,并记 $\bar{\mathbf{v}}_n = \mathbf{Q}^{-1} \mathbf{v}_n$,得到

$$\bar{\mathbf{v}}_{n+1} = (\mathbf{I} - 2\mu\mathbf{\Lambda}) \bar{\mathbf{v}}_n \quad (5-157)$$

递归应用式(5-157)可得

$$\begin{aligned} \bar{\mathbf{v}}_{n+1} &= (\mathbf{I} - 2\mu\mathbf{\Lambda})^2 \bar{\mathbf{v}}_{n-1} \\ &= \begin{bmatrix} (1 - 2\mu\lambda_1)^n & & & \\ & (1 - 2\mu\lambda_2)^n & & \\ & & \ddots & \\ & & & (1 - 2\mu\lambda_{N-1})^n \end{bmatrix} \bar{\mathbf{v}}_0 \end{aligned} \quad (5-158)$$

要使算法稳定,应有

$$\lim_{n \rightarrow \infty} \bar{\mathbf{v}}_n = 0 \quad (5-159)$$

这时有 $\lim_{n \rightarrow \infty} \mathbf{Q}^{-1} \mathbf{v}_n = 0$,即

$$\lim_{n \rightarrow \infty} \mathbf{Q}^{-1} (\mathbf{h}_n - \mathbf{h}_{\text{opt}}) = 0 \quad (5-160)$$

即

$$\lim_{n \rightarrow \infty} \mathbf{h}_n = \mathbf{h}_{\text{opt}} \quad (5-161)$$

其收敛到最优解。

要使式(5-159)成立,必须有

$$\lim_{n \rightarrow \infty} (1 - 2\mu\lambda_i)^n = 0, \quad i = 0, 1, 2, \dots, N-1 \quad (5-162)$$

即

$$|1 - 2\mu\lambda_{\max}| < 1 \quad (5-163)$$

所以

$$0 < \mu < \frac{1}{\lambda_{\max}} \quad (5-164)$$

式中, λ_{\max} 是 \mathbf{R}_x 最大的特征值,这就是梯度法收敛的充分必要条件。一般而言,LMS算法收敛速度取决于 μ 和 λ_{\min} , μ 越大,收敛速度越快, λ_{\min} 越大,收敛速度越快。通常定义 $d = \lambda_{\max}/\lambda_{\min}$ 为谱动态范围, d 越大,收敛时间越长。

5.7.3 RLS 自适应滤波

RLS(Recursive Least Squares,递推最小二乘法)自适应滤波器的设计准则是从滤波器开始运行到当前时间,滤波器系数的更新是令总的平方误差达到最小,令 n 表示当前时间,总的平方误差一般定义为

$$\xi(n) = \sum_{j=0}^n \lambda^{n-j} |e(j)|^2 = \sum_{j=0}^n \lambda^{n-j} |d(j) - \mathbf{h}_n^T \mathbf{x}(j)|^2 \quad (5-165)$$

其中, $e(j)$ 是瞬时误差, $\lambda(0 < \lambda \leq 1)$ 称为遗忘因子,用于降低“旧”数据的影响,提高滤波器的跟踪能力。

最优滤波器系数将使得 $\xi(n)$ 最小,即满足以下方程

$$\frac{\partial \xi(n)}{\partial \mathbf{h}_n} = 0 \quad (5-166)$$

基于式(5-165)可得到包含最优滤波器系数的方程

$$\mathbf{R}(n)\mathbf{h}_n = \mathbf{c}(n) \quad (5-167)$$

其中

$$\mathbf{R}(n) = \sum_{j=0}^n \lambda^{n-j} \mathbf{x}(j)\mathbf{x}^T(j) \quad (5-168)$$

$$\mathbf{c}(n) = \sum_{j=0}^n \lambda^{n-j} \mathbf{x}(j)d(j) \quad (5-169)$$

如果 $\mathbf{R}(n)$ 是满秩矩阵, 求解方程(5-167)可得到滤波器系数, 相应的解称为最小二乘解。但是直接求解上述方程需要求矩阵的逆矩阵。下面推导求解的迭代形式, 首先 $\mathbf{R}(n)$ 的更新可以表示为

$$\mathbf{R}(n) = \lambda\mathbf{R}(n-1) + \mathbf{x}(n)\mathbf{x}^T(n) \quad (5-170)$$

$$\mathbf{c}(n) = \lambda\mathbf{c}(n-1) + \mathbf{x}(n)d(n) \quad (5-171)$$

将上述两式代入式(5-167), 得到

$$[\mathbf{R}(n) - \mathbf{x}(n)\mathbf{x}^T(n)]\mathbf{h}_{n-1} = \mathbf{c}(n) - \mathbf{x}(n)d(n) \quad (5-172)$$

经过简单计算, 得到

$$\mathbf{R}(n)\mathbf{h}_{n-1} + \mathbf{x}(n)[d(n) - \mathbf{h}_{n-1}^T\mathbf{x}(n)] = \mathbf{c}(n) \quad (5-173)$$

注意 $d(n) - \mathbf{h}_{n-1}^T\mathbf{x}(n)$ 刚好是应用前一时刻滤波器所造成的误差, 记为 $d(n) - \mathbf{h}_{n-1}^T\mathbf{x}(n) = e(n-1)$ 并在式(5-173)两边乘以 $\mathbf{R}^{-1}(n)$, 得到

$$\mathbf{h}_{n-1} + \mathbf{R}^{-1}(n)\mathbf{x}(n)e(n-1) = \mathbf{R}^{-1}(n)\mathbf{c}(n) = \mathbf{h}_n \quad (5-174)$$

定义如下自适应增益向量

$$\mathbf{g}(n) = \mathbf{R}^{-1}(n)\mathbf{x}(n) \quad (5-175)$$

最终得到

$$\mathbf{h}_n = \mathbf{h}_{n-1} + \mathbf{g}(n)e(n-1) \quad (5-176)$$

式(5-176)的更新依然涉及矩阵求逆, 但其中的求逆可用如下迭代公式求解:

$$(\mathbf{A} - \mathbf{x}\mathbf{x}^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{x}\mathbf{x}^T\mathbf{A}^{-1}}{1 + \mathbf{x}^T\mathbf{A}^{-1}\mathbf{x}} \quad (5-177)$$

将 $\mathbf{R}(n) = \lambda\mathbf{R}(n-1) + \mathbf{x}(n)\mathbf{x}^T(n)$ 代入式(5-177)得到

$$[\lambda\mathbf{R}(n-1) - \mathbf{x}(n)\mathbf{x}^T(n)]^{-1} = \lambda\mathbf{R}^{-1}(n-1) - \frac{\lambda^{-1}\mathbf{R}^{-1}(n-1)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{R}^{-1}(n-1)\lambda^{-1}}{1 + \lambda^{-1}\mathbf{x}^T(n)\mathbf{R}^{-1}(n-1)\mathbf{x}(n)} \quad (5-178)$$

定义

$$\mathbf{P}(n) = \mathbf{R}^{-1}(n) \quad (5-179)$$

可得到如下迭代更新

$$\mathbf{P}(n) = \lambda\mathbf{P}(n-1) - \frac{\lambda^{-1}\mathbf{P}(n-1)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{P}^T(n-1)\lambda^{-1}}{1 + \lambda^{-1}\mathbf{x}^T(n)\mathbf{P}(n-1)\mathbf{x}(n)} \quad (5-180)$$

最终得到最小二乘解的递归实现, 称为递归最小二乘(Recursive Least Square, RLS)算法, 具体步骤如下:

步骤 1: 初始化, 设 $\mathbf{R}(-1) = \mathbf{I}$, $\mathbf{P}(-1) = \mathbf{I}$, 选择合适的遗忘因子 λ 初始化的值(一般为小于 1 但接近 1 的实数), 初始化滤波器系数, 记为 \mathbf{h}_{-1} , 设 $n=0$;

步骤 2: 迭代更新,如下式所示:

$$\begin{aligned} e(n-1) &= d(n) - \mathbf{h}_{n-1}^T \mathbf{x}(n) \\ \mathbf{P}(n) &= \lambda \mathbf{P}(n-1) - \frac{\lambda^{-1} \mathbf{P}(n-1) \mathbf{x}(n) \mathbf{x}^T(n) \mathbf{P}^T(n-1) \lambda^{-1}}{1 + \lambda^{-1} \mathbf{x}^T(n) \mathbf{P}(n-1) \mathbf{x}(n)} \\ \mathbf{g}(n) &= \mathbf{P}(n) \mathbf{x}(n) \\ \mathbf{h}_n &= \mathbf{h}_{n-1} + \mathbf{g}(n) e(n-1) \end{aligned}$$

步骤 3: 判断是否收敛,如果不收敛,令 $n=n+1$,返回步骤 2,收敛判断可以用和 LMS 算法相同的准则。

总之,RLS 算法的计算复杂度高于 LMS 算法,但收敛速度较快。

本章习题

1. 设噪声中存在 L 个具有随机相位的复正弦信号

$$x_n = \sum_{i=1}^L A_i e^{j(\omega_i n + \phi_i)} + v_n$$

式中, $\phi_i, i=1, 2, \dots, L$ 为均匀分布的随机相位,它们是互相独立的; v_k 为零均值与 σ_v^2 方差的白噪声,且与 ϕ_i 互相独立。

(1) 证明 $E[e^{j\phi_i} e^{-j\phi_i}] = \delta_{ik}, k=1, 2, \dots, L$;

(2) 证明 x_n 的自相关函数

$$r_{xx}(k) = E[x_n x_{n-k}^*] = \sigma_v^2 \delta(k) + \sum_{i=1}^L |A_i|^2 e^{j\omega_i k};$$

(3) 将 x_n 通过一个传递函数为 $A(z) = a_0 + a_1 z^{-1} + \dots + a_M z^{-M}$ 的滤波器滤波,滤波后的输出为 y_n ,证明输出功率为

$$P_y = E[y_n^* y_n] = \mathbf{a}^H \mathbf{R}_{xx} \mathbf{a} = \sigma_v^2 \mathbf{a}^H \mathbf{a} + \sum_{i=1}^L |A_i|^2 |A(\omega_i)|^2$$

式中

$$\begin{aligned} \mathbf{a} &= [a_0, a_1, \dots, a_M]^T \\ \mathbf{R}_{xx} &= [r_{xx}(i, j)] = [r_{xx}(i - j)] \\ A(\omega_i) &= \sum_{m=0}^M a_m e^{-j\omega_i m} \end{aligned}$$

2. 接收信号为 $x(t) = s(t) + v(t)$, $s(t)$ 、 $v(t)$ 均为零均值平稳信号,目标信号自相关函数为 $R_{ss}(\tau) = e^{-|\tau|}$,噪声的自相关函数为 $R_{vv}(\tau) = e^{-2|\tau|}$,信号与噪声互不相关,请根据 5.2.2 节的思路导出因果连续 IIR Wiener 滤波器。

3. 设系统模型为 $x_n = 0.6x_{n-1} + \omega_n$,观测方程为 $z_n = x_n + v_n$,其中 ω_n 为方差 $\sigma_\omega^2 = 0.82$ 的白噪声, v_n 为方差 $\sigma_v^2 = 1$ 的白噪声, v_n 与 x_n 不相关。试求其阶数为 2 的离散 FIR Wiener 滤波器并计算其均方误差。

4. 利用第 3 题的已知条件,计算因果 IIR Wiener 滤波器和非因果 IIR Wiener 滤波器,计算相应的均方误差并和第 3 题的 FIR 滤波结果进行比较。

5. 非因果 FIR 线性滤波器 $h(n)$ 的输入/输出关系为

$$y(n) = \sum_{k=-L}^L h(k)x(n-k)$$

设目标信号为 $d(n)$, 误差信号为 $e(n) = d(n) - y(n)$, 试推导该非因果 FIR 线性滤波器的 Wiener-Hopf 正则方程。

6. 在以下 AR(p) 模型中

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

$u(n)$ 是零均值, 方差为 σ_u^2 的高斯白噪声。证明以下线性预测是线性最优预测并求出预测误差。

$$\hat{x}(n) = - \sum_{k=1}^p a_k x(n-k)$$

7. 证明阶数相同的线性前向预测系数和后向预测系数互为共轭。

8. 一维平稳随机信号 $x(n)$ 的系统状态方程和观测方程分别为

$$\begin{aligned} s(n) &= 0.5s(n-1) + w(n) \\ x(n) &= s(n) + v(n) \end{aligned}$$

其中, $w(n)$ 、 $v(n)$ 为白噪声, 方差均为 1。 $s(n)$ 、 $w(n)$ 、 $v(n)$ 两两互不相关。试导出相应的卡尔曼迭代滤波公式, 分析当 $k \rightarrow \infty$ 时, 预测参数的收敛值, 并和最佳线性预测器进行比较。

9. $x(n)$ 的标量系统状态方程和观测方程定义为

$$\begin{aligned} s(n) &= as(n-1) + w(n) \\ x(n) &= s(n) + v(n) \end{aligned}$$

写出从 $x(n)$ 恢复出 $s(n)$ 的卡尔曼滤波迭代计算过程, 求出等效的系统传输函数(由 $x(n)$ 到 $s(n)$ 的传输函数)。

10. 对第 9 题中的标量状态空间模型, 试分析其预测误差随时间增长的变化情况, 是否存在随时间增长, 预测误差逐渐下降的趋势?

11. 如下时变状态空间模型

$$\begin{aligned} \mathbf{x}_{n+1} &= \mathbf{F}_n \mathbf{x}_n + \mathbf{w}_n \\ \mathbf{y}_{n+1} &= \mathbf{H}_{n+1} \mathbf{x}_{n+1} + \mathbf{v}_{n+1} \end{aligned}$$

证明存在无观测噪声的等价状态空间模型, 如下式所示

$$\begin{aligned} \mathbf{x}'_{n+1} &= \mathbf{F}'_n \mathbf{x}'_n + \mathbf{w}'_n \\ \mathbf{y}_{n+1} &= \mathbf{H}'_{n+1} \mathbf{x}'_{n+1} \end{aligned}$$

12. 基于式(5-105)的加权最小二乘滤波误差为

$$E = \mathbf{e}^H \mathbf{e} = (\mathbf{d}^H - \mathbf{h}^H \mathbf{X}^H) \mathbf{W} (\mathbf{d} - \mathbf{X} \mathbf{h})$$

试分别针对实数模型和复数模型, 推导最佳滤波系数及其对应的误差。

13. 线性滤波器的输入为 $x(t) = s(t) + v(t)$, $v(t)$ 为高斯白噪声, 功率谱密度为 N_0 , $s(t)$ 为方波信号, 定义如下:

$$s(t) = \begin{cases} A, & 0 \leq t \leq A \\ 0, & \text{其他} \end{cases}$$

设计匹配滤波器并求最大输出信噪比。当采用如下滤波器进行降噪时,输出最大信噪比是多大? 要使信噪比尽量大,如何选取参数 a ?

$$h(t) = \begin{cases} e^{-at}, & 0 \leq t \leq A \\ 0, & \text{其他} \end{cases}$$

14. 考虑随机信号的匹配滤波问题,假设接收信号模型为

$$x(n) = s(n) + v(n)$$

其中, $s(n)$ 为平稳随机信号, $w(n)$ 为平稳噪声, $s(n)$ 与 $w(n)$ 相互独立且其自相关矩阵已知,分别记为 \mathbf{R}_{ss} 和 \mathbf{R}_{vv} 。考虑线性因果 FIR 滤波器 $h(n)$, $0 \leq n \leq L$ 。证明滤波器输出可表示为

$$y(n) = \mathbf{h}^H \mathbf{x}(n)$$

其中, $\mathbf{h} = [h(0), \dots, h(L)]^T$, $\mathbf{x}(n) = [x(n), \dots, x(n-L)]^T$, 证明滤波器输出信号的信噪比可表示为

$$\text{SNR} = \frac{\mathbf{h}^H \mathbf{R}_{ss} \mathbf{h}}{\mathbf{h}^H \mathbf{R}_{vv} \mathbf{h}}$$

15. 利用第 14 题的结论,证明当噪声是白噪声时,使输出信噪比达到最大的匹配滤波器为 \mathbf{R}_{ss} 最大特征值对应的特征向量。以上结论是否可拓展到有色噪声?

16. 考虑如图 5-19 所示单权自适应线性组合器。设开关 S 是断开的, $E[x^2(n)] = 1$, $E[x(n)x(n-1)] = 0.5$, $E[d^2(n)] = 4$, $E[d(n)x(n)] = -1$, $E[d(n)x(n-1)] = 1$, 试导出性能函数表达式,并给出性能函数的图形。

17. 如图 5-19 所示,设开关 S 闭合,按第 16 题的要求再做一次。当只有一个权因子时,性能函数表达式表示什么类型的曲线?

18. 考虑如图 5-20 所示自适应噪声对消系统,试:

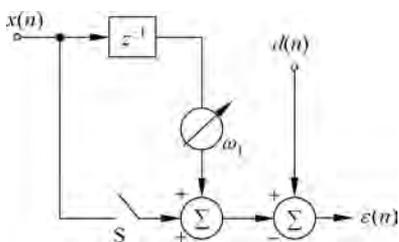


图 5-19 单权自适应线性组合器

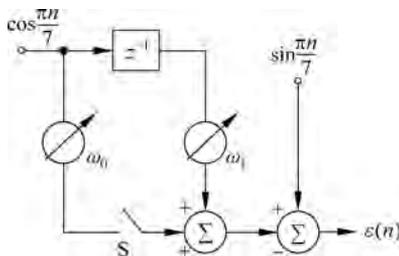


图 5-20 自适应噪声对消系统

- (1) 写出性能表面表达式;
- (2) 确定自适应增益的范围;
- (3) 写出这种情况下的 LMS 算法。

19. 给定如图 5-21 所示系统辨识结构,试给出递归 LMS 算法。

20. 对于如图 5-22 所示逆模拟系统,设对所有 $i \neq n$,有 $s(n)$ 与 $s(i)$ 相互独立,且 $R_{ss}(0) = 1$,同时设 $n_0(n)$ 与 $s(n)$ 是相互独立的白噪声,且 $R_{nn}(0) = P_n$,试推导以下功率谱的表达式: $S_{ss}(z)$ 、 $S_{xx}(z)$ 、 $S_{dd}(z)$ 、 $S_{dx}(z)$ 。

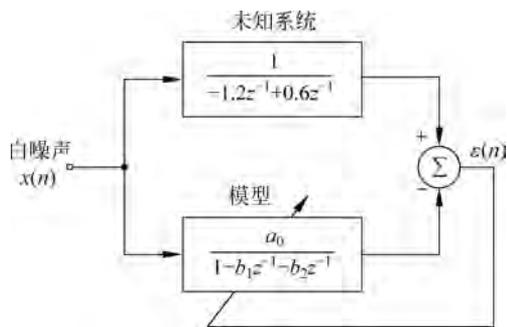


图 5-21 系统辨识结构

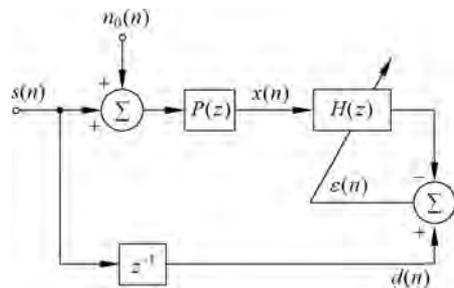


图 5-22 逆模拟系统

21. 信号系统结构如图 5-23 所示,有一部分信号泄漏到参考信道,设 $n_0(n)$ 是总功率为 N 的白噪声(独立)。试用输入功率谱 $S_{ss}(z)$ 表示最佳 $W^n(z)$ 。

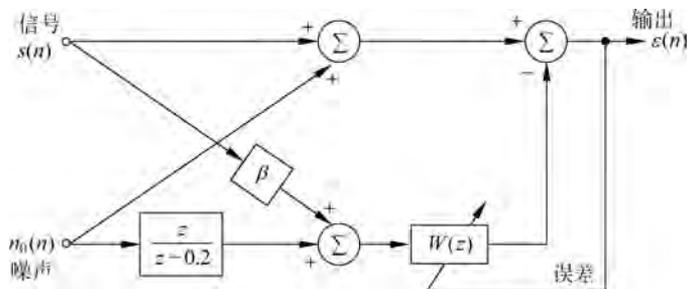


图 5-23 信号系统结构