

## 背景与预览

### 滤波问题

估计器或滤波器这一术语通常用来称呼一个系统，设计这样的系统是为了从含有噪声的数据中提取人们感兴趣的、接近规定质量的信息。由于这样一个宽目标，估计理论应用于诸如通信、雷达、声纳、导航、地震学、生物医学工程、金融工程等众多不同领域。例如，考虑一个数字通信系统，其基本形式由发射机、信道和接收机连接组成，如图 1 所示。发射机的作用是把数字源（例如计算机）产生的由 0、1 符号序列组成的消息信号变换为适合于信道上传送的波形。典型地，信道主要受到下列两种损伤：

- 符号间干扰 理想上 线性传输媒介定义为

$$h(t) = A\delta(t - \tau) \quad (1)$$

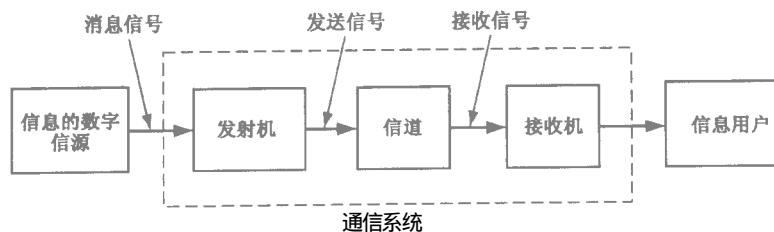
式中  $t$  表示连续时间， $h(t)$  表示脉冲响应， $A$  为幅度标度因子， $\delta(t)$  是 Dirac  $\delta$  函数（即单位脉冲函数）， $\tau$  表示信号沿信道传送过程中所产生的传播时延。式 (1) 是某一理想传输媒介的时域描述。等效地，我们可以在频域表征它，并写为

$$H(j\omega) = A \exp -j\omega\tau \quad (2)$$

式中  $j$  是  $-1$  的平方根， $\omega$  表示角频率， $H(j\omega)$  是传输媒介的频率响应，而  $\exp \cdot$  表示指数函数。实际上，对任何物理信道，不可能满足由式 (1) 给出的理想时域描述 [ 或式 (2) 等效频域描述 ] 所包含的严格要求。我们尽最大努力所能做到的，也只是在表示发送信号基本谱内容的频带上逼近式 (2)，它将使物理信道发生色散 (dispersive)。在数字通信系统中，这种信道损伤将引起符号间干扰 (intersymbol interference)，由此造成紧接着的脉冲（表示 1、0 发送序列）相互间模糊不清，以致于它们不再可区分。

- 噪声 某种形式的噪声出现在每个通信信道的输出端。该噪声可以是系统内部的（由接收机前端放大器产生的热噪声）或系统外部的（由其他信源产生的干扰信号）

这两种损伤的最终结果是，信道输出端收到的信号是含有噪声的或失真的发送信号。接收机的作用是，操作接收信号并把原消息信号的一个可靠估值传递给系统输出端的某个用户。



作为涉及滤波器理论应用的另一个例子，考虑图 2 所示的情况。图中表示一个连续时间

动态系统 其  $t$  时刻的状态由多维向量  $\mathbf{x}(t)$  表示。描述状态  $\mathbf{x}(t)$  演变的方程通常受到系统误差的影响。滤波问题是复杂的, 因为  $\mathbf{x}(t)$  是隐蔽的而且能够观测它的惟一方法是通过间接测量 其测量方程是状态  $\mathbf{x}(t)$  自身的函数。再则, 该测量方程不可避免地受到它所拥有的噪声的影响。图 2 所描述的动态系统可以是飞行中的飞机, 在这种情况下, 飞机的位置和速度组成状态  $\mathbf{x}(t)$  的元素, 而测量系统可以是跟踪雷达。总之, 给定测量系统在间隔  $[0, T]$  内产生的可测向量  $\mathbf{y}(t)$  及先验信息 我们的要求是估计动态系统的状态  $\hat{\mathbf{x}}(t)$ 。

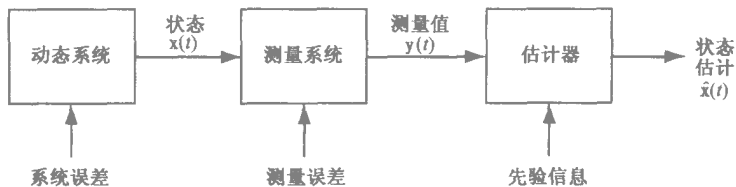


图 2 描绘状态估计中所涉及成分的框图

上述两个例子所阐明的估计理论实际上是统计的, 因为不可避免地存在影响所研究系统正常运行的噪声或系统误差。

### 三种基本估计

三种基本的信息处理运算是滤波、平滑和预测 每一种运算由某一估计器来完成。这些运算之间的差异用图 3 来说明。

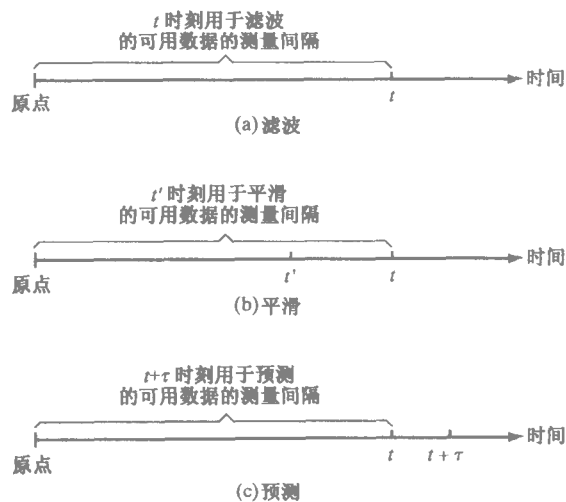


图 3 估计基本形式示意图

- 滤波 (filtering) 它是用  $t$  时刻及以前的数据来提取或估计  $t$  时刻感兴趣信息的一种运算过程。
- 平滑 (smoothing) 它是一种后验形式的估计, 因为在感兴趣时刻之后观测数据用于这种估计。特别地, 时刻  $t'$  的平滑估计可利用时间间隔  $[0, t']$  ( $t' < t$ ) 内的观测数据来获得。因而在平滑估计的计算中, 包含有  $t - t'$  的时延。由等待更多数据积累所获得的效益是, 平滑能够得到比滤波更精确的估值。
- 预测 (prediction) 它是估计的预测方面。其目的是用  $t$  时刻及以前的数据来推出未来

某些  $t + \tau$  (对某些  $\tau > 0$ ) 时刻感兴趣信息的一种估计。

由图 3 可以明显看出, 滤波和预测是实时运算, 而平滑是非实时的。实时运算意指在该运算中, 以当前可得到的数据为基础完成人们感兴趣的估计。

## 线性最优滤波器

滤波器可分为线性滤波器和非线性滤波器两种。若滤波器输出端滤波、平滑或预测的量是它的输入观测量的线性函数 则认为该滤波器是线性的 否则 认为该滤波器是非线性的。

在解线性滤波问题的统计方法中, 通常假设已知有用信号及其附加噪声的某些统计参数 (例如 均值和自相关函数) 而且需要设计含噪数据作为其输入的线性滤波器 使得根据某种统计准则噪声对滤波器的影响最小。实现该滤波器优化问题的一个有用方法是使误差信号 (定义为期望响应与滤波器实际输出之差) 的均方值最小化。对于平稳输入, 其解决方案通常称为维纳滤波器 (Wiener filter)。该滤波器在均方误差意义上是最优的。误差信号均方值相对于线性滤波器可调参数的曲线通常称为误差性能曲面。该曲面的极小点即为维纳解。

维纳滤波器不适合于应对信号和 / 或噪声的非平稳性是问题所固有的情况。在这种情况下, 必须假设最优滤波器为时变形式。对于这个更加困难的问题, 十分成功的一个解决方案是采用卡尔曼滤波器 (Kalman filter)。该滤波器在各种工程应用中是一个强有力的系统。

包括维纳滤波器和卡尔曼滤波器的线性滤波器理论已经在连续时间信号和离散时间信号文献中获得广泛的研究。然而, 由于数字计算机的广泛普及和数字信号处理器件与日俱增的应用等技术原因, 离散时间线性滤波器通常更为人们所乐意使用。因此, 在后续章节中, 我们仅仅考虑离散时间形式的维纳滤波器和卡尔曼滤波器。在这种形式中, 输入和输出信号, 以及滤波器自身特征都定义在时间的离散时刻。在任何情况下, 连续时间信号总可由均匀时间间隔观测信号定义的一系列样值来表示。在从连续时间信号到离散时间信号的变换过程中并不会发生信息丢失, 只需要满足众所周知的取样定理。该定理表明, 取样率必须高于两倍的连续时间信号最高频率。因此 我们用序列  $u(n) (n=0, \pm 1, \pm 2, \dots)$  来表示离散时间信号  $u(t)$ 。为方便起见 这里把取样周期归一化为 1。这一约定将贯穿全书。

## 自适应滤波器

维纳滤波器的设计要求所要处理的数据统计方面的先验知识。只有当输入数据的统计特性与滤波器设计所依赖的某一先验知识匹配时, 该滤波器才是最优的。当这个信息完全未知时, 就不可能设计维纳滤波器, 或者该设计不再是最优的。在这种情况下, 可采用的一个直接方法是“估计和插入”过程。该过程包含两个步骤 首先是“估计”有关信号的统计参数 然后将所得到的结果“插入 (plug into)”非递归公式以计算滤波器参数。对于实时运算 该过程的缺点是要求特别精心制作, 而且要求价格昂贵的硬件。为了消除这个限制, 可采用自适应滤波器 (adaptive filter)。采用这样一种系统, 意味着滤波器是自设计的, 因为自适应滤波器依靠递归算法进行其运算, 这样使得它有可能在有关信号特征的完整知识不能得到的环境下, 完满地完成滤波运算。该算法将从某些预先确定的初始条件集出发, 这些初始条件代表了人们所知道的上述环境的任何一种情况。我们还发现, 在平稳环境下, 该算法经一些成功迭代后收敛于某

种统计意义上的最优维纳解。在非平稳环境下，该算法提供了一种跟踪能力，因为它能够跟踪输入数据统计特性随时间的变化，只要这种变化是足够缓慢的。

作为递归算法应用的一个直接结果，自适应滤波器的参数将借此从一次迭代到另一迭代进行更新，滤波器参数变成与数据相关。因此，这意味着自适应滤波器实际上是一个非线性系统。从这个意义上讲，它不遵循叠加原理。尽管有这个特性，自适应滤波器通常还是分为线性和非线性两种。如果输入-输出映射遵循叠加原理，则认为这个自适应滤波器是线性的；否则，认为该滤波器是非线性的。

在线性自适应滤波器文献中，已经研究了大量的递归算法。在下面的分析中，自适应算法的选择取决于如下一个或多个因素：

- 收敛速率 它定义为算法在响应平稳输入时足够接近地收敛于均方误差意义上最优维纳解所需要的迭代数。快速收敛允许算法快速自适应于统计意义上未知的平稳环境。
- 失调 (misadjustment) 对于一个感兴趣的算法，这个参数提供了自适应滤波器集平均的最终均方误差与维纳滤波器所产生的最小均方误差之间偏离程度的一个定量测量。
- 跟踪 (tracking) 当一个自适应滤波算法运行在非平稳环境时，该算法需要跟踪环境的统计量变化。然而，算法的跟踪性能受到两个相互矛盾的特性的影响：(a) 收敛速率；(b) 由算法噪声引起的稳态波动。
- 鲁棒性 (robustness) 对于一个鲁棒的自适应滤波器，小的扰动 (disturbance) 只会产生小的估计误差。这些扰动来源于各种因素，包括来自滤波器内部或外部的因素。
- 计算要求 这里关心的问题包括：(a) 完成算法的一次完整迭代所需要的运算量（即乘法、除法、加法和减法）；(b) 存储数据和程序所需要的存储器位置的大小；(c) 在计算机上对算法编程所需要的投资。
- 结构 涉及算法的信息流结构以及硬件实现的方式。例如，其结构呈现高度模块化、并行或并发的算法很适合于使用超大规模集成电路 (VLSI, very large-scale integration)<sup>①</sup> 实现。
- 数值特性 当一个算法数值实现时，将产生由量化误差所引起的不精确性，该量化误差依次由输入数据的模/数变换和内部计算的数字表示所产生。一般来说，造成严重设计问题的是后者的量化误差源。特别地，存在人们所关心的两种基本问题：数值稳定性和数值精确性问题。数值稳定性是自适应滤波算法固有的特征。另一方面，数值精确性由表示数据样值和滤波器系数的位数（即二进制数字）确定。当某种算法对其数字实现的字长变化不敏感时，就说该自适应滤波算法是数值鲁棒的。

这些因素以其特有的方式也出现在非线性自适应滤波器设计中，除非在维纳滤波器形式中不再有一个明确规定的参考格式。反之，值得一提的是，非线性滤波算法可能收敛于误差性能曲面某一局部极小点，或者可望收敛于它的某一全局极小点。

除了第 17 章和后记外，本书只关心线性自适应滤波器及其有关课题。

<sup>①</sup> VLSI 技术更适合于高度模块化、并行或并发算法的实现。当一个结构由类似的基本节级级联组成时就认为这个结构是模块化的。并行，意指大量运算并列进行的性能。并发，意指在相同时间内大量类似运算的性能。关于自适应滤波器 VLSI 实现的讨论见 Shanbhag 和 Parhi(1994) 的书。该书强调使用流水线这是用来增加自适应滤波算法吞吐量的一种体系结构技术。

## 线性滤波器结构

线性滤波算法的运行涉及两个基本过程：(1) 滤波过程，用来对一系列输入数据产生输出响应；(2) 自适应过程，其目的是提供滤波过程中可调参数自适应控制的一个机制（算法）。这两个过程相互影响地工作。自然，滤波过程结构的选择总体上对算法的运行具有深刻的影响。

线性滤波器的脉冲响应决定了滤波器的记忆能力。另一方面，可把线性滤波器分为有限脉冲响应 (FIR, finite-duration impulse response) 和无限脉冲响应 (IIR, infinite-duration impulse response) 滤波器，分别由有限记忆和无限长但衰落记忆来表征。

### 具有有限记忆的线性滤波器

在有限记忆自适应滤波器中，存在三种类型很容易区分的滤波器结构，如下所述。

#### 横向滤波器

横向滤波器 (transversal filter)<sup>①</sup>，也称为抽头延迟线滤波器或有限脉冲响应滤波器，它由图 4 所示的三个基本单元组成：(a) 单位延迟单元；(b) 乘法器；(c) 加法器。用在滤波器中的延迟单元的个数确定了脉冲响应的有限持续时间。延迟单元个数（如图中  $M$  所示）通常称为滤波器的阶数。在该图中，每个延迟单元用单位延迟算子  $z^{-1}$  表示。特别地，当对  $u(n)$  进行  $z^{-1}$  运算时，其结果输出为  $u(n-1)$ 。滤波器中每个乘法器的作用是用滤波器系数 [ 也称为抽头权值 (tap weight) ] 乘以其相连接的抽头输入。于是，连到第  $k$  个抽头输入  $u(n-k)$  的乘法器产生  $w_k^* u(n-k)$  的输出，其中  $w_k$  是抽头权值， $k=0, 1, \dots, M$ ，星号表示复数共轭。这里假设抽头输入和抽头权值都是复数。滤波器中加法器的合并作用是对各个乘法器输出求和，并产生总的滤波器输出。对于所示的横向滤波器，其输出为

$$y(n) = \sum_{k=0}^M w_k^* u(n-k) \quad (3)$$

上式叫做有限卷积和，因为它将滤波器的有限脉冲响应  $w_k^*$  与滤波器输入  $u(n)$  卷积以便产生滤波器的输出  $y(n)$ 。

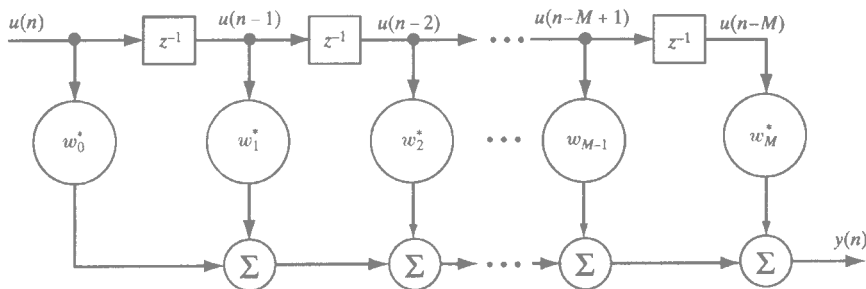


图 4 横向滤波器

① Kallmann(1940) 首先把横向滤波器描述为连续时间系统，其输出为非离散 (nondispersive) 延迟线的均匀间隔抽头上电压的线性组合构成。横向滤波器典型地利用数字电路、电荷耦合器件或声表面波器件来实现。由于通用和易于实现，横向滤波器作为一种基本信号处理结构，获得广泛应用。

## 格型预测器

格型预测器 (lattice predictor)<sup>①</sup> 具有模块结构。在这种模块结构中，它由大量独立的级组成，每一级具有格型的外表，从而取名“格型”作为构造的表征。图 5 表示由  $M$  级组成的格型预测器， $M$  称为预测器阶数。图中示出的第  $m$  级格型预测器由下列一对输入输出关系（假设使用复数值、广义平稳输入数据）描述

$$f_m(n) = f_{m-1}(n) + \kappa_m^* b_{m-1}(n-1) \quad (4)$$

和

$$b_m(n) = b_{m-1}(n-1) + \kappa_m f_{m-1}(n) \quad (5)$$

式中  $m=1, 2, \dots, M$ ， $M$  为最终的预测器阶数。 $f_m(n)$  是第  $m$  级前向预测误差， $b_m(n)$  是第  $m$  级后向预测误差。 $\kappa_m$  叫做  $m$  阶发射系数。前向预测误差  $f_m(n)$  定义为输入  $u(n)$  与基于一组  $m$  个过去输入  $u(n-1), \dots, u(n-m)$  所做出的预测值之差。相应地，后向预测误差  $b_m(n)$  定义为输入  $u(n-m)$  与一组基于  $m$  个未来输入  $u(n), \dots, u(n-m+1)$  所做出的后向预测值之差。考虑图中第 1 级输入的条件，我们有

$$f_0(n) = b_0(n) = u(n) \quad (6)$$

式中  $u(n)$  为  $n$  时刻格型预测器输入。于是从式 (6) 的初始条件出发并给定一组发射系数  $\kappa_1, \kappa_2, \dots, \kappa_M$ ，则可经由格型预测器一级一级地向前推进，最终确定出一对输出  $f_M(n)$  和  $b_M(n)$ 。

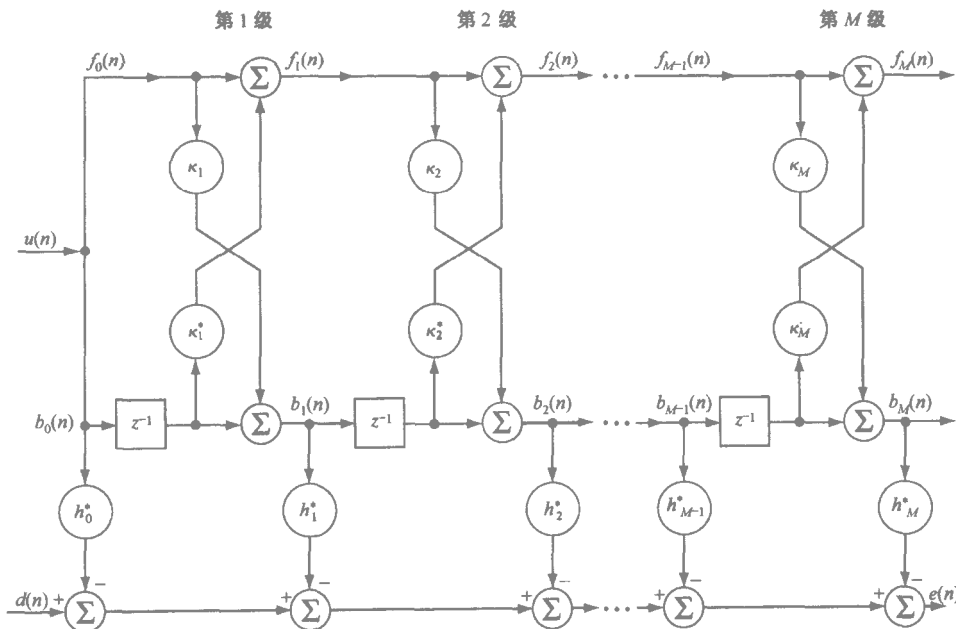


图 5 多级格型滤波器

对于由统计过程获得的相关输入序列  $u(n), u(n-1), \dots, u(n-M)$  后向预测误差  $b_0(n), b_1(n), \dots, b_M(n)$  组成一个非相关随机变量序列。再则，两个随机变量序列之间在如

下意义上存在着——对应关系：如果给定其中的一个，可以惟一地确定另一个，反之亦然。因此后向预测误差  $b_0(n), b_1(n), \dots, b_M(n)$  的线性组合可用于提供期望响应序列  $d(n)$  的估计，如图 5 下半部所示。 $d(n)$  与其估计值之间的差表示估计误差  $e(n)$ 。这里描述的过程称为联合过程估计。当然，可以直接应用原输入序列  $u(n), u(n-1), \dots, u(n-M)$  产生期望响应的估计。然而，图中所述的间接法具有简化抽头权值  $h_0, h_1, \dots, h_M$  计算的优点，因为它利用了估计中所采用的后向预测误差的非相关特性。

### 脉动阵列

脉动阵列 (systolic array)<sup>①</sup> 表示并行计算网络，它非常适合于映射大量重要的代数计算，诸如矩阵乘法、三角化和反向替代。脉动阵列的基本处理单元有两种：边界单元和内部单元，其功能分别画在图 6(a) 和图 6(b) 中。在每一种情况中，参数  $r$  表示存储在单元中的一个数值。边界单元的作用是产生等于  $u/r$ ，即输入  $u$  除以存储在单元中的数  $r$  的输出。内部单元的作用有两方面：(a) 用存储在单元中的  $r$  乘以输入  $s$  (来自顶部)，再从第二个输入 (来自左边) 中减去乘积  $rs$ ，从而产生差  $u - rs$  作为该单位右边的输出；(b) 不用交替地向下发送第一个输入  $s$ 。

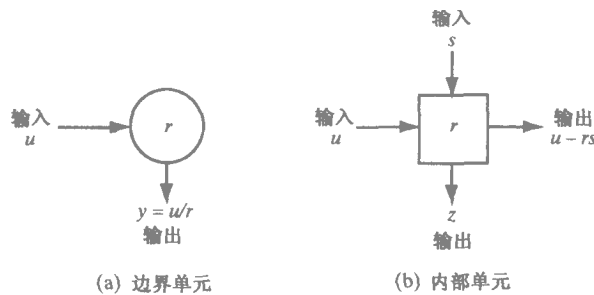


图 6 脉动阵列的两种基本单元

例如，考虑图 7 所示的  $3 \times 3$  三角波前阵列。该阵列涉及边界单元和内部单元的组合。在这种情况下，三角阵列计算与输入向量  $\mathbf{u}$  有关的输出向量  $\mathbf{y}$

$$\mathbf{y} = \mathbf{R}^{-T} \mathbf{u} \quad (7)$$

式中  $\mathbf{R}^{-T}$  是转置矩阵  $\mathbf{R}^T$  的逆矩阵。 $\mathbf{R}^T$  的元素是三角阵列中各自单元的内容。零加到图中阵列的输入中是为了提供式 (7) 流水线计算所需要的延迟。

如前所述，脉动阵列结构提供了模块化、局部互连、高度流水线和同步并行处理，其中同步借助于全球时钟获得。

### 具有无限记忆的线性滤波器

注意图 4 的横向滤波器、图 5 的格型预测器联合过程估计和图 7 的三角脉动阵列都具有共同的特性：它们都是由有限脉冲响应表征的。换句话说，它们都是 FIR 滤波器的一些例子，其结构只含前馈支路。另一方面，图 8 所示的结构是 IIR 滤波器的一个例子。IIR 滤波器不同于 FIR 滤波器的特征是它包含反馈支路。反馈的存在使得 IIR 滤波器的脉冲响应持续时间变

脉动阵列最早由 Kung 和 Leiserson (1978) 提出。在特殊情况下，使用脉动阵列有可能产生较大的吞吐量，这在实时操作中需要运用大量高级信号处理算法时是需要的。

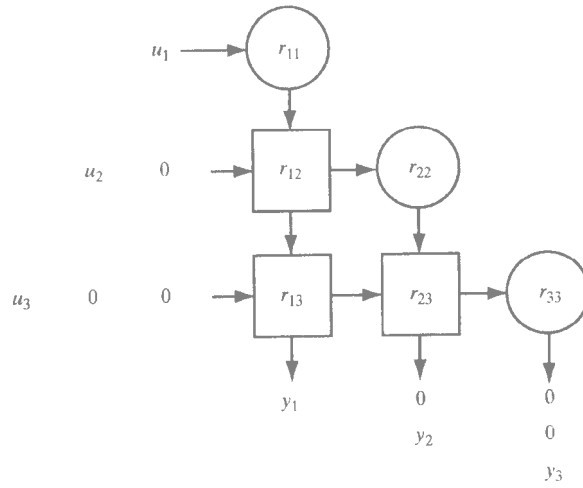


图 7 三角脉动阵列

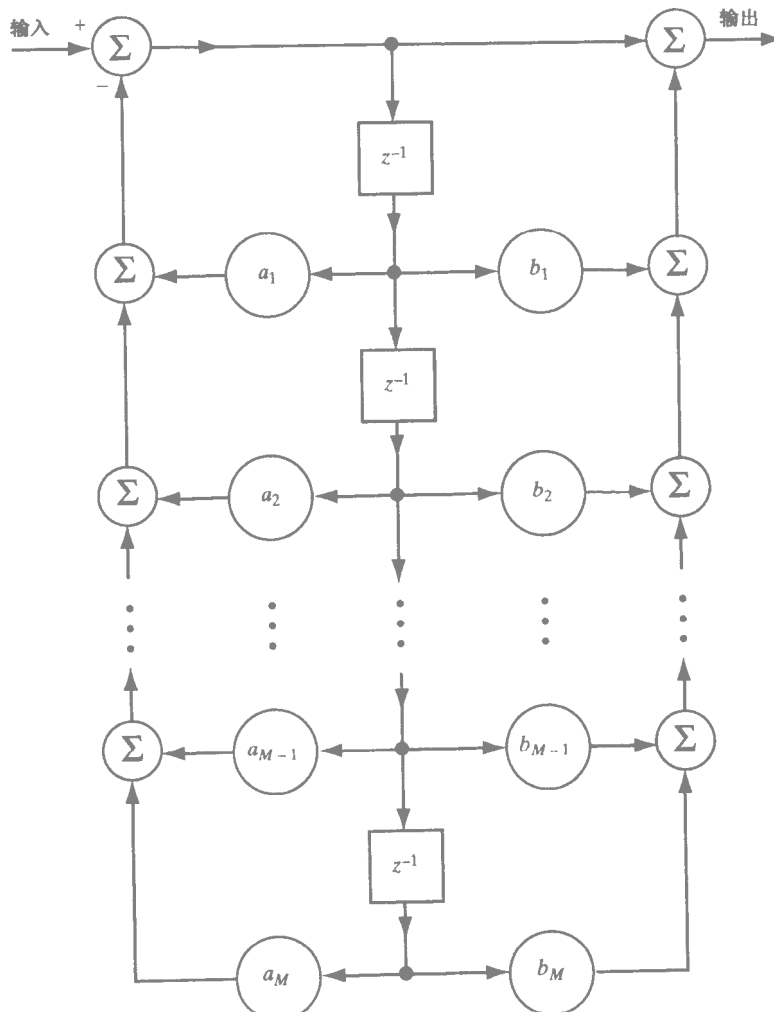


图 8 IIR 滤波器 (假定实值数据)

为无限长。此外，反馈的存在引入新的问题：潜在的不稳定。特别地，IIR 滤波器有可能变成不稳定（即发生振荡），除非在选择反馈系数时采取特殊的预防措施。相反地，FIR 滤波器总是稳定的。这就解释了为什么流行使用 FIR 滤波器，并以一种形式或另一种形式作为设计线性自适应滤波器的结构基础。

构造 IIR 自适应滤波器并以很简单的方式满足稳定性要求的一种新颖方法是使用横向 Laguerre 滤波器。图 9 示出这种滤波器的结构，其中使用 Laguerre 函数代替单位延迟。特别地，我们有

$$L_0(z) = \frac{\sqrt{1-a^2}}{1-az^{-1}} \tag{8}$$

和

$$L_1(z) = \frac{z^{-1}-a}{1-az^{-1}} \tag{9}$$

$L_0(z)$  和  $L_1(z)$  由  $a$  参数化，为了保证该滤波器的稳定性，参数  $a$  的选择受到如下约束

$$0 < |a| < 1$$

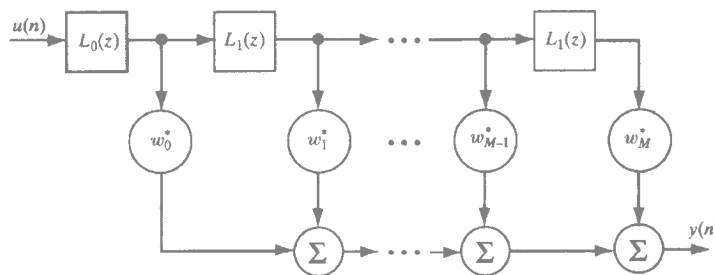


图 9 Laguerre 横向滤波器

图 10(a)和图 10(b)分别示出  $a$  取正值和负值时  $L_0(z)$  的零极点模式。类似地，图 11(a)和图 11(b)分别示出  $a$  取正值和负值时  $L_1(z)$  的零-极点模式。由图可见：

- 无论  $L_0(z)$  还是  $L_1(z)$  都是稳定的滤波器，因为它们惟一的极点  $z = a$  在  $z$  平面单位圆内。
- 零阶 Laguerre 滤波器  $L_0(z)$  是一个一阶低通滤波器（当  $a$  为正值时）或一阶高通滤波器（当  $a$  为负值时）。
- 一阶 Laguerre 滤波器  $L_1(z)$  是一个一阶全通滤波器。

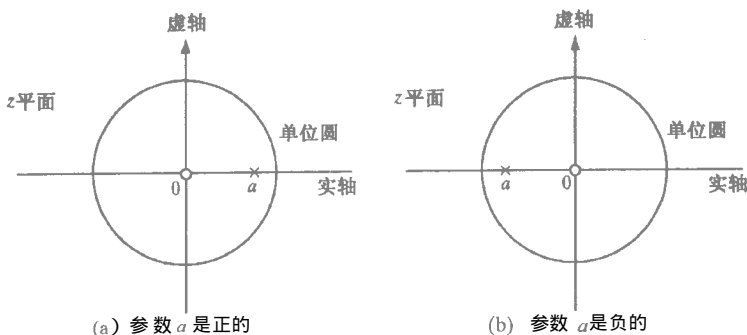


图 10  $L_0(z)$  的零-极点模式

图 10 和图 11 的参数  $a$  假设是实数；如果需要的话，也可以假设是复数值。

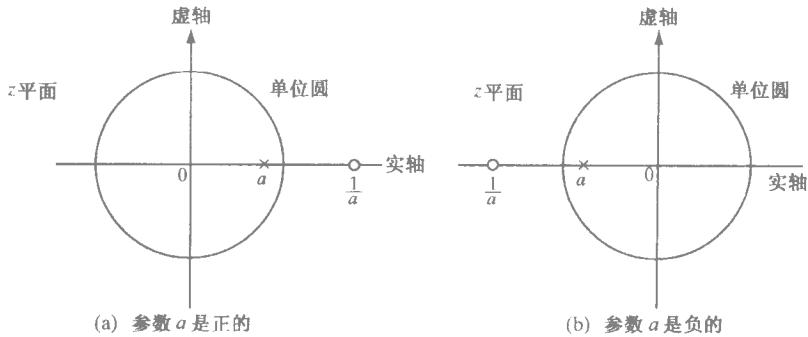


图 11  $L_1(z)$  的零极点模式

## 线性自适应滤波器的研究方法

对线性自适应滤波问题不存在惟一的解决方案。相反地，我们有一套表示各种递归算法的工具，其中的每一个提供了它所拥有的期望特性。自适应滤波的用户所面临的挑战首先是了解各种自适应滤波算法的能力和限制，其次是把了解到的知识用于选择合适的算法以满足目前的应用需要。

基本上，可认为存在两种不同的推导线性自适应滤波器递归算法的方法。

### 随机梯度法

随机梯度法 (stochastic gradient approach) 使用抽头延迟线或者横向滤波器作为实现线性自适应滤波器的构造基础。对于平稳输入的情况，代价函数<sup>①</sup> (也称为性能指标 定义为均方误差 即期望响应与横向滤波器输出之差的均方值)。代价函数恰好是横向滤波器中抽头权值的二次函数。未知抽头权值均方误差的函数关系看做是具有惟一确定的极小点的多维抛物面。如前所述，我们把这个抛物面称为误差性能曲面；对应于该曲面极小点的抽头权值定义了最优维纳解。

为了推导更新自适应横向滤波器抽头权值的递归算法，我们分两步进行这项工作。首先，使用迭代方法求解维纳-霍夫方程 (Wiener-Hopf equation)；迭代过程以最优化理论中人们所熟知的最速下降法 (method of steepest descent) 为基础。这个方法需要使用梯度向量，其值取决于两个参量：横向滤波器中的抽头输入的自相关矩阵以及期望响应与该抽头输入之间的互相关向量。其次，我们使用这些相关的瞬态值，以便推导梯度向量的估计值，推导中假设该向量是随机的。基于上述思想的算法，通常称为最小均方 (LMS, least-mean-square) 算法。当横向滤波器运行在实数据的情况下，该算法大体上可描述为

$$(\text{抽头权向量更新值}) = (\text{老的抽头权向量值}) + (\text{学习速率参数}) \times (\text{抽头输入向量}) \times (\text{误差信号})$$

其中误差信号定义为期望向量与抽头输入向量所产生的横向滤波器的实际向量之差。

<sup>①</sup> 在函数的一般定义中 我们经常说到从向量空间到实数/复数 标量的变换 (Luenberger, 1969; Dorny, 1975)。代价函数提供了分配性能指标的定量度量，从而限制它为标量。

LMS 算法很简单，而且在正确条件下可能获得满意的性能。其主要缺点是收敛速率相当缓慢，而且对抽头输入相关矩阵条件数（矩阵的条件数定义为其最大特征值与最小特征值之比）的变化比较敏感。即使这样，LMS 算法仍然十分流行且应用广泛。

在非平稳环境下，误差性能曲面的方向随时间连续变化。在这种情况下，LMS 算法有一个连续跟踪误差性能曲面极小点的附加任务。实际上，只要输入数据与 LMS 算法学习速率相比是慢的，就会发生跟踪问题。

随机梯度方法也用于格型结构。产生的自适应滤波算法叫做梯度自适应格型 (GAL, gradient adaptive lattice) 算法。LMS 和 GAL 算法是线性自适应滤波器随机梯度族的两个成员，而迄今为止，LMS 仍然是这个家族中最流行的一员。

## 最小二乘估计

第二种线性自适应滤波算法方法以最小二乘法为基础。根据这个方法，我们对加权误差平方和形式的代价函数进行最小化，其中误差或残差定义为期望响应与实际滤波器输出之差。最小二乘法可用块估计或递归估计来表示。在块估计中，输入数据流以等长度块的形式排列，而且一块一块地对输入数据进行滤波处理；而在递归估计中，一个样值一个样值地对感兴趣的估计（例如，横向滤波器的抽头权值）进行更新。通常，递归估计比块估计要求较少的存储量，这就是为什么在实际中递归估计的使用要多得多的原因。

递归最小二乘 (RLS, Recursive least-squares) 可看做卡尔曼滤波的一个特例。卡尔曼滤波器著名的特点是状态概念，它提供了特定时刻加到滤波器抽头所有输入的一个度量。于是，在卡尔曼滤波算法的核心部分，存在一种递推关系，它可用文字表述为

$$(\text{状态递推值}) = (\text{旧的状态值}) + (\text{卡尔曼增益}) \times (\text{新息向量})$$

其中新息向量表示在计算时刻进入滤波过程的新的信息。现在，我们有足够的把握说，在卡尔曼变量与 RLS 变量之间的确存在一一对应关系。这个对应关系使得我们可以从大量卡尔曼滤波器文献中选择一些方法用来设计基于最小二乘估计的线性自适应滤波器。此外，取决于所采用的方法，递归最小二乘线性自适应滤波算法族可分为三类：

- a) 标准 RLS 算法 它假设使用横向滤波器作为线性自适应滤波器的结构基础。标准 RLS 算法的推导依赖于线性代数中的一个基本结果——矩阵求逆引理。更重要的是，该算法既有相同于标准卡尔曼滤波算法的优点，也受到相同的限制，如缺乏数值鲁棒性和计算复杂性过高。正是这两个限制促使人们研究开发另两类 RLS 算法。
- b) 平方根 RLS 算法 它基于输入数据矩阵的 QR 分解。完成这个分解的两种众所周知的技术是豪斯霍尔德 (Householder) 变换和吉文斯 (Givens) 旋转。此二者都是数据自适应的变换。就所讨论的问题而言，基于豪斯霍尔德变换或吉文斯旋转的 RLS 算法是数值稳定的和鲁棒的。由此产生的线性自适应滤波器叫做平方根自适应滤波器，因为在矩阵意义上它们表示了标准 RLS 算法的平方根形式。
- c) 快速 RLS 算法 标准 RLS 算法和平方根 RLS 算法都具有随  $M^2$  增加的计算复杂性，其中  $M$  是算法中可调权值的数目（即自由度），这种算法通常称为  $O(M^2)$  算法，其中  $O(x)$  代表  $x$  的数量级。相反，LMS 是  $O(M)$  算法，其中的计算复杂性随  $M$  线性增加。

当  $M$  很大时，从硬件实现的观点看， $O(M^2)$  算法的计算复杂性将会变得不可接受。因此存在一个强烈的动因，以假设计算复杂性为  $O(M)$  形式的方式修正 RLS 算法的计算公式。在瞬态处理的情况下，这个目标的确可达到。首先利用输入数据矩阵的托伯利兹 (Toeplitz) 结构所固有的冗余度这个优点，其次通过利用前向和后向两个方向线性最小平方预测中的冗余度。这样产生的算法都称为快速 RLS 算法。这些算法组合了具有  $O(M)$  计算复杂性的递归线性最小二乘估计的期望特性。取决于所采用的滤波结构，一般存在两种形式的快速 RLS 算法：

- 阶更新自适应滤波器 它以类似于线性前向和后向预测中采用的格型结构为基础。
- 快速横向滤波器 在该结构中，线性前向和后向预测用分离的横向滤波器来完成。

某些（但不是全部）阶更新自适应滤波器的实现已经证明是数值稳定的，而快速横向滤波器将会遇到数值稳定性问题，故其实际应用受到了限制。

线性自适应滤波器介绍性的初步讨论将是不完整的，更不用说关于其跟踪特性的某些内容。在这个范围内我们注意到诸如 LMS 算法等随机梯度算法是模型无关的。一般来说，人们期望它呈现好的跟踪特性，实际上它也做到了这一点。相反，RLS 算法是模型相关的，这意味着其跟踪性能可能比大量的随机梯度算法族来得差，除非人们注意使作为算法基础的数学模型与产生输入数据的基础物理过程之间的失配最小化。

## 如何选择自适应滤波器

给出对系统设计者可用的各种自适应滤波器，如何选择自适应滤波器以解决感兴趣的应用问题呢？很清楚，无论何种选择，都必须是费用有效的。牢记这个目标，我们可以看清要注意的三个重要问题：计算费用、性能和鲁棒性。计算机模拟提供了详细研究这些问题的良好的第一步。可以从使用 LMS 算法出发，并把它作为自适应滤波工具以便满足研究需要。LMS 算法的实现相对简单，而且足以强有力地评价现有问题各种自适应应用的实际效益。再则，它还可作为评价进一步性能改进提供实际的参考框架，这种改进可通过使用更复杂的自适应滤波算法来获得。最后，该研究还必须包括对现实世界数据的测试，而且该测试是不可替代的。

自适应滤波的实际应用是千变万化的，而且每种应用具有各自的特殊性。所以，对于一种应用的解决方案不一定适用于另外一种应用。不过，要获得成功，必须从物理上理解与感兴趣实际应用有关的滤波器运行环境。

## 自适应滤波器的实数和复数形式

在研究自适应滤波算法时，不管其来源如何，习惯上都假设输入数据是基带形式的。“基带”用来指明由信息源产生的源消息信号所表示的频带。

在诸如通信、雷达、声纳等应用中，接收机输入的承载信息的信号典型地由调制在载波上的消息信号组成。消息信号的带宽通常小于载波频率。这意味着调制信号是窄带信号。为了获得窄带信号的基带表示，信号以完全消除载波影响的方式进行向下频谱搬移，在搬移过程中完全保护包含在消息信号中的信息。一般来说，这样得到的基带信号是复数的。换句话说，信号样值可以写为

$$u(n) = u_I(n) + ju_Q(n) \quad (10)$$

式中  $u_I(n)$  是同相 (in-phase) (实数) 分量,  $u_Q(n)$  是正交 (虚数) 分量。等价地, 我们有

$$u(n) = |u(n)|e^{j\phi(n)} \quad (11)$$

式中  $|u(n)|$  和  $\phi(n)$  分别是信号的幅度和相角。

因此, 在本书后续章节中研究的自适应滤波器, 无论是线性还是非线性滤波器, 理论都假设使用复数信号。这样研究的自适应滤波算法被说成是复数形式的。复数自适应滤波器的重要优点在于, 它保护了后面提到的应用领域中所遇到的复数信号的数学表达式和完美的结构。

如果要处理的信号是实数的, 我们自然使用感兴趣的实数形式自适应滤波算法。给出自适应滤波器的复数形式, 将其变为相应的实数形式将是很简单的。特别地, 我们做如下两件事情:

- 每当复数共轭运算发生在算法中时, 移走该复数共轭运算。
- 每当算法中发生埃尔米特 (Hermitian) 转置 (即共轭转置) 时, 用原来的转置代替埃尔米特转置。

做这样的简单处理后, 复数自适应滤波器就包括实数自适应滤波器作为特例。

## 自适应波束形成

至今讨论的自适应滤波方法和结构都是属于时间类别的。在这类滤波器中, 滤波运算在时域进行。当然, 自适应滤波也可以是空间类别的, 在这种类别中, 独立传感器阵列代替了空间的各点, 以便“听到”来自某一距离的信号。实际上, 传感器提供了在空间对接收信号取样的一种手段。在某一特殊时刻收集的传感器输出集构成信源的一个“快照”。当传感器均匀地位于某一直线上的时候, 空间滤波器中的数据快照起到了类似于某一特定时刻横向滤波器中的一组连续抽头输入的作用<sup>①</sup>。

阵列信号处理的重要应用包括如下方面:

- 雷达 这里的传感器由对入射电磁波做出响应的天线元素组成, 而且要求能对辐射电磁波做出响应的信号源做出检测, 估计波到达角, 并提取有关信号源的信息。
- 声纳 这里的传感器由用来对入射声波做出响应的水中听音器组成, 而且要求其信号处理类似于雷达中的信号处理。
- 语音增强 这里的传感器由麦克风组成, 而且要求在有背景噪声的情况下能听到期望说话者的声音。

在这些应用中, 所谓波束形成 (beamforming) 的目的在于区分信号和噪声的空间特性。用来做波束形成的系统叫做波束形成器 (beamformer)。“波束形成器”这个术语来自这样一个事实: 早期的天线被设计成射束的形式, 以便接收从特定方向发射的信源信号并衰减不感兴趣的其他方向的信号 (Van Veen & Buckley, 1988)。波束形成应用于能量的发射和接收。

图 12 示出使用相同传感器的线性阵列的自适应波束形成器框图。传感器输出 (假设是基

<sup>①</sup> 时间域与空间域信号处理之间相似性的讨论, 见 Bracewell(1986)、Van Veen 和 Buckley(1988) 的论文。

带形式) 信号各自加权并求和, 以便产生波束形成器的总输出。波束形成器必须满足两个要求:

- 转向 steering 能力 借此 目标 源 信号总是受保护。
- 干扰消除 这样, 使得输出信噪比最大。

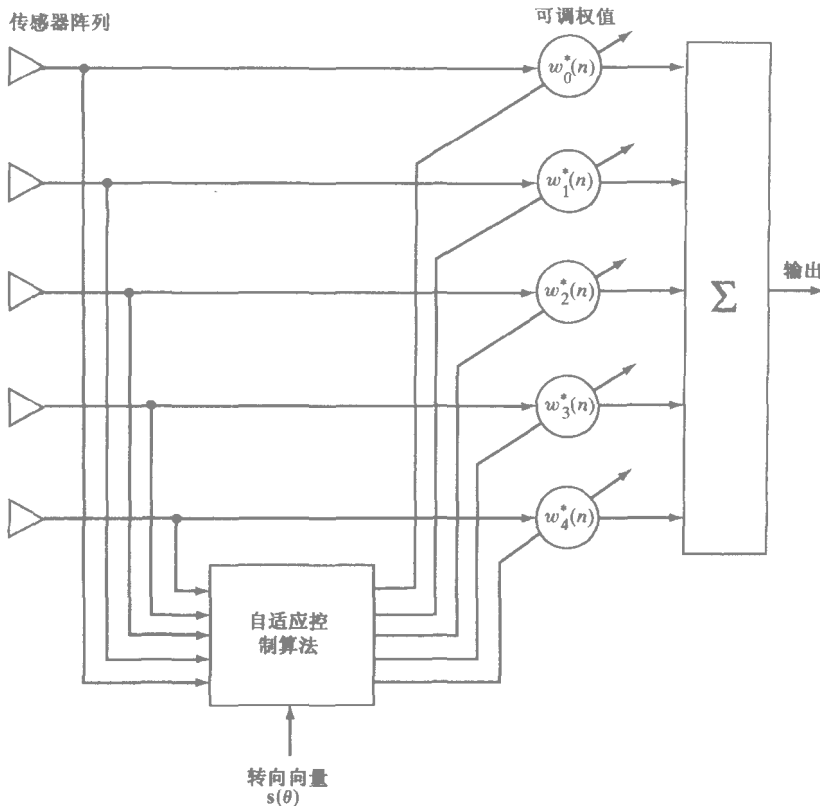


图 12 五个传感器阵列的自适应波束形成器

满足这些要求的一种方法是 自适应过程中 在  $M \times 1$  权向量  $\mathbf{w}(n)$  满足如下条件

$$\mathbf{w}^H(n)\mathbf{s}(\theta) = 1 \quad \text{对于所有 } n \text{ 和 } \theta = \theta_i \quad (12)$$

的约束下 使波束形成器输出的方差 平均功率 极小化。式中  $\mathbf{s}(\theta)$  是  $M \times 1$  转向向量 上标  $H$  表示埃尔米特转置。假设基带信号是复数值, 因而需要复数共轭运算。角度  $\theta = \theta_i$  的值由目标的方向来确定。角度  $\theta$  以传感器 0 作为参考点进行自我测量。

转向向量对  $\theta$  的依赖性用如下关系来定义

$$\mathbf{s}(\theta) = [1, e^{-j\theta}, \dots, e^{-j(M-1)\theta}]^T \quad (13)$$

令  $\phi$  表示平面波实际入射角, 它是由相对于线性阵列的法线测得的。则根据图 13 容易看出

$$\theta = \frac{2\pi d}{\lambda} \sin \phi, \quad -\pi/2 \leq \theta \leq \pi/2 \quad (14)$$

式中  $d$  是该阵列相邻传感器之间的间隔,  $\lambda$  是入射波的波长。由于  $\phi$  限制在  $[-\pi/2, \pi/2]$  范围内而  $\theta$  的允许值在  $[-\pi, \pi]$  范围内 故由式 (12) 可见  $d$  必须小于  $\lambda/2$  以使得  $\phi$  与  $\theta$  值之间存

在一一对应关系。于是要求  $d < \lambda/2$  可看做取样定理的空间域表示。如果这个要求不满足，则波束形成器的辐射（天线）模式将呈现格栅型波瓣（grating lobes）（Skolnik, 1980）。该辐射模式就是波束形成器输出功率作为其测量方向函数的一个关系图。

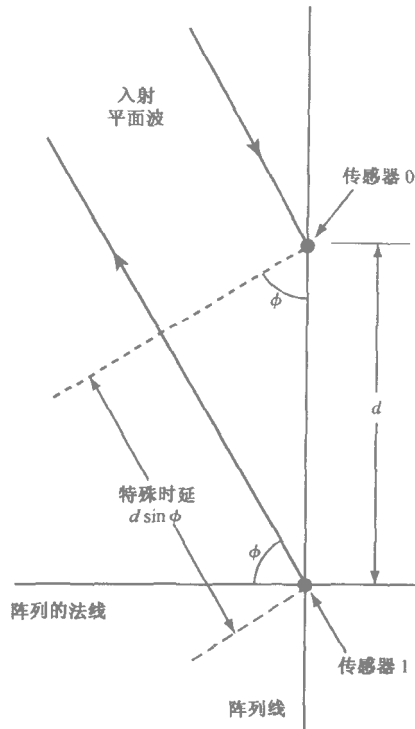


图 13 当平面波入射到线性阵列上时产生的空间时延

对式 (12) 所强加的信号保护约束保证：对于规定的观察方向  $\theta = \theta_0$ ，不管权向量  $\mathbf{w}$  的元素取什么值，阵列响应将维持不变（等于 1）。在这个约束下，使波束形成器输出方差最小化的算法很自然地称为最小方差失真响应 MVDR, minimum-variance distortionless response 波束形成算法。注意，强加的信号保护约束使可用自由度变为  $M - 2$ ，其中  $M$  是阵列中传感器的数目。因此，由 MVDR 产生的独立空值数（即能够消除独立干扰的数目）为  $M - 2$ 。

## 自适应滤波器的四种应用

自适应滤波器具有在未知环境下良好运行并跟踪输入统计量随时间变化的能力，使得自适应滤波器成为信号处理和自动控制应用领域强大的设备。实际上，自适应滤波器已经成功地应用于通信、雷达、声纳、地震学和生物医学工程等领域。尽管这些应用在特性方面是千变万化的，但它们都有一个共同的基本特征：输入向量和期望响应被用来计算估计误差，该误差依次用来控制一组可调滤波器系数。取决于所采用的滤波器结构，可调系数可取抽头权值、反射系数或旋转参数等形式。然而，自适应滤波器各种应用之间的本质不同在于其期望响应的提取方式不同。就此，可把自适应滤波器应用分为四种类型，如图 14 所示。为方便起见，图中使用如下记号：

$u$  = 加到自适应滤波器的输入

$y$  = 自适应滤波器的输出

$d$  = 期望响应

$e = d - y$  = 估计误差

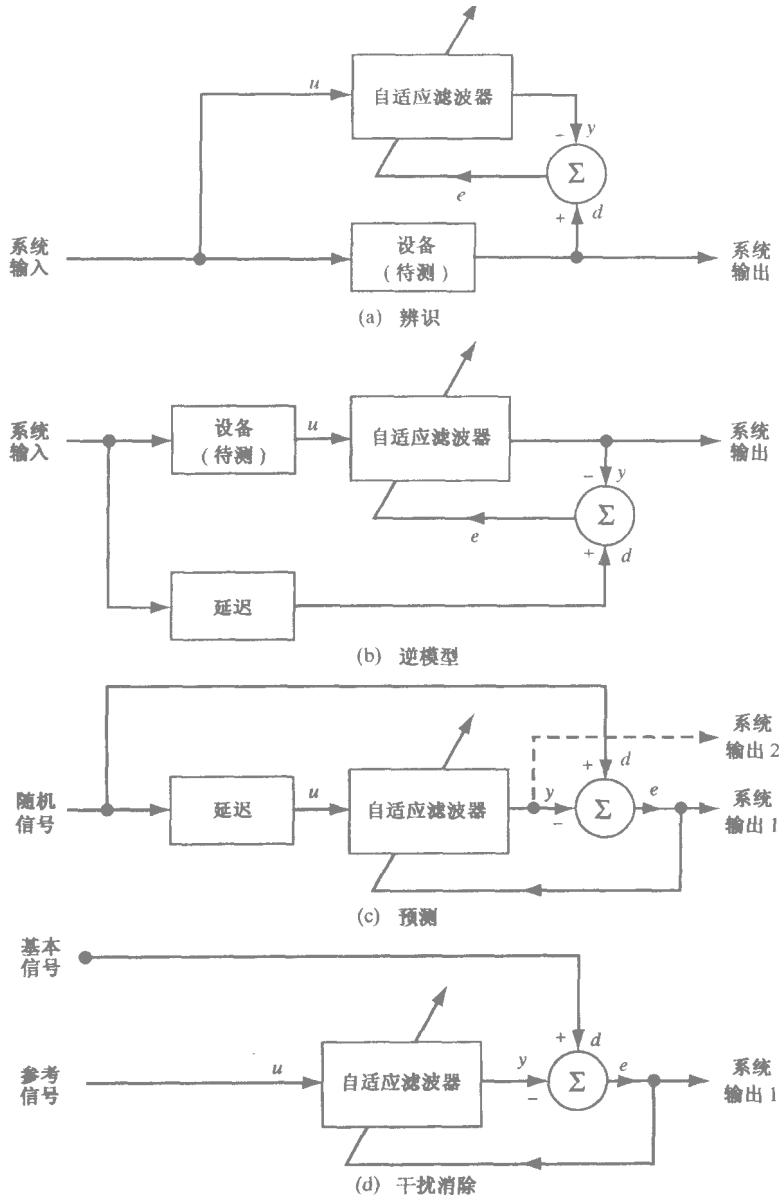


图 14 自适应滤波应用的四种基本类型

自适应滤波器应用的四种基本类型的作用描述如下：

- 1) 辨识 [图 14(a)] 数学模型概念是科学与工程的基础。在这类涉及辨识的应用中，自适应滤波器用来提供一个在某种意义上能够最好拟合未知装置的线性模型。该装置和自适应滤波器由相同的输入激励。该装置的输出提供作为自适应滤波器的期望响

应。如果该装置具有动态特性，则自适应滤波器所提供的模型将是时变的。

- 2) 逆模型 [图 14(b)] 在第二类应用中，自适应滤波器的作用是提供一个逆模型，该模型可在某种意义上最好拟合未知噪声装置。理想地，在线性系统的情况下，该逆模型具有等于未知装置转移函数倒数的转移函数，使得二者的组合构成一个理想的传输媒介。该装置（系统）输入的延迟构成自适应滤波器的期望响应。在某些应用中，该装置（系统）输入不加延迟地用做期望响应。
- 3) 预测 [图 14(c)] 这里，自适应滤波器的作用是对随机信号的当前值提供某种意义上最好的预测。于是，信号的当前值用做自适应滤波器的期望响应。信号的过去值加到滤波器的输入端。取决于感兴趣的应用，自适应滤波器的输出或估计、预测误差均可作为系统的输出。在第一种情况下，系统作为一个预测器；而在后一种情况下，系统作为预测误差滤波器。
- 4) 干扰消除 [图 14(d)] 在最后一类应用中，自适应滤波器以某种意义上的最优化方式消除包含在基本信号中的未知干扰（类似于一个承载信息的信号分量）。基本信号用做自适应滤波器的期望响应，参考信号用做滤波器的输入。参考信号来自定位的某一传感器或一组传感器，并以承载信息的信号是微弱的或基本不可预测的方式，供给基本信号上。

表 1 列出某些特定的应用以便说明自适应滤波器应用的四种基本类型。所列出的应用来自控制系统、地震学、心电图学、通信和雷达等领域。其用途列在表中最后一列。

表 1 自适应滤波器的应用

自适应滤波的类型	应用	用途
辨识	系统辨识	给定一个未知的动态系统，系统辨识的目的是设计一个自适应滤波器以逼近该动态系统
	分层地球建模	在地震探测中，开发研究地球的分层模型以解释地球表面的复杂问题
逆模型	均衡	给定一个未知脉冲响应的信道，自适应均衡器的用途是通过操作信道输出，使得信道与均衡器的级联提供理想传输媒介的一个良好逼近
	预测编码	自适应预测用来开发感兴趣信号（如语音信号）的模型，而不是直接对信号编码。在该编码中，预测误差被编码以便传输和存储。典型地，预测误差具有比原信号更小的方差，从而作为改进编码的基础
干扰消除	谱分析	在这项应用中，预测模型用于估计某一感兴趣信号的功率谱
	噪声消除	自适应噪声消除器的目的是从接收信号中减去噪声以改善信噪比。电话电路中遇到的回音消除是噪声消除的一种特殊形式。噪声消除也用于心电图学中
	波束形成	波束形成器是由具有可调加权系数的天线阵元组成的空间滤波器。自适应波束形成器的双重目的是，自适应地控制加权系数以便消除与未知方向阵列密切有关的干扰信号，并同时感兴趣目标信号提供保护

## 历史回顾

——为了理解一种科学，必须知道它的历史 [Auguste Comte(1798 ~ 1857)]

我们将通过回顾与本书课题紧密相关的三个研究领域来完成这个导引性章节。这三个领