

# 第1章 绪论

## 1.1 基于麦克风阵列的信号处理技术研究

麦克风阵列是以特定方式排列，从而能够准确获取空间信息的一组麦克风。类比于无线通信的空间分集，这里讨论麦克风阵列的空间多样性。这种多样性通常由辐射源到传感器之间的声脉冲响应来表示，并且可以通过不同的方式加以理解和应用，本书将对此进行详细阐释。这些采用有限冲激响应（Finite Impulse Response, FIR）滤波器建模的声信道通常是非唯一的，尤其是当 FIR 滤波器具有共同的零点时问题最大。

对于空间多样性所提供的丰富信息需要做进一步处理。因此，麦克风阵列信号处理的主要目的就是：根据应用的不同，利用麦克风阵列输出信号中包含的空域-时域（也可能是频域）信息，估计某些参数或提取感兴趣的信号。尽管本书中也涉及单麦克风系统的特殊情况（第 2 章中的最优滤波器，第 6 章中的频域噪声抑制），但本书的重点主要集中在多传感器系统的应用方面，其原因在于多传感系统更加灵活，因而能够更好地解决许多重要的实际问题。

根据麦克风应用场合不同，麦克风阵列可以采用不同的几何排列，麦克风阵列的几何排列形状对处理算法的构建具有重要影响。例如，在声源定位中，阵列的几何排列形状必须是已知的，由此才能正确地对声源进行定位。在某些情况下，适当的规则排列可以简化估计问题，这就是均匀线性阵列和圆形阵列得到广泛应用的原因<sup>[148]</sup>。尽管目前这两种排列形状的阵列占据了大部分的市场，但同时也出现了一些能够更好地获取声场信息的高复杂度三维球状阵列<sup>[163, 164]</sup>。然而，在另外一些关键问题，例如噪声抑制或者源分离应用中，阵列几何形状对算法的影响较小。在这种情况下，可以将系统称为多麦克风系统，而不称为麦克风阵列。二者的区别在具体场合下是显而易见的，因此不需要对这两种情况严格区分。

麦克风阵列中遇到的许多问题看似容易解决，因为类似的问题在窄带天线阵列中已经得到了相当长时间的研究。事实却并非如此。实际上，麦克风阵列与天线阵列（如雷达和声纳）在实际应用中的工作方式是不同的，其原因如下<sup>[105, 215]</sup>：

- (1) 语音信号是宽带信号。
- (2) 室内混响（或者多径）效应显著。

- (3) 环境特性和信号均高度非平稳。
- (4) 噪声与目标语音信号可能具有相同的频谱特性。
- (5) 传感器的数目通常有限。
- (6) 人耳听力具有极宽的动态范围(听力正常情况下高达120dB),并且对通道冲激响应的弱拖尾极其敏感。因此,滤波器模型的长度很长(抽头数目数以千计的情况并不少见)。

正是由于这些问题的存在,现有的许多算法在一些实际问题的解决中性能并不好。

许多麦克风阵列处理算法都是借用窄带阵列处理算法或者对其进行简单推广后得到的<sup>[51]</sup>。这种策略的优点在于,能够简便地对数十年来形成的天线阵列算法进行扩展。但是,缺点在于,这些算法并不是针对实际的声学环境而设计,其性能往往受到很大限制。简单来说,麦克风阵列需要的是宽带处理。总的来说,这也正是本书所采用的研究方法。

通过麦克风阵列能够解决的实际问题主要包括:

- (1) 噪声抑制。
- (2) 回声抑制。
- (3) 去混响。
- (4) 单声源定位。
- (5) 声源数目估计。
- (6) 多声源定位。
- (7) 源分离。
- (8) 鸡尾酒会效应。

图1.1对上述的大多数问题进行了描述,其中,所有麦克风拾取得到的信号被送给一组给定的滤波器,而这些滤波器需要根据上述实际问题进行优化设计<sup>[20]</sup>。

噪声抑制算法的目的就是要从受到加性噪声影响的观测信号中估计出目标语音信号。对于这一问题,已经存在许多基于单麦克风的技术<sup>[16,154,156]</sup>。然而,所有单麦克风噪声抑制算法的主要问题在于它会使语音信号失真<sup>[41,42]</sup>。尽管语音质量得到改善,但可懂度却降低了。然而,采用麦克风阵列,却可以在不太影响语音信号的条件下抑制噪声(至少在理论上可以实现)。

在免提通信中,由于扬声器与麦克风之间的声场耦合以及系统延时的存在,导致回声效应的产生,使得实时交谈变得非常困难<sup>[10,29,84,98,99,121]</sup>。并且,声学系统可能变得极不稳定。利用麦克风阵列技术,使阵列指向目标声源,同时在扬声器方向陷零,将能够显著降低回声强度。尽管这种思路简单直接且极具吸引力,但是在实际中并没有得到应用。对于这一问题,当前最好的解决方案仍然是采用声学回声抵消<sup>[10]</sup>。

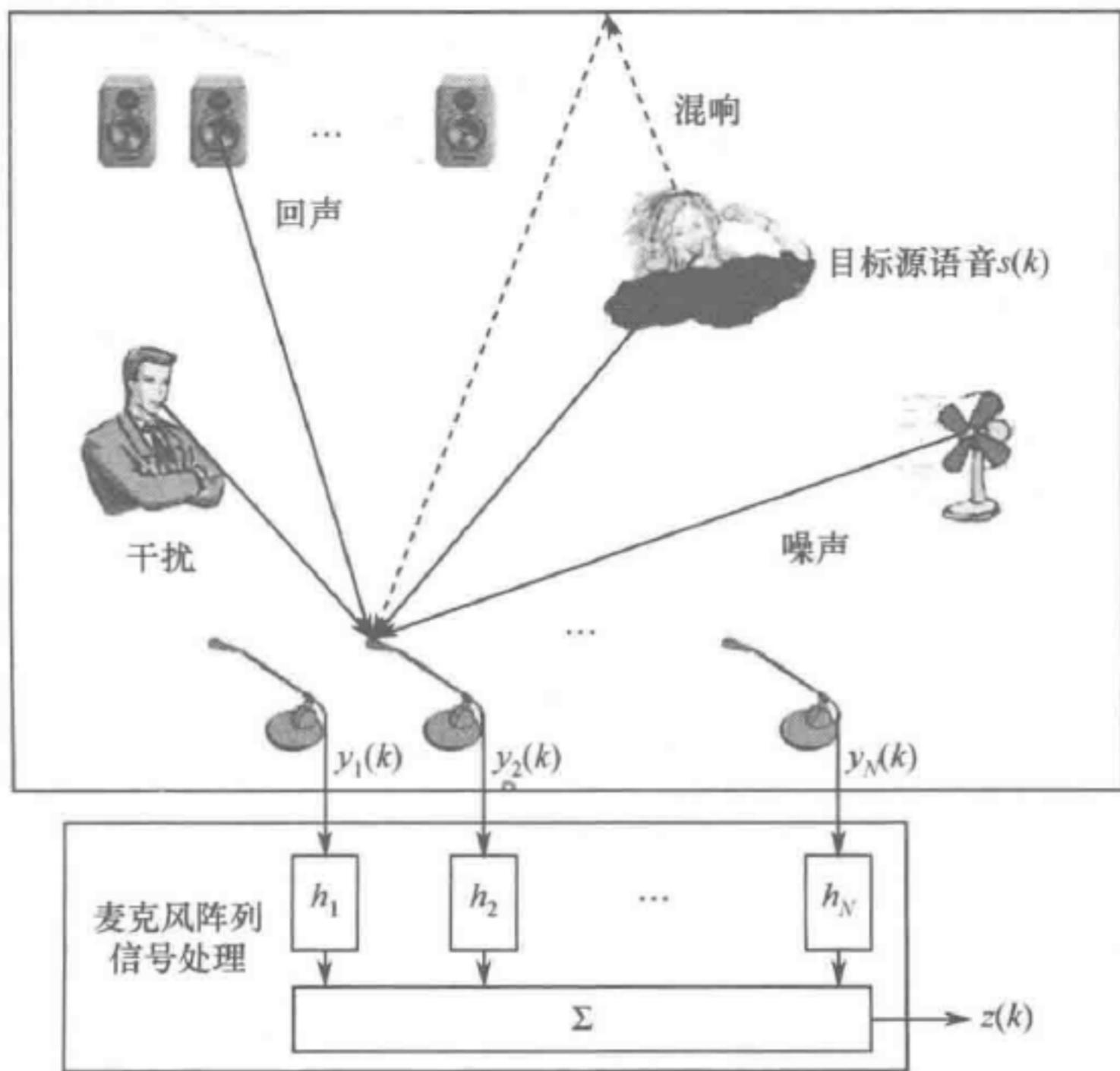


图 1.1 麦克风阵列信号处理

在房间内的免提通信场景下，麦克风采集到的语音信号不仅包含直接路径传播的信号，而且包含了由于房间边界以及房间内物体的反射而产生的延迟衰减信号。这种多径传播效应在观测信号中引入了回声和谱失真，称为混响，这可能会严重影响源信号，造成信号质量和可懂度下降。因此，去除混响对提高语音信号的可懂度有很大的作用<sup>[125]</sup>。四十多年来，为了利用麦克风阵列，研究者做了大量的工作寻求切实可行的解决方案。

在声学环境下，声源位置信息对于许多实际应用发挥着重要作用，例如视频会议中的自动摄像机跟踪、噪声抑制和去混响中波束方向的调整。信号源位置的估计，通常也称为声源定位问题，数十年来受到了极大的关注<sup>[26,117,175,222]</sup>。要实现对声源的入射角度或者其在直角坐标系中位置的估计，需要采用二维或三维的麦克风阵列。对声源数目的估计和对多个声源的定位是两个相互联系的问题，对于窄带信号源已有一些相关的算法，但是，这些问题在宽带信号源中的研究才刚刚起步。

在多麦克风声源分离中，需要同时对来自不同方向的信号进行分离。由于不能直接获取声通道和声源信号的信息，因此所有这些方法本质上都属于盲处理。独立成分分析 (Independent Component Analysis, ICA) 是盲源分离 (Blind Source Separation, BSS) 问题中使用最为广泛的工具<sup>[127]</sup>，它充分利用了源信号的独立性。在信号瞬时混合的情况下，多数基于 ICA 的算法都能够获得良好的性能，然而，在混响（卷积）环境下，此类算法的性能并不是很好。对于这一问题，近年来的研究取得很大进展，但如何将其应用于语音和声学环境，

以及可实用化的程度并不明确。有关 ICA 方面的文献已经相当丰富<sup>[159]</sup>（参见文献[182]），因此本书不再从这一角度讨论 BSS。

众所周知，人类具有在嘈杂的谈话环境或背景噪声中专注于某一特定的话音或声音的能力。这一有趣的心理声学现象称为鸡尾酒会效应<sup>[45,46]</sup>。从文献[45,46]中的心理声学实验中可以观察到一个重要的现象，即空间听觉发挥着重要作用。人类对语音的感知很大程度上得益于空间听觉能力，而这种能力又主要归因于双耳的作用。这一点可以通过日常经验直观地验证，同时还可以观察单耳和双耳在多个说话人同时存在的密闭空间环境中对声音的理解力差异来进一步说明<sup>[125]</sup>。具有正常听力和大脑处理能力的人都可以毫不费力地有效解决这种鸡尾酒会问题，但对于麦克风阵列信号处理仍然是一个极其复杂的问题。这一问题成为该领域所面临的诸多挑战的根源所在，直到今天仍然没有一个明确的解决思路。

无论麦克风阵列的尺寸、几何形状或阵元数目如何，上述问题都是极难解决的。某些情况下，采用特定几何形状的阵列或增加麦克风的数目可以更准确地解决问题。然而，这种改善可能非常有限，甚至可以忽略不计。因此，出现了一些根本性问题：如何利用空间信息？对于一个具体问题，能够解决到何种程度？合理的模型是怎样的？可以达到的极限是什么及其理由？除了空间信息还可以利用其他信息吗？如果可以，如何利用？

本书的目的并不是为了针对上述问题提出不同解决方案，也不是为了描述最新的研究成果，而是为了提出总体框架、介绍重要工具以及信号模型，以帮助读者更加直观而精确地理解多麦克风信号的处理。

最后，这里简要列出一些麦克风阵列的典型应用：

- (1) 电话会议；
- (2) 多方通信；
- (3) 免提耳机声音交互；
- (4) 对话系统；
- (5) 计算机游戏；
- (6) 指挥和控制接口；
- (7) 听写系统；
- (8) 高品质录音；
- (9) 声学监视；
- (10) 听觉场景分析；
- (11) 传感器网络技术。

可以看出，麦克风阵列的应用范围很广并且发展迅猛。显然，优秀的麦克风阵列技术及其解决方案符合市场需求，必将得到广泛应用。

## 1.2 本书的内容与结构

本书共 10 章（包括本章）。后面的 9 章尝试从一个崭新的视角对麦克风阵列信号处理所涉及的重要问题进行研究与探索。每章都从基本概念逐步深入，使读者可以顺着从基本理论到实际应用的思路进行学习研究。各章之间虽然存在连贯性，但又是自包含的，可以独立地进行学习。

线性最优滤波器在麦克风阵列信号处理等许多应用中都具有根本性的作用。最优滤波的概念本身容易理解，对本书的其他章节也非常重要。第 2 章主要研究维纳滤波器、Frost 滤波器和卡尔曼滤波器。本章还提出了 Pearson 相关系数的概念，并将其作为均方误差的替代准则。这一研究得出了许多有意义的成果。

传统的应用于窄带天线阵列空间滤波的波束形成方法已经相当成熟。第 3 章以信号增强为应用背景，通过采用一个简单信号传播模型，讨论了这一方面的相关典型技术。第 3 章还介绍了宽带波束形成的理念，这一理念对于语音信号具有重要意义。

线性约束最小方差 (Linear Constrained Minimum Variance, LCMV) 滤波器在天线阵列中的应用非常广泛。这种最优滤波器非常强大，其原因在于它可以将所有的约束与其代价函数联系起来。第 4 章通过使用三种不同的信号模型，介绍了在室内声学环境下如何使用 LCMV 滤波器进行噪声抑制和混响去除。

第 5 章专门介绍了多麦克风噪声抑制问题，在一个统一的框架内推导了几种经典的多通道处理方法。此外，本章还对语音增强涉及的重要内容进行了讨论，包括噪声抑制能力和语音失真大小。

第 6 章介绍了非因果（频域）维纳滤波及其在降噪方面的应用，分为单通道和多通道两种情况。为帮助读者更好地理解频域算法的工作原理，本章还介绍了有关语音增强方面的许多基本内容，尤其是针对多个麦克风的情况。

第 7 章以目标信号和干扰信号作为输入，麦克风信号作为输出，从而将麦克风阵列视为一个多输入多输出 (Multiple Input Multiple Output, MIMO) 系统，并基于 MIMO 通道冲激响应建立一个总体框架，用于分析波束形成在信号提取、混响去除及干扰抑制方面的性能。

作为第 7 章的延续，第 8 章指明了如何采用分两步处理的策略，将干扰和混响这两个问题分离开，分别采用不同的方法来解决。并且，清楚地说明了实现分离的前提条件。这种分离的实现能够有助于更好地理解源分离和去混响之间的相互关系。

第9章研究了波达方向(Direction of Arrival, DOA)估计和波达时间差(Time Difference of Arrival, TDOA)估计这一重要问题。由于DOA估计问题在本质上与TDOA估计问题是相同的，因此本章更多地将研究重点放在TDOA估计算法上面，并介绍了许多具体算法：既介绍了经典算法，如互相关方法等，也介绍了许多新颖算法，如最小熵技术等。本章还讨论了多源情况下TDOA估计的原理。

最后，第10章对本书中未涉及的问题进行了讨论。

## 第2章 经典最优滤波器

### 2.1 引言

Norbert Wiener 是最早将加性噪声污染后的随机过程估计作为一个滤波问题开始研究的学者之一，他在其具有里程碑意义的文稿中，提出了平稳时间序列的外插、内插和平滑<sup>[234]</sup>。在他推导的最优估计中，需要对一个积分方程进行求解，该方程即为著名的 Wiener-Hopf 方程<sup>[233]</sup>。继 Wiener 发表其著作以后，Levinson 在离散时间域提出了相同的问题<sup>[152]</sup>，这对于该领域具有重大影响。事实上，正是由于 Levinson 的贡献，才使得 Wiener 的思想变得易于被工程师们所理解，因而也变得更加实用。文献[136]很好地概述了线性滤波理论以及该领域内多种新技术的发展历史。

由于维纳滤波既简洁又有效，因此在实际中有着非常广泛的应用。然而，这一最优滤波器并不适用于非平稳信号，并且在许多情况下，它还会对目标信号造成失真，这在本章会有进一步阐述。

1960 年，R. E. Kalman 发表了他的著名文献，其中描述了离散数据线性滤波问题的递归求解方法<sup>[137]</sup>。所谓的卡尔曼滤波器是基于以下事实：目标信号服从一个状态模型，与维纳滤波器相比，它在非平稳环境下也能很好地工作。这种序贯滤波器的另一个优点在于：在准确建模的情况下，不会造成目标信号的失真。

本章主要研究麦克风阵列中常见的三种重要滤波器：维纳滤波器、线性约束最小方差滤波器和卡尔曼滤波器。此外，本章还提出了一种新的基于 Pearson 相关系数的准则，用于替代均方误差（Mean Square Error, MSE）准则（该准则用于推导维纳滤波器），并且分析了这一新准则更便于使用的原因。

### 2.2 维纳滤波器

假设一个零均值的纯净语音信号  $x(k)$ ，受到一个零均值的噪声信号  $v(k)$ （与  $x(k)$  不相关的白噪声或有色噪声）的干扰，则在离散采样时刻  $k$ ，带噪语音信号表示为

$$y(k) = x(k) + v(k) \quad (2.1)$$

假设所有的信号都是平稳的，本节的目标就是要找到维纳意义下  $x(k)$  的最优估计<sup>[234]</sup>。

定义时刻  $k$  的纯净语音信号样值与其估计值之间的误差信号为

$$\begin{aligned} e(k) &= x(k) - z(k) \\ &= x(k) - \mathbf{h}^T \mathbf{y}(k) \end{aligned} \quad (2.2)$$

式中

$$\mathbf{h} = [h_0 \ h_1 \ \cdots \ h_{L-1}]^T$$

为长度是  $L$  的 FIR 滤波器，上标“T”表示矢量或矩阵的转置；

$$\mathbf{y}(k) = [y(k) \ y(k-1) \ \cdots \ y(k-L+1)]^T$$

是由观测信号  $y(k)$  的最近  $L$  个样值组成的矢量；

$$z(k) = \mathbf{h}^T \mathbf{y}(k) \quad (2.3)$$

为滤波器  $\mathbf{h}$  的输出。

MSE 准则可以表示为<sup>[103]</sup>

$$\begin{aligned} J(\mathbf{h}) &= E[e^2(k)] \\ &= \mathbf{h}^T \mathbf{R}_{yy} \mathbf{h} - 2\mathbf{r}_{yx}^T \mathbf{h} + \sigma_x^2 \end{aligned} \quad (2.4)$$

式中：  $E[\cdot]$  表示数学期望；

$$\mathbf{R}_{yy} = E[y(k)y^T(k)] \quad (2.5)$$

为观测信号  $y(k)$  的自相关矩阵，并假设其为满秩；

$$\mathbf{r}_{yx} = E[y(k)x(k)] \quad (2.6)$$

为带噪语音信号和纯净语音信号之间的互相关矢量；  $\sigma_x^2 = E[x^2(k)]$  为信号  $x(k)$  的方差，则得到最优维纳滤波器为

$$\mathbf{h}_W = \arg \min_{\mathbf{h}} J(\mathbf{h}) = \mathbf{R}_{yy}^{-1} \mathbf{r}_{yx} \quad (2.7)$$

然而，  $x(k)$  实际上是不可观测的，因此难以获得对  $\mathbf{r}_{yx}$  的估计。但是，

$$\begin{aligned} \mathbf{r}_{yx} &= E[y(k)x(k)] \\ &= E\{y(k)[y(k) - v(k)]\} \\ &= E[y(k)y(k)] - E\{[x(k) + v(k)]v(k)\} \\ &= E[y(k)y(k)] - E[v(k)v(k)] \\ &= \mathbf{r}_{yy} - \mathbf{r}_{vv} \end{aligned} \quad (2.8)$$

此时可以看出， $\mathbf{r}_{yx}$  取决于相关矢量  $\mathbf{r}_{yy}$  和  $\mathbf{r}_{vv}$ ，矢量  $\mathbf{r}_{yy}$ （也就是矩阵  $\mathbf{R}_{yy}$  的第一列）可以在语音和噪声混合的时间段估计得到，而  $\mathbf{r}_{vv}$  可以在只有噪声存在的时间段估计得到。

考虑长度  $L$  的特殊滤波器

$$\mathbf{h}_1 = [1 \ 0 \cdots 0]^T \quad (2.9)$$

其对应的 MSE 为

$$\begin{aligned} J(\mathbf{h}_1) &= E\left\{\left[x(k) - \mathbf{h}_1^T y(k)\right]^2\right\} \\ &= E\left\{\left[x(k) - y(k)\right]^2\right\} \\ &= E\left\{v^2(k)\right\} = \sigma_v^2 \end{aligned} \quad (2.10)$$

这意味着观测信号  $y(k)$  将不经改变地通过滤波器  $\mathbf{h}_1$  (无噪声抑制)。

利用式 (2.8) 以及  $\mathbf{h}_1 = \mathbf{R}_{yy}^{-1} \mathbf{r}_{yy}$ , 可以得到另一种形式的维纳滤波<sup>[15]</sup>:

$$\begin{aligned} \mathbf{h}_W &= \mathbf{h}_1 - \mathbf{R}_{yy}^{-1} \mathbf{r}_{vv} \\ &= \left[ \mathbf{I} - \mathbf{R}_{yy}^{-1} \mathbf{r}_{vv} \right] \mathbf{h}_1 \\ &= \left[ \frac{\mathbf{I}}{\text{SNR}} + \tilde{\mathbf{R}}_{vv}^{-1} \tilde{\mathbf{R}}_{xx} \right]^{-1} \tilde{\mathbf{R}}_{vv}^{-1} \tilde{\mathbf{R}}_{xx} \mathbf{h}_1 \end{aligned} \quad (2.11)$$

式中

$$\text{SNR} = \frac{\sigma_x^2}{\sigma_v^2} \quad (2.12)$$

表示输入信噪比 (Signal To Noise Ratio, SNR),  $\mathbf{I}$  表示单位矩阵, 并且

$$\begin{aligned} \tilde{\mathbf{R}}_{xx} &= \frac{\mathbf{R}_{xx}}{\sigma_x^2} = \frac{E[\mathbf{x}(k)\mathbf{x}^T(k)]}{\sigma_x^2} \\ \tilde{\mathbf{R}}_{vv} &= \frac{\mathbf{R}_{vv}}{\sigma_v^2} = \frac{E[\mathbf{v}(k)\mathbf{v}^T(k)]}{\sigma_v^2} \end{aligned}$$

得到

$$\lim_{\text{SNR} \rightarrow \infty} \mathbf{h}_W = \mathbf{h}_1 \quad (2.13)$$

$$\lim_{\text{SNR} \rightarrow 0} \mathbf{h}_W = \mathbf{0}_{L \times 1} \quad (2.14)$$

式中:  $\mathbf{0}_{L \times 1}$  为全零矢量, 其长度与  $\mathbf{h}_W$  的长度相同。最小均方误差 (Minimum Mean Squared Error, MMSE) 为

$$\begin{aligned} J(\mathbf{h}_W) &= \sigma_x^2 - \mathbf{r}_{yx}^T \mathbf{h}_W \\ &= \sigma_v^2 - \mathbf{r}_{vv}^T \mathbf{R}_{yy}^{-1} \mathbf{r}_{vv} \\ &= \mathbf{r}_{vv}^T \mathbf{h}_W \\ &= \mathbf{h}_1^T (\mathbf{R}_{yy} - \mathbf{R}_{yy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{yy}) \mathbf{h}_1 \end{aligned} \quad (2.15)$$

从式 (2.15) 可以清晰地看出  $J(\mathbf{h}_W) < J(\mathbf{h}_1)$ , 因此有可能实现噪声抑制。

归一化的 MMSE 可表示为

$$\tilde{J}(\mathbf{h}_W) = \frac{J(\mathbf{h}_W)}{J(\mathbf{h}_1)} = \frac{J(\mathbf{h}_W)}{\sigma_v^2} \quad (2.16)$$

式中:  $0 < \tilde{J}(\mathbf{h}_W) < 1$ 。

因此, 在维纳意义下对纯净语音  $x(k)$  的最优估计为

$$\begin{aligned} z_W(k) &= \mathbf{h}_W^T \mathbf{y}(k) \\ &= y(k) - \mathbf{r}_{vv}^T \mathbf{R}_{yy}^{-1} \mathbf{y}(k) \end{aligned} \quad (2.17)$$

因此, 估计信号的方差为

$$\begin{aligned} E[z_W^2(k)] &= \mathbf{h}_W^T \mathbf{R}_{yy} \mathbf{h}_W \\ &= \mathbf{h}_W^T \mathbf{R}_{xx} \mathbf{h}_W + \mathbf{h}_W^T \mathbf{R}_{vv} \mathbf{h}_W \end{aligned} \quad (2.18)$$

式 (2.18) 是两项之和, 其中第一项为纯净语音衰减后的功率, 第二项为残留噪声的功率 (始终大于零)。由式 (2.18) 可以看出, 维纳滤波有可能实现噪声抑制, 但同时付出的代价是引起纯净语音的衰减, 从而导致语音失真。

定义 (维纳滤波的) 降噪因子为<sup>[15]</sup>

$$\begin{aligned} \xi_{nr}(\mathbf{h}_W) &= \frac{\mathbf{h}_1^T \mathbf{R}_{vv} \mathbf{h}_1}{\mathbf{h}_W^T \mathbf{R}_{xx} \mathbf{h}_W} \\ &= \frac{\mathbf{h}_1^T \mathbf{R}_{vv} \mathbf{h}_1}{\mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{R}_{yy}^{-1} \mathbf{R}_{vv} \mathbf{R}_{yy}^{-1} \mathbf{R}_{xx} \mathbf{h}_1} \end{aligned} \quad (2.19)$$

语音失真指数为<sup>[15]</sup>

$$\begin{aligned} v_{sd}(\mathbf{h}_W) &= \frac{E\left\{ [x(k) - \mathbf{h}_W^T x(k)]^2 \right\}}{\sigma_x^2} \\ &= \frac{(\mathbf{h}_1 - \mathbf{h}_W)^T \mathbf{R}_{xx} (\mathbf{h}_1 - \mathbf{h}_W)}{\mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h}_1} \end{aligned} \quad (2.20)$$

降噪因子始终大于 1, 且  $\xi_{nr}(\mathbf{h}_W)$  的值越大, 降噪程度越大。并且

$$\lim_{SNR \rightarrow 0} \xi_{sd}(\mathbf{h}_W) = \infty \quad (2.21)$$

$$\lim_{SNR \rightarrow \infty} \xi_{nr}(\mathbf{h}_W) = 1 \quad (2.22)$$

维纳滤波的语音失真指数通常介于 0~1 之间, 并且

$$\lim_{SNR \rightarrow 0} v_{sd}(\mathbf{h}_W) = 1 \quad (2.23)$$

$$\lim_{SNR \rightarrow \infty} v_{sd}(\mathbf{h}_W) = 0 \quad (2.24)$$

因此, 当  $v_{sd}(\mathbf{h}_W)$  趋近于 1 时, 语音信号严重失真, 而当  $v_{sd}(\mathbf{h}_W)$  趋近于 0 时, 语音信号轻微失真。由此可以看出, 当信噪比较低的时候, 维纳滤波会对语音信号造成极为严重的影响。

根据参考文献[77], 若  $\mathbf{R}_{vv}$  为正定矩阵, 则可以对对称矩阵  $\mathbf{R}_{xx}$  和  $\mathbf{R}_{vv}$  进行联合对角化, 从而得到

$$\mathbf{R}_{xx} = \mathbf{B}^T \Lambda \mathbf{B} \quad (2.25)$$

$$\mathbf{R}_{vv} = \mathbf{B}^T \mathbf{B} \quad (2.26)$$

$$\mathbf{R}_{yy} = \mathbf{B}^T [\mathbf{I} + \mathbf{A}] \mathbf{B} \quad (2.27)$$

式中:  $\mathbf{B}$  为满秩但不一定正交的方阵; 对角阵

$$\mathbf{A} = \text{diag}[\lambda_1 \lambda_2 \dots \lambda_L] \quad (2.28)$$

由矩阵  $\mathbf{R}_{yy}^{-1} \mathbf{R}_{xx}$  的特征值组成, 且  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$ 。将式 (2.25) ~ 式 (2.27) 代入式 (2.19) 后得到:

$$\xi_{nr}(\mathbf{h}_W) = \frac{\sum_{l=1}^L b_{l1}^2}{\sum_{l=1}^L \frac{\lambda_l^2}{(1+\lambda_l)} b_{l1}^2} \quad (2.29)$$

式中: 元素  $b_{l1}, l=1, 2, \dots, L$  组成矩阵  $\mathbf{B}$  的第一列, 并且满足  $\sum_{l=1}^L b_{l1}^2 = \sigma_v^2$ 。

利用式 (2.25) ~ 式 (2.27) 对矩阵的分解, 输入信噪比可表示为

$$\begin{aligned} \text{SNR} &= \frac{\mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h}_1}{\mathbf{h}_1^T \mathbf{R}_{yy} \mathbf{h}_1} \\ &= \frac{\sum_{l=1}^L \lambda_l b_{l1}^2}{\sum_{l=1}^L b_{l1}^2} \end{aligned} \quad (2.30)$$

利用式 (2.30), 式 (2.29) 可重写为

$$\begin{aligned} \xi_{nr}(\mathbf{h}_W) &= \frac{1}{\text{SNR}} \cdot \frac{\sum_{l=1}^L \lambda_l b_{l1}^2}{\sum_{l=1}^L \frac{\lambda_l^2}{(1+\lambda_l)^2} b_{l1}^2} \\ &= \frac{1}{\text{SNR}} \cdot \frac{\sum_{l=1}^L \frac{(1+\lambda_l)^2}{(1+\lambda_l)^2} \lambda_l b_{l1}^2}{\sum_{l=1}^L \frac{\lambda_l^2}{(1+\lambda_l)^2} b_{l1}^2} \\ &= \frac{1}{\text{SNR}} \cdot \left[ \frac{\sum_{l=1}^L \frac{\lambda_l + \lambda_l^3}{(1+\lambda_l)^2} b_{l1}^2}{\sum_{l=1}^L \frac{\lambda_l^2}{(1+\lambda_l)^2} b_{l1}^2} + 2 \right] \end{aligned} \quad (2.31)$$

根据  $\lambda_l + \lambda_l^3 \geq \lambda_l^3$ , 容易从式 (2.31) 推导出

$$\xi_{nr}(\mathbf{h}_W) \geq \frac{1}{\text{SNR}} \cdot \left[ \frac{\sum_{l=1}^L \frac{\lambda_l^3}{(1+\lambda_l)^2} b_{l1}^2}{\sum_{l=1}^L \frac{\lambda_l^2}{(1+\lambda_l)^2} b_{l1}^2} + 2 \right] \quad (2.32)$$

可以证明如下不等式（参考下面定理的证明过程）：

$$\frac{\sum_{l=1}^L \frac{\lambda_l^3}{(1+\lambda_l)^2} b_{ll}^2}{\sum_{l=1}^L \frac{\lambda_l^2}{(1+\lambda_l)^2} b_{ll}^2} \geq \frac{\sum_{l=1}^L \lambda_l b_{ll}^2}{\sum_{l=1}^L b_{ll}^2} = \text{SNR} \quad (2.33)$$

当且仅当所有对应于非零  $b_{ll}$  的  $\lambda_l$  ( $l=1, 2, \dots, L$ ) 相等时，式 (2.33) 中的等号成立。由此得出<sup>[41,42]</sup>：

$$\xi_{nr}(\mathbf{h}_W) \geq \frac{\text{SNR} + 2}{\text{SNR}} \geq 1 \quad (2.34)$$

从式 (2.34) 可以得出，降噪因子的下限是关于信噪比的单调递减函数。当信噪比趋近于 0 时，降噪因子接近于无穷大，当信噪比趋近于无穷时，降噪因子接近于 1。这表明，随着信噪比的减小，维纳滤波可以消除更多的噪声，这也正好满足实际需要，因为随着信噪比的降低，会有更多的噪声需要消除。

语音失真指数的上限可以由式 (2.25) ~ 式 (2.27) 所给出的特征值分解推导得到。事实上，将式 (2.25) ~ 式 (2.27) 代入式 (2.20)，可得<sup>[41,42]</sup>

$$\begin{aligned} v_{sd}(\mathbf{h}_W) &= \frac{\sum_{l=1}^L \frac{\lambda_l}{(1+\lambda_l)^2} b_{ll}^2}{\sum_{l=1}^L \lambda_l b_{ll}^2} \\ &\leq \frac{\sum_{l=1}^L \frac{\lambda_l}{1+2\lambda_l} \lambda_l b_{ll}^2}{\sum_{l=1}^L \frac{\lambda_l+2\lambda_l^2}{1+2\lambda_l} b_{ll}^2} \\ &\leq \frac{1}{2 \cdot \text{SNR} + 1} \end{aligned} \quad (2.35)$$

式中：利用到如下不等式

$$\frac{\sum_{l=1}^L \frac{\lambda_l^2}{1+2\lambda_l} b_{ll}^2}{\sum_{l=1}^L \frac{\lambda_l}{1+2\lambda_l} b_{ll}^2} \geq \frac{\sum_{l=1}^L \lambda_l b_{ll}^2}{\sum_{l=1}^L b_{ll}^2} = \text{SNR} \quad (2.36)$$

这一不等式可以用归纳法进行证明。

图 2.1 所示为降噪因子的下限(式(2.34))和语音失真指数的上限(式(2.35))，二者都是输入信噪比的函数。

从前面的分析可以看出，维纳滤波器是以语音信号的衰减为代价实现噪声抑制的。因此，仅以降噪因子作为评判标准是难以令人满意的。事实上，最为恰当的评价指标是输出信噪比，其定义为

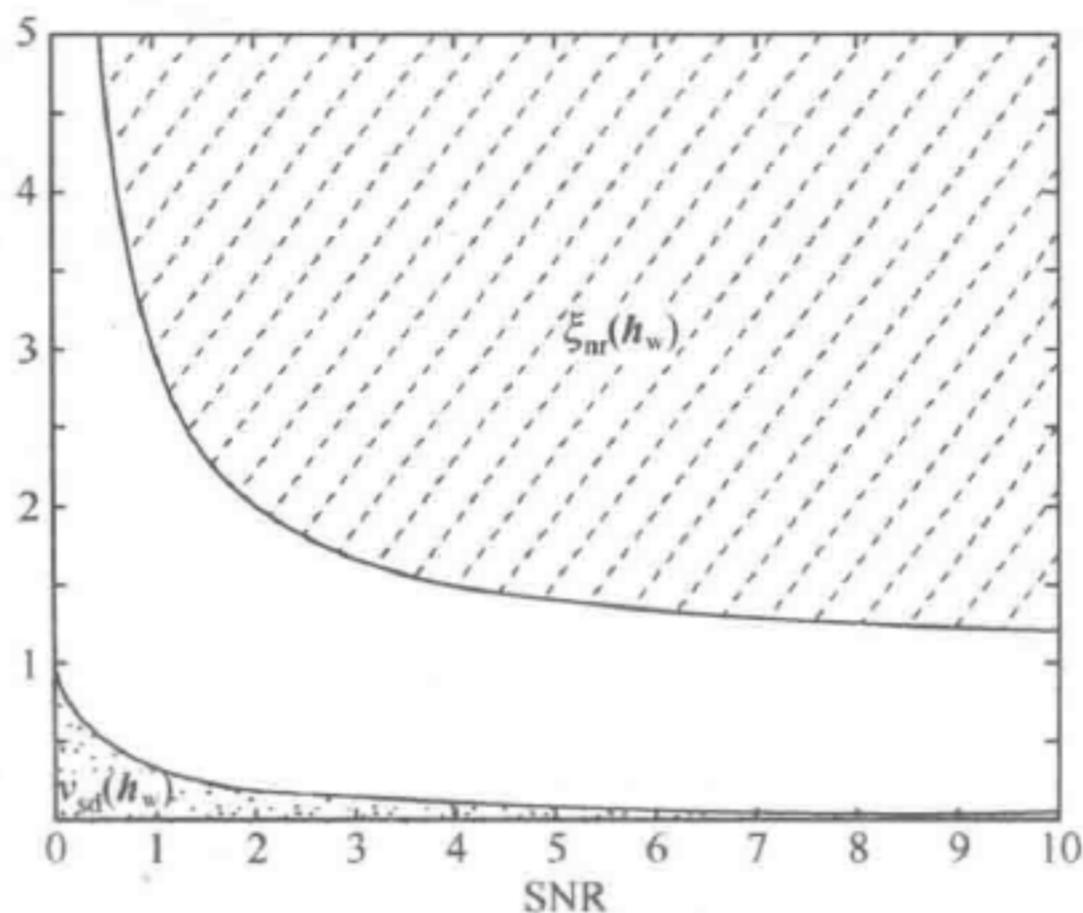


图 2.1 以输入信噪比为参数,  $\xi_{nr}(\mathbf{h}_w)$  和  $v_{sd}(\mathbf{h}_w)$  取值范围的示意图, 其中实线以上是  $\xi_{nr}(\mathbf{h}_w)$  的取值范围, 虚线以下是  $v_{sd}(\mathbf{h}_w)$  的取值范围

$$\text{SNR}(\mathbf{h}_w) = \frac{\mathbf{h}_w^T \mathbf{R}_{xx} \mathbf{h}_w}{\mathbf{h}_w^T \mathbf{R}_{vv} \mathbf{h}_w} \quad (2.37)$$

如果  $\text{SNR}(\mathbf{h}_w) > \text{SNR}$ , 则说明与语音的衰减相比, 维纳滤波器对噪声的抑制方面更为有效。此时存在一个关键问题, 即维纳滤波是否能够提高信噪比。为了回答这一问题, 这里给出以下定理<sup>[15, 41, 42]</sup>。

**定理** 式(2.7)给出的最优维纳滤波的输出信噪比(式(2.37))始终大于或等于其输入信噪比(式(2.12))。

**证明:** 若噪声  $v(k)$  为零, 则维纳滤波器等价于  $\mathbf{h}_1$ , 其对语音信号没有任何影响。对式(2.37)应用矩阵分解(式(2.25)~式(2.27)), 则输出信噪比可以重写为

$$\text{SNR}(\mathbf{h}_w) = \frac{\sum_{l=1}^L \frac{\lambda_l^3}{(\lambda_l+1)^2} b_{l1}^2}{\sum_{l=1}^L \frac{\lambda_l^2}{(\lambda_l+1)^2} b_{l1}^2} \quad (2.38)$$

则由此得出

$$\frac{\text{SNR}(\mathbf{h}_w)}{\text{SNR}} = \frac{\sum_{l=1}^L b_{l1}^2 \cdot \sum_{l=1}^L \frac{\lambda_l^3}{(\lambda_l+1)^2} b_{l1}^2}{\sum_{l=1}^L \lambda_l b_{l1}^2 \cdot \sum_{l=1}^L \frac{\lambda_l^2}{(\lambda_l+1)^2} b_{l1}^2} \quad (2.39)$$

由于所有的求和项  $\sum_{l=1}^L \frac{\lambda_l^3}{(\lambda_l+1)^2} b_{l1}^2$ ,  $\sum_{l=1}^L \frac{\lambda_l^2}{(\lambda_l+1)^2} b_{l1}^2$ ,  $\sum_{l=1}^L \lambda_l b_{l1}^2$  和  $\sum_{l=1}^L b_{l1}^2$  均为非负数,

因此, 只要能够说明不等式

$$\sum_{l=1}^L \frac{\lambda_l^3}{(1+\lambda_l)^2} b_{l1}^2 \sum_{l=1}^L b_{l1}^2 \geq \sum_{l=1}^L \frac{\lambda_l^2}{(1+\lambda_l)^2} b_{l1}^2 \sum_{l=1}^L \lambda_l b_{l1}^2 \quad (2.40)$$

成立，则  $\text{SNR}(\mathbf{h}_w) \geq \text{SNR}$ 。现在通过归纳法证明这一不等式。

(1) 基本步骤：若  $L=2$ ，则

$$\sum_{l=1}^2 \frac{\lambda_l^3}{(1+\lambda_l)^2} b_{ll}^2 \sum_{l=1}^2 b_{ll}^2 = \frac{\lambda_1^3}{(1+\lambda_1)^2} b_{11}^4 + \frac{\lambda_2^3}{(1+\lambda_2)^2} b_{21}^4 + \left[ \frac{\lambda_1^3}{(1+\lambda_1)^2} + \frac{\lambda_2^3}{(1+\lambda_2)^2} \right] b_{11}^2 b_{21}^2$$

由于  $\lambda_l \geq 0$ ，容易证明

$$\frac{\lambda_1^3}{(1+\lambda_1)^2} + \frac{\lambda_2^3}{(1+\lambda_2)^2} \geq \frac{\lambda_1^2 \lambda_2}{(1+\lambda_1)^2} + \frac{\lambda_1 \lambda_2^2}{(1+\lambda_2)^2}$$

其中，等号在  $\lambda_1 = \lambda_2$  时成立。因此

$$\begin{aligned} \sum_{l=1}^2 \frac{\lambda_l^3}{(1+\lambda_l)^2} b_{ll}^2 \sum_{l=1}^2 b_{ll}^2 &\geq \frac{\lambda_1^3}{(1+\lambda_1)^2} b_{11}^4 + \frac{\lambda_2^3}{(1+\lambda_2)^2} b_{21}^4 + \left[ \frac{\lambda_1^2 \lambda_2}{(1+\lambda_1)^2} + \frac{\lambda_1 \lambda_2^2}{(1+\lambda_2)^2} \right] b_{11}^2 b_{21}^2 \\ &= \sum_{l=1}^2 \frac{\lambda_l^2}{(1+\lambda_l)^2} b_{ll}^2 \sum_{l=1}^2 \lambda_l b_{ll}^2 \end{aligned}$$

因此，该定理对  $L=2$  是成立的，其中的等号在  $b_{11}$  和  $b_{21}$  中任意一个为 0（注意  $b_{11}$  和  $b_{21}$  不能同时为 0，否则  $\mathbf{B}$  不可逆）时或者  $\lambda_1 = \lambda_2$  时成立。

(2) 推演步骤：假设该定理对  $L=P$  是成立的，即

$$\sum_{l=1}^P \frac{\lambda_l^3}{(1+\lambda_l)^2} b_{ll}^2 \sum_{l=1}^P b_{ll}^2 \geq \sum_{l=1}^P \frac{\lambda_l^2}{(1+\lambda_l)^2} b_{ll}^2 \sum_{l=1}^P \lambda_l b_{ll}^2$$

下面必须证明它对于  $L=P+1$  也是成立的。事实上，

$$\begin{aligned} \sum_{l=1}^{P+1} \frac{\lambda_l^3}{(1+\lambda_l)^2} b_{ll}^2 \sum_{l=1}^{P+1} b_{ll}^2 &= \left[ \sum_{l=1}^P \frac{\lambda_l^3}{(1+\lambda_l)^2} b_{ll}^2 + \frac{\lambda_{P+1}^3}{(1+\lambda_{P+1})^2} b_{(P+1)1}^2 \right] \left[ \sum_{l=1}^P b_{ll}^2 + b_{(P+1)1}^2 \right] \\ &= \left[ \sum_{l=1}^P \frac{\lambda_l^3}{(1+\lambda_l)^2} b_{ll}^2 \right] \left[ \sum_{l=1}^P b_{ll}^2 \right] + \frac{\lambda_{P+1}^3}{(1+\lambda_{P+1})^2} b_{(P+1)1}^4 + \\ &\quad \sum_{l=1}^P \left[ \frac{\lambda_l^3}{(1+\lambda_l)^2} + \frac{\lambda_{P+1}^3}{(1+\lambda_{P+1})^2} \right] b_{ll}^2 b_{(P+1)1}^2 \end{aligned}$$

利用归纳假设，以及

$$\frac{\lambda_l^3}{(1+\lambda_l)^2} + \frac{\lambda_{P+1}^3}{(1+\lambda_{P+1})^2} \geq \frac{\lambda_l^2 \lambda_{P+1}}{(1+\lambda_l)^2} + \frac{\lambda_l \lambda_{P+1}^2}{(1+\lambda_{P+1})^2}$$

得到

$$\begin{aligned} \sum_{l=1}^{P+1} \frac{\lambda_l^3}{(1+\lambda_l)^2} b_{ll}^2 \sum_{l=1}^{P+1} b_{ll}^2 &\geq \sum_{l=1}^P \frac{\lambda_l^2}{(1+\lambda_l)^2} b_{ll}^2 \sum_{l=1}^P \lambda_l b_{ll}^2 + \frac{\lambda_{P+1}^3}{(1+\lambda_{P+1})^2} b_{(P+1)1}^4 + \\ &\quad \sum_{l=1}^P \left[ \frac{\lambda_l^2 \lambda_{P+1}}{(1+\lambda_l)^2} + \frac{\lambda_l \lambda_{P+1}^2}{(1+\lambda_{P+1})^2} \right] b_{ll}^2 b_{(P+1)1}^2 \\ &= \sum_{l=1}^{P+1} \frac{\lambda_l^2}{(1+\lambda_l)^2} b_{ll}^2 \sum_{l=1}^{P+1} \lambda_l b_{ll}^2 \end{aligned}$$

其中, 当所有对应于非零  $b_l$  的  $\lambda_l (l=1, 2, \dots, P+1)$  均相等时, 等号成立。证明完毕。

尽管维纳滤波能够改进信噪比, 但它并没有使输出信噪比最大化。事实上, 式 (2.37) 正是著名的广义瑞利商。因此, 能够使输出信噪比最大化的滤波器为矩阵  $\mathbf{R}_{yy}^{-1}\mathbf{R}_{xx}$  最大特征值所对应的特征矢量 (参见 2.5 节)。然而, 这一滤波器通常会导致大的语音失真。

此外, 更具一般性的多通道维纳滤波噪声抑制将在第 5 章研究。

## 2.3 Frost 滤波器

线性约束最小方差 (LCMV) 滤波器<sup>[76]</sup>可以看作维纳滤波器的一种特殊形式, 本书也称其为 Frost 滤波器。

### 2.3.1 算法

在许多实际应用中, 参考信号难以获得, 甚至根本不存在。因此, 式 (2.2) 定义的误差信号实际上是没有意义的。

假设参考信号  $x(k)$  为 0, MSE 准则 (式 (2.4)) 变为

$$J(\mathbf{h}) = \mathbf{h}^T \mathbf{R}_{yy} \mathbf{h} \quad (2.41)$$

显然, 使  $J(\mathbf{h})$  取得最小值的  $\mathbf{h}$  为  $\mathbf{h} = \mathbf{0}_{L \times 1}$ 。但是, 在许多应用中, 滤波器  $\mathbf{h}$  具有如下形式的约束:

$$\mathbf{C}^T \mathbf{h} = \mathbf{u} \quad (2.42)$$

式中:  $\mathbf{C}$  为大小是  $L \times L_c$  的约束矩阵, 且

$$\mathbf{u} = [u_0 \ u_1 \ \cdots \ u_{L_c-1}]^T$$

是一个长度为  $L_c$  的矢量, 其中的元素需要提前确定。

在这种情况下, 要获得最优滤波器, 需要求解以下优化问题:

$$\min_{\mathbf{h}} J(\mathbf{h}) \quad \text{s.t.} \quad \mathbf{C}^T \mathbf{h} = \mathbf{u} \quad (2.43)$$

利用拉格朗日乘子将约束条件与代价函数结合起来, 容易得到 Frost 滤波器为<sup>[76]</sup>

$$\mathbf{h}_F = \mathbf{R}_{yy}^{-1} \mathbf{C} (\mathbf{C}^T \mathbf{R}_{yy}^{-1} \mathbf{C})^{-1} \mathbf{u} \quad (2.44)$$

需要注意的是, 为了保证该滤波器存在, 必须使相关矩阵  $\mathbf{R}_{yy}$  可逆, 且  $\mathbf{C}$  必须为列满秩, 这也就意味着  $L_c \leq L$ 。下面假设  $\mathbf{C}$  的秩等于  $L_c$ 。在  $L_c = L$  这一特定情况下, 可以由式 (2.42) 直接求解得到:  $\mathbf{h}_F = (\mathbf{C}^T)^{-1} \mathbf{u}$ , 此时它不再依赖于观测信号。而在  $L_c = 1$  的情况下, 约束矩阵  $\mathbf{C}$  变为一个约束矢量  $\mathbf{c}$ , 此时求解结果的形式类似于最小方差无失真响应滤波器<sup>[35,149]</sup>:

$$\mathbf{h}_F = u_0 \frac{\mathbf{R}_{yy}^{-1} \mathbf{c}}{\mathbf{c}^T \mathbf{R}_{yy}^{-1} \mathbf{c}} \quad (2.45)$$

### 2.3.2 广义旁瓣抵消结构

与 LCMV 方法所解决的问题相同, 广义旁瓣抵消结构通过将滤波器矢量  $\mathbf{h}_F$  分解为正交子空间上的两个分量来解决这一问题<sup>[31,54,94,230]</sup>:

$$\mathbf{h}_F = \mathbf{f} - \mathbf{B}_c \mathbf{w}_{GSC} \quad (2.46)$$

式中

$$\mathbf{f} = \mathbf{C}(\mathbf{C}^T \mathbf{C})^{-1} \mathbf{u} \quad (2.47)$$

为  $\mathbf{C}^T \mathbf{f} = \mathbf{u}$  的最小范数解;  $\mathbf{B}_c$  为阻塞矩阵, 它张成了  $\mathbf{C}^T$  的零空间, 即

$$\mathbf{C}^T \mathbf{B}_c = \mathbf{0}_{L_c \times (L-L_c)} \quad (2.48)$$

$\mathbf{w}_{GSC}$  为权重矢量, 其推导过程将在下面给出。矩阵  $\mathbf{B}_c$  的大小为  $L \times (L-L_c)$ , 其中  $L-L_c$  为  $\mathbf{C}^T$  零空间的维数。因此, 矢量  $\mathbf{w}_{GSC}$  的长度为  $L-L_c$ 。阻塞矩阵并不唯一, 最直观的选择是

$$\mathbf{B}_c = \begin{bmatrix} \mathbf{I}_{(L-L_c) \times (L-L_c)} \\ \mathbf{0}_{L_c \times (L-L_c)} \end{bmatrix} - \mathbf{C}(\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \begin{bmatrix} \mathbf{I}_{(L-L_c) \times (L-L_c)} \\ \mathbf{0}_{L_c \times (L-L_c)} \end{bmatrix} \quad (2.49)$$

采用广义旁瓣消除 (Generalized Sidelobe Concroller, GSC) 方法获得滤波器  $\mathbf{w}_{GSC}$ , 可以表示为以下无约束优化问题:

$$\min_{\mathbf{w}} (\mathbf{f} - \mathbf{B}_c \mathbf{w})^T \mathbf{R}_{yy} (\mathbf{f} - \mathbf{B}_c \mathbf{w}) \quad (2.50)$$

它的解为

$$\mathbf{w}_{GSC} = (\mathbf{B}_c^T \mathbf{R}_{yy} \mathbf{B}_c)^{-1} \mathbf{B}_c^T \mathbf{R}_{yy} \mathbf{f} \quad (2.51)$$

定义两个滤波器  $\mathbf{f}$  和  $\mathbf{B}_c \mathbf{w}$  的输出之间的误差信号为

$$e(k) = \mathbf{y}^T(k) \mathbf{f} - \mathbf{y}^T(k) \mathbf{B}_c \mathbf{w} \quad (2.52)$$

容易看出,  $E[e^2(k)]$  关于  $\mathbf{w}$  的最小化等价于式 (2.50)。

现在需要确认, 滤波器 LCMV 和 GSC 是否确实等价, 即

$$\mathbf{u}^T (\mathbf{C}^T \mathbf{R}_{yy}^{-1} \mathbf{C})^{-1} \mathbf{C}^T \mathbf{R}_{yy}^{-1} = \mathbf{f}^T [\mathbf{I} - \mathbf{R}_{yy} \mathbf{B}_c (\mathbf{B}_c^T \mathbf{R}_{yy} \mathbf{B}_c)^{-1} \mathbf{B}_c^T] \quad (2.53)$$

为此, 这里采用文献[28]中给出的证明。

式 (2.53) 右边方括号内的矩阵可以重写为

$$[\mathbf{I} - \mathbf{R}_{yy} \mathbf{B}_c (\mathbf{B}_c^T \mathbf{R}_{yy} \mathbf{B}_c)^{-1} \mathbf{B}_c^T] = \mathbf{R}_{yy}^{1/2} (\mathbf{I} - \mathbf{P}_1) \mathbf{R}_{yy}^{-1/2} \quad (2.54)$$

式中

$$\mathbf{P}_1 = \mathbf{R}_{yy}^{1/2} \mathbf{B}_c (\mathbf{B}_c^T \mathbf{R}_{yy} \mathbf{B}_c)^{-1} \mathbf{B}_c^T \mathbf{R}_{yy}^{1/2} \quad (2.55)$$

为矩阵  $\mathbf{R}_{yy}^{1/2} \mathbf{B}_c$  的列张成的子空间上的投影算子。由于  $\mathbf{B}_c^T \mathbf{C} = \mathbf{B}_c^T \mathbf{R}_{yy}^{1/2} \mathbf{R}_{yy}^{-1/2} \mathbf{C} = \mathbf{0}_{(L-L_c) \times L_c}$ , 这表明  $\mathbf{B}_c^T$  的行与  $\mathbf{C}$  的列正交, 且  $\mathbf{R}_{yy}^{1/2} \mathbf{B}_c$  的列张成的子空间与  $\mathbf{R}_{yy}^{-1/2} \mathbf{C}$  的列张成的子空间正交。由于  $\mathbf{B}_c$  的秩为  $L-L_c$ , 其中  $L_c$  为  $\mathbf{C}$  的秩, 则两个子空间的维数之和为  $L$ , 并且这两个子空间是互补的。这意味着

$$\mathbf{P}_1 + \mathbf{P}_2 = \mathbf{I} \quad (2.56)$$

式中

$$\mathbf{P}_2 = \mathbf{R}_{yy}^{-1/2} \mathbf{C} (\mathbf{C}^T \mathbf{R}_{yy}^{-1} \mathbf{C})^{-1} \mathbf{C}^T \mathbf{R}_{yy}^{-1/2} \quad (2.57)$$

对式 (2.53) 进行变量替换，并应用约束  $\mathbf{u}^T = \mathbf{f}^T \mathbf{C}$  之后，式 (2.53) 变为

$$\begin{aligned} \mathbf{u}^T (\mathbf{C}^T \mathbf{R}_{yy}^{-1} \mathbf{C})^{-1} \mathbf{C}^T \mathbf{R}_{yy}^{-1} &= \mathbf{f}^T \mathbf{R}_{yy}^{1/2} \mathbf{P}_2 \mathbf{R}_{yy}^{-1/2} \\ &= \mathbf{f}^T \mathbf{R}_{yy}^{1/2} (\mathbf{I} - \mathbf{P}_1) \mathbf{R}_{yy}^{-1/2} \\ &= \mathbf{f}^T [\mathbf{I} - \mathbf{R}_{yy} \mathbf{B}_c (\mathbf{B}_c^T \mathbf{R}_{yy} \mathbf{B}_c)^{-1} \mathbf{B}_c^T] \end{aligned} \quad (2.58)$$

由此可见，LCMV 和 GSC 是严格等价的。

### 2.3.3 在线性内插中的应用

本节主要阐述线性内插与 Frost 滤波器之间的关系。

线性内插是前向和后向线性预测的一种直接推广。实际上，线性内插通过信号的过去值和将来值来预测采样值  $y(k-i)$ <sup>[140, 183]</sup>。定义内插误差为

$$\begin{aligned} e_i(k) &= y(k-i) - \hat{y}(k-i) \\ &= y(k-i) - \sum_{l=0, l \neq i}^{L-1} h_{i,l} y(k-l) \\ &= \mathbf{h}_i^T \mathbf{y}(k), \quad i = 0, 1, \dots, L-1 \end{aligned} \quad (2.59)$$

式中： $\hat{y}(k-i)$  为内插值；

$$\mathbf{h}_i = [-h_{i,0} \cdots h_{i,1} \cdots -h_{i,L-1}]^T$$

为长度是  $L$  的矢量，它由内插系数组成且  $h_{i,i} = 1$ 。 $i = 0$  和  $i = L-1$  两种情况分别对应前向和后向预测误差。

为了得到最优内插，需要最小化以下代价函数：

$$J_i(\mathbf{h}_i) = E[e_i^2(k)] = \mathbf{h}_i^T \mathbf{R}_{yy} \mathbf{h}_i \quad (2.60)$$

并且满足以下约束条件：

$$\mathbf{c}_i^T \mathbf{h}_i = h_{i,i} = 1 \quad (2.61)$$

式中

$$\mathbf{c}_i = [0 \ 0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0]^T$$

为长度是  $L$  的约束矢量，它的第  $(i+1)$  个元素为 1，其余元素均为 0。利用拉格朗日乘子，容易看出，这一优化问题的解为

$$\mathbf{R}_{yy} \mathbf{h}_{o,i} = E_i \mathbf{c}_i \quad (2.62)$$

式中

$$E_i = \mathbf{h}_{o,i}^T \mathbf{R}_{yy} \mathbf{h}_{o,i} = \frac{1}{\mathbf{c}_i^T \mathbf{R}_{yy}^{-1} \mathbf{c}_i} \quad (2.63)$$

为内插误差的功率。因此，

$$\mathbf{h}_{o,i} = \frac{\mathbf{R}_{yy}^{-1} \mathbf{c}_i}{\mathbf{c}_i^T \mathbf{R}_{yy}^{-1} \mathbf{c}_i} \quad (2.64)$$

将式(2.64)与式(2.44)进行比较,可以清楚地看出,最优内插是Frost滤波器的一种特殊情况。

由式(2.62)得到

$$\frac{\mathbf{h}_{o,i}}{E_i} = \mathbf{R}_{yy}^{-1} \mathbf{c}_i \quad (2.65)$$

因此,  $\mathbf{R}_{yy}^{-1}$  的第  $(i+1)$  列为  $\mathbf{h}_{o,i}/E_i$ 。可以看出,  $\mathbf{R}_{yy}^{-1}$  能够进行以下因子分解<sup>[12]</sup>:

$$\mathbf{R}_{yy}^{-1} = \begin{bmatrix} 1 & -h_{o,1,0} & \cdots & -h_{o,L-1,0} \\ -h_{o,0,1} & 1 & \cdots & -h_{o,L-1,1} \\ \vdots & \vdots & \ddots & \vdots \\ -h_{o,0,L-1} & -h_{o,1,L-1} & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1/E_0 & 0 & \cdots & 0 \\ 0 & 1/E_1 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/E_{L-1} \end{bmatrix} = \mathbf{H}_o^T \mathbf{D}_e^{-1} \quad (2.66)$$

此外,由于  $\mathbf{R}_{yy}^{-1}$  为对称矩阵,式(2.66)可以表示为

$$\mathbf{R}_{yy}^{-1} = \begin{bmatrix} 1/E_0 & 0 & \cdots & 0 \\ 0 & 1/E_1 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/E_{L-1} \end{bmatrix} \begin{bmatrix} 1 & -h_{o,0,1} & \cdots & -h_{o,0,L-1} \\ -h_{o,1,0} & 1 & \cdots & -h_{o,1,L-1} \\ \vdots & \vdots & \ddots & \vdots \\ -h_{o,L-1,0} & -h_{o,L-1,1} & \cdots & 1 \end{bmatrix} = \mathbf{D}_e^{-1} \mathbf{H}_o \quad (2.67)$$

因此,可以推导出

$$\frac{h_{o,i,l}}{E_i} = \frac{h_{o,l,i}}{E_l}, \quad i, l = 0, 1, \dots, L \quad (2.68)$$

$\mathbf{R}_{yy}^{-1}$  的第一列和最后一列分别为前向和后向归一化预测系数,其余的列均为归一化内插系数。

下面分析内插系数对相关矩阵条件数的影响。矩阵  $\mathbf{R}_{yy}$  的条件数定义为<sup>[89]</sup>

$$\chi[\mathbf{R}_{yy}] = \|\mathbf{R}_{yy}\| \|\mathbf{R}_{yy}^{-1}\| \quad (2.69)$$

式中:  $\|\cdot\|$  表示矩阵的任一范数。需要注意  $\chi[\mathbf{R}_{yy}]$  与所采用的范数有关。这里采用 Frobenius 范数计算  $\chi[\mathbf{R}_{yy}]$ ,

$$\|\mathbf{R}_{yy}\|_F = \{\text{tr}[\mathbf{R}_{yy}^T \mathbf{R}_{yy}]\}^{1/2} = \{\text{tr}[\mathbf{R}_{yy}^2]\}^{1/2} \quad (2.70)$$

$$\|\mathbf{R}_{yy}^{-1}\|_F = \{\text{tr}[\mathbf{R}_{yy}^{-2}]\}^{1/2} \quad (2.71)$$

从式(2.65)得到

$$\frac{\mathbf{h}_{o,i}^T \mathbf{h}_{o,i}}{E_i^2} = \mathbf{c}_i^T \mathbf{R}_{yy}^{-2} \mathbf{c}_i \quad (2.72)$$

这意味着

$$\sum_{i=0}^{L-1} \frac{\mathbf{h}_{o,i}^T \mathbf{h}_{o,i}}{E_i^2} = \sum_{i=0}^{L-1} \mathbf{c}_i^T \mathbf{R}_{yy}^{-2} \mathbf{c}_i = \text{tr}[\mathbf{R}_{yy}^{-2}] \quad (2.73)$$

同时，容易验证

$$\text{tr}[\mathbf{R}_{yy}^2] = Lr_{yy}^2(0) + 2 \sum_{l=1}^{L-1} (L-l)r_{yy}^2(l) \quad (2.74)$$

式中： $r_{yy}(l), l = 0, 1, \dots, L-1$  为 Toeplitz 矩阵  $\mathbf{R}_{yy}$  的元素。因此，采用 Frobenius 范数时相关矩阵条件数的平方为

$$\chi_F^2[\mathbf{R}_{yy}] = \left[ Lr_{yy}^2(0) + 2 \sum_{l=1}^{L-1} (L-l)r_{yy}^2(l) \right] \sum_{i=0}^{L-1} \frac{\mathbf{h}_{o,i}^T \mathbf{h}_{o,i}}{E_i^2} \quad (2.75)$$

此外，关于前向预测系数与条件数的其他一些有意思的关系，可以在文献[17] 中找到。

基于 LCMV 滤波器的噪声抑制和语音去混响将在本书第 4 章、第 5 章和第 7 章进行研究。

## 2.4 卡尔曼滤波

卡尔曼滤波是针对非平稳信号<sup>[137-139]</sup>，在维纳滤波<sup>[234]</sup>基础上的一种自然的推广。卡尔曼滤波也是对被噪声污染的信号进行序贯（迭代）MMSE 估计。这里，信号通过状态模型来描述。

考虑一个观测信号

$$\begin{aligned} y(k) &= x(k) + v(k) \\ &= \mathbf{h}_1^T \mathbf{x}(k) + v(k) \end{aligned} \quad (2.76)$$

式中： $\mathbf{h}_1$  的定义如式 (2.9)； $\mathbf{x}(k)$  为长度为  $L$  的状态矢量； $v(k)$  为零均值高斯白噪声，且  $\sigma_v^2 = E[v^2(k)]$ 。需要注意的是，噪声的方差可以随时间变化。

假设语音信号可以表示为

$$x(k) = \sum_{l=1}^L a_l x(k-l) + v_x(k) \quad (2.77)$$

式中： $a_l, l = 1, 2, \dots, L$  可以视为信号  $x(k)$  的预测系数； $v_x(k)$  为与  $v(k)$  不相关的零均值高斯白噪声，且  $\sigma_{v_x}^2 = E[v_x^2(k)]$ 。式 (2.77) 称为状态模型。

利用状态矢量可以将式 (2.77) 重写为

$$\mathbf{x}(k) = A\mathbf{x}(k-1) + v_x(k)\mathbf{h}_1 \quad (2.78)$$

式中

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_{L-1} & a_L \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad (2.79)$$

为大小是  $L \times L$  的（非奇异）状态转移矩阵。

给定下列方程式：

$$\mathbf{x}(k) = A\mathbf{x}(k-1) + v_x(k)\mathbf{h}_1 \quad (2.80)$$

$$y(k) = \mathbf{h}_1^T \mathbf{x}(k) + v(k) \quad (2.81)$$

并假设  $A$ ,  $\sigma_v^2$  和  $\sigma_{v_x}^2$  均已知，则卡尔曼滤波的目标就是要找到  $\mathbf{x}(k)$  的最优线性 MMSE。这可以分两步实现。下面将借用文献[102]中巧妙而直观的方法来推导卡尔曼滤波。

令  $\hat{\mathbf{x}}(k|k-1)$  表示给定观测值  $y(1), y(2), \dots, y(k-1)$  的条件下在时刻  $k$  对  $\mathbf{x}(k)$  的线性 MMSE 估计。对应的状态估计误差为

$$\mathbf{e}(k|k-1) = \mathbf{x}(k) - \hat{\mathbf{x}}(k|k-1) \quad (2.82)$$

误差的协方差矩阵为

$$\mathbf{R}_{ee}(k|k-1) = E[\mathbf{e}(k|k-1)\mathbf{e}^T(k|k-1)] \quad (2.83)$$

第一步，不使用新的观测值，而是利用状态方程 (2.80) 预测  $\mathbf{x}(k)$ 。由于无法采用新的信息，因而最可能的估计为

$$\hat{\mathbf{x}}(k|k-1) = A\hat{\mathbf{x}}(k-1|k-1) \quad (2.84)$$

估计误差为

$$\begin{aligned} \mathbf{e}(k|k-1) &= \mathbf{x}(k) - \hat{\mathbf{x}}(k|k-1) \\ &= A\mathbf{x}(k-1) + v_x(k)\mathbf{h}_1 - A\hat{\mathbf{x}}(k-1|k-1) \\ &= A\mathbf{e}(k-1|k-1) + v_x(k)\mathbf{h}_1 \end{aligned} \quad (2.85)$$

假定  $E[\mathbf{e}(k-1|k-1)] = \mathbf{0}_{L \times 1}$ （这一零均值条件表明在最优线性估计中不存在固定偏差<sup>[82]</sup>），则  $E[\mathbf{e}(k|k-1)] = \mathbf{0}_{L \times 1}$ 。由于  $\mathbf{e}(k-1|k-1)$  与  $v_x(k)$  不相关，因而

$$\mathbf{R}_{ee}(k|k-1) = A\mathbf{R}_{ee}(k-1|k-1)A^T + \sigma_{v_x}^2(k)\mathbf{h}_1\mathbf{h}_1^T \quad (2.86)$$

这正是 Riccati 方程。

第二步，利用新的观测值  $y(k)$  来估计  $\mathbf{x}(k)$ 。基于  $\hat{\mathbf{x}}(k|k-1)$  和  $y(k)$  的线性估计具有如下形式：

$$\hat{\mathbf{x}}(k|k) = \mathbf{K}'(k)\hat{\mathbf{x}}(k|k-1) + \mathbf{k}(k)y(k) \quad (2.87)$$

式中：  $\mathbf{K}'(k)$  和  $\mathbf{k}(k)$  为待确定的矩阵和矢量。矢量  $\mathbf{k}(k)$  称为卡尔曼增益。此时，估计误差为

$$\begin{aligned}
\mathbf{e}(k|k) &= \mathbf{x}(k) - \hat{\mathbf{x}}(k|k) \\
&= \mathbf{x}(k) - \mathbf{K}'(k)\hat{\mathbf{x}}(k|k-1) - \mathbf{k}(k)y(k) \\
&= \mathbf{x}(k) - \mathbf{K}'(k)[\mathbf{x}(k) - \mathbf{e}(k|k-1)] - \mathbf{k}(k)[\mathbf{h}_1^T \mathbf{x}(k) + v(k)] \\
&= [\mathbf{I} - \mathbf{K}'(k) - \mathbf{k}(k)\mathbf{h}_1^T]\mathbf{x}(k) + \mathbf{K}'(k)\mathbf{e}(k|k-1) - \mathbf{k}(k)v(k)
\end{aligned} \tag{2.88}$$

由于  $E[\mathbf{e}(k|k-1)] = \mathbf{0}_{L \times 1}$ , 因而只有当

$$\mathbf{K}'(k) = \mathbf{I} - \mathbf{k}(k)\mathbf{h}_1^T \tag{2.89}$$

时,  $E[\mathbf{e}(k|k)] = \mathbf{0}_{L \times 1}$ 。基于这一约束条件, 可以得到

$$\begin{aligned}
\hat{\mathbf{x}}(k|k) &= [\mathbf{I} - \mathbf{k}(k)\mathbf{h}_1^T]\hat{\mathbf{x}}(k|k-1) + \mathbf{k}(k)y(k) \\
&= \hat{\mathbf{x}}(k|k-1) + \mathbf{k}(k)[y(k) - \mathbf{h}_1^T \hat{\mathbf{x}}(k|k-1)]
\end{aligned} \tag{2.90}$$

$$\begin{aligned}
\mathbf{e}(k|k) &= \mathbf{K}'(k)\mathbf{e}(k|k-1) - \mathbf{k}(k)v(k) \\
&= [\mathbf{I} - \mathbf{k}(k)\mathbf{h}_1^T]\mathbf{e}(k|k-1) - \mathbf{k}(k)v(k)
\end{aligned} \tag{2.91}$$

由于  $v(k)$  与  $v_x(k)$  和  $y(k-1)$  均不相关, 则  $v(k)$  与  $\mathbf{x}(k)$  和  $\hat{\mathbf{x}}(k|k-1)$  也不相关。因而  $E[\mathbf{e}(k|k-1)v(k)] = \mathbf{0}_{L \times 1}$ 。因此, 误差  $\mathbf{e}(k|k)$  的协方差矩阵为

$$\begin{aligned}
\mathbf{R}_{ee}(k|k) &= E[\mathbf{e}(k|k)\mathbf{e}^T(k|k)] \\
&= [\mathbf{I} - \mathbf{k}(k)\mathbf{h}_1^T]\mathbf{R}_{ee}(k|k-1)[\mathbf{I} - \mathbf{k}(k)\mathbf{h}_1^T]^T + \sigma_v^2(k)\mathbf{k}(k)\mathbf{k}^T(k)
\end{aligned} \tag{2.92}$$

最后, 需要找到最小化 MSE 的卡尔曼增益矢量  $\mathbf{k}(k)$ , 即

$$J(k) = \text{tr}[\mathbf{R}_{ee}(k|k)] \tag{2.93}$$

取  $J(k)$  关于  $\mathbf{k}(k)$  的微分, 得到

$$\frac{\partial J(k)}{\partial \mathbf{k}(k)} = -2[\mathbf{I} - \mathbf{k}(k)\mathbf{h}_1^T]\mathbf{R}_{ee}(k|k-1)\mathbf{h}_1 + 2\sigma_v^2(k)\mathbf{k}(k) \tag{2.94}$$

令式 (2.94) 等于 0, 则得到卡尔曼增益为

$$\mathbf{k}(k) = \frac{\mathbf{R}_{ee}(k|k-1)\mathbf{h}_1}{\mathbf{h}_1^T \mathbf{R}_{ee}(k|k-1)\mathbf{h}_1 + \sigma_v^2(k)} \tag{2.95}$$

误差协方差矩阵可以简化为

$$\begin{aligned}
\mathbf{R}_{ee}(k|k) &= [\mathbf{I} - \mathbf{k}(k)\mathbf{h}_1^T]\mathbf{R}_{ee}(k|k-1) - \\
&\quad \{[\mathbf{I} - \mathbf{k}(k)\mathbf{h}_1^T]\mathbf{R}_{ee}(k|k-1)\mathbf{h}_1 + \sigma_v^2(k)\mathbf{k}(k)\}\mathbf{k}^T(k)
\end{aligned} \tag{2.96}$$

根据式 (2.94), 式 (2.96) 中的第二项等于 0。因此

$$\mathbf{R}_{ee}(k|k) = [\mathbf{I} - \mathbf{k}(k)\mathbf{h}_1^T]\mathbf{R}_{ee}(k|k-1) \tag{2.97}$$

卡尔曼滤波的步骤总结如下<sup>[102]</sup>:

(1) 状态方程:

$$\mathbf{x}(k) = A\mathbf{x}(k-1) + v_x(k)\mathbf{h}_1$$

(2) 观测方程:

$$y(k) = \mathbf{h}_1^T \mathbf{x}(k) + v(k)$$

(3) 初始化:

$$\hat{\mathbf{x}}(0|0) = E[\mathbf{x}(0)]$$

$$\mathbf{R}_{ee}(0|0) = E[\mathbf{x}(0)\mathbf{x}^T(0)]$$

(4) 计算过程: 对于  $k=1, 2, \dots$

$$\hat{\mathbf{x}}(k|k-1) = A\hat{\mathbf{x}}(k-1|k-1)$$

$$\mathbf{R}_{ee}(k|k-1) = A\mathbf{R}_{ee}(k-1|k-1)A^T + \sigma_{v_x}^2(k)\mathbf{h}_1\mathbf{h}_1^T$$

$$k(k) = \frac{\mathbf{R}_{ee}(k|k-1)\mathbf{h}_1}{\mathbf{h}_1^T\mathbf{R}_{ee}(k|k-1)\mathbf{h}_1 + \sigma_v^2(k)}$$

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + k(k)[y(k) - \mathbf{h}_1^T\hat{\mathbf{x}}(k|k-1)]$$

$$\mathbf{R}_{ee}(k|k) = [\mathbf{I} - k(k)\mathbf{h}_1^T]\mathbf{R}_{ee}(k|k-1)$$

最后, 利用卡尔曼滤波对时刻  $k$  的语音样值  $x(k)$  做出的估计为

$$z_K(k) = \mathbf{h}_1^T\hat{\mathbf{x}}(k|k) \quad (2.98)$$

类似于维纳滤波, 这里定义卡尔曼滤波的语音失真指数为

$$v_{sd}(k) = \frac{E\{[x(k) - \mathbf{h}_1^T\hat{\mathbf{x}}(k|k)]^2\}}{\sigma_x^2(k)} = \frac{\mathbf{h}_1^T\mathbf{R}_{ee}(k|k)\mathbf{h}_1}{\sigma_x^2(k)} \quad (2.99)$$

当卡尔曼滤波收敛时,  $\mathbf{R}_{ee}(k|k)$  将越来越小, 因此  $v_{sd}(k)$  也逐渐减小。显然, 卡尔曼滤波能够得到比维纳滤波小得多的语音失真。

## 2.5 可替代 MSE 的准则

许多最优滤波器的性能, 例如信噪比性能等, 难以通过 MSE 表示出来。本节将给出一种基于 Pearson 相关系数 (Pearson Correlation Coefficient, PCC) 的新准则。平方 PCC (squared PCC, SPCC) 具有许多吸引人的性质, 可以将它作为优化代价函数。类似于 MSE, 基于这种新的代价函数能够推导出维纳滤波和许多其他最优或次优滤波器。相对于 MSE, 采用平方 PCC 的明显优势在于, 易于分析最优滤波器的性能 (尤其是输出信噪比)。

### 2.5.1 Pearson 相关系数

设  $x$  和  $y$  是两个均值为 0 的实随机变量。Pearson 系数 (PCC) 的定义为<sup>①</sup>[64, 181, 191]

$$\rho(x, y) = \frac{E[xy]}{\sigma_x \sigma_y} \quad (2.100)$$

式中:  $E[xy]$  为  $x$  和  $y$  的互相关系数;  $\sigma_x^2 = E[x^2]$  和  $\sigma_y^2 = E[y^2]$  分别为信号  $x$  和  $y$  的方差。为便于分析, 下面使用平方 Pearson 系数 (SPCC):

$$\rho^2(x, y) = \frac{E^2[xy]}{\sigma_x^2 \sigma_y^2} \quad (2.101)$$

<sup>①</sup> 这一相关系数以 Karl Pearson 的名字来命名, 他对该系数的诸多性质进行了阐述。

SPCC 的一个重要性质是

$$0 \leq \rho^2(x, y) \leq 1 \quad (2.102)$$

SPCC 表明了两个随机变量  $x$  和  $y$  之间线性相关性的程度。若  $\rho^2(x, y) = 0$ ，则称  $x$  和  $y$  不相关。 $\rho^2(x, y)$  的值越接近于 1，两个变量之间的相关性越强。如果两个变量相互独立，则  $\rho^2(x, y) = 0$ 。但反之则不一定成立，其原因在于 SPCC 仅表示两个变量  $x$  和  $y$  之间的线性相关性。而对于非线性相关性，SPCC 可能等于 0。但是，在  $x$  和  $y$  服从联合正态分布的特殊情况下，“独立”等价于“不相关”。

## 2.5.2 SPCC 的重要性质

这一小节讨论关于  $x, v, y$  和  $z$  四个信号之间的 SPCC 所具有的一系列性质。 $x(k)$  和  $y(k)$ （定义如式（2.1））之间的 SPCC 为

$$\rho^2(x, y) = \frac{\sigma_x^2}{\sigma_y^2} = \frac{\text{SNR}}{1 + \text{SNR}} \quad (2.103)$$

式中： $\sigma_y^2 = E[y^2(k)] = \sigma_x^2 + \sigma_v^2$  为信号  $y(k)$  的方差。

$x(k)$  和  $z(k)$ （定义如式（2.3））之间的 SPCC 为

$$\rho^2(x, z) = \frac{(\mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h})^2}{\sigma_x^2 (\mathbf{h}^T \mathbf{R}_{yy} \mathbf{h})} = \frac{(\mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h})^2}{\sigma_x^2 (\mathbf{h}^T \mathbf{R}_{xx} \mathbf{h})} \cdot \frac{\text{SNR}(\mathbf{h})}{1 + \text{SNR}(\mathbf{h})} \quad (2.104)$$

式中

$$\text{SNR}(\mathbf{h}) = \frac{\mathbf{h}^T \mathbf{R}_{xx} \mathbf{h}}{\mathbf{h}^T \mathbf{R}_{vv} \mathbf{h}} \quad (2.105)$$

为滤波器  $\mathbf{h}$  的输出信噪比。

### 性质 1

$$\rho^2(x, z) = \rho^2(x, \mathbf{h}^T y) = \rho^2(x, \mathbf{h}^T x) \rho^2(\mathbf{h}^T x, \mathbf{h}^T y) \quad (2.106)$$

式中

$$\rho^2(x, \mathbf{h}^T x) = \frac{(\mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h})^2}{\sigma_x^2 (\mathbf{h}^T \mathbf{R}_{xx} \mathbf{h})} \quad (2.107)$$

$$\rho^2(\mathbf{h}^T x, \mathbf{h}^T y) = \frac{\text{SNR}(\mathbf{h})}{1 + \text{SNR}(\mathbf{h})} \quad (2.108)$$

SPCC  $\rho^2(x, \mathbf{h}^T x)$  可以视为一种语音失真指数。若  $\mathbf{h} = \mathbf{h}_1$ （无语音失真），则  $\rho^2(x, \mathbf{h}^T x) = 1$ 。 $\rho^2(x, \mathbf{h}^T x)$  的值越接近于 0，语音信号的失真越大（除了简单的时延滤波器）。SPCC  $\rho^2(\mathbf{h}^T x, \mathbf{h}^T y)$  表明信噪比的改进步度，因此它可以被视为一种噪声抑制指数，当  $\text{SNR}(\mathbf{h})$  最大时， $\rho^2(\mathbf{h}^T x, \mathbf{h}^T y)$  取得最大值。

性质 1 在噪声抑制问题中具有基础性作用。它表明 SPCC  $\rho^2(x, \mathbf{h}^T y)$  是两个重要指数的乘积，这两个重要指数分别用于反映噪声抑制大小和语音失真程度，

下面将阐述 SPCC 实际上是一个代价函数。与之相比，MSE 具有更加复杂的形式，且它在语音增强中没有实际的物理意义。

### 性质 2

$$\rho^2(x, \mathbf{h}^\top y) \leq \frac{\text{SNR}(\mathbf{h})}{1 + \text{SNR}(\mathbf{h})} \quad (2.109)$$

当  $\mathbf{h} = \mathbf{h}_1$  时，等号成立。

### 性质 3

$$\rho^2(\mathbf{h}^\top x, y) = \rho^2(x, \mathbf{h}^\top x) \rho^2(x, y) \quad (2.110)$$

### 性质 4

$$\rho^2(\mathbf{h}^\top x, y) \leq \frac{\text{SNR}}{1 + \text{SNR}} \quad (2.111)$$

当  $\mathbf{h} = \mathbf{h}_1$  时，等号成立。

$v(k)$  和  $y(k)$ （定义如式（2.1））之间的 SPCC 为

$$\rho^2(v, y) = \frac{\sigma_v^2}{\sigma_y^2} = \frac{1}{1 + \text{SNR}} \quad (2.112)$$

$v(k)$  和  $z(k)$ （定义如式（2.3））之间的 SPCC 为

$$\rho^2(v, z) = \frac{(\mathbf{h}_1^\top \mathbf{R}_{vv} \mathbf{h})^2}{\sigma_v^2 (\mathbf{h}^\top \mathbf{R}_{yy} \mathbf{h})} = \frac{(\mathbf{h}_1^\top \mathbf{R}_{vv} \mathbf{h})^2}{\sigma_{vv}^2 (\mathbf{h}^\top \mathbf{R}_{vv} \mathbf{h})} \cdot \frac{1}{1 + \text{SNR}(\mathbf{h})} \quad (2.113)$$

### 性质 5

$$\rho^2(v, z) = \rho^2(v, \mathbf{h}^\top y) = \rho^2(v, \mathbf{h}^\top v) \cdot \rho^2(\mathbf{h}^\top v, \mathbf{h}^\top y) \quad (2.114)$$

式中

$$\rho^2(v, \mathbf{h}^\top v) = \frac{(\mathbf{h}_1^\top \mathbf{R}_{vv} \mathbf{h})^2}{\sigma_v^2 (\mathbf{h}^\top \mathbf{R}_{vv} \mathbf{h})} \quad (2.115)$$

$$\rho^2(\mathbf{h}^\top v, \mathbf{h}^\top y) = \frac{1}{1 + \text{SNR}(\mathbf{h})} \quad (2.116)$$

### 性质 6

$$\rho^2(v, \mathbf{h}^\top y) \leq \frac{1}{1 + \text{SNR}(\mathbf{h})} \quad (2.117)$$

当  $\mathbf{h} = \mathbf{h}_1$  时，等号成立。

### 性质 7

$$\rho^2(\mathbf{h}^\top v, y) = \rho^2(v, \mathbf{h}^\top v) \cdot \rho^2(x, y) \quad (2.118)$$

### 性质 8

$$\rho^2(\mathbf{h}^\top v, y) \leq \frac{1}{1 + \text{SNR}} \quad (2.119)$$

当  $\mathbf{h} = \mathbf{h}_1$  时，等号成立。

### 性质 9

$$\text{SNR}(\mathbf{h}) = \frac{\rho^2(\mathbf{h}^\top \mathbf{x}, \mathbf{h}^\top \mathbf{y})}{\rho^2(\mathbf{h}^\top \mathbf{v}, \mathbf{h}^\top \mathbf{y})} \quad (2.120)$$

2.5.3 节将会介绍，正量  $\rho^2(x, \mathbf{h}^\top y)$  可以作为一个用来推导不同最优滤波器的准则。这里列出的许多性质与此密切相关，并可以帮助读者更好地理解 SPCC 在语音增强应用中的根本作用。

### 2.5.3 由 SPCC 推导最优滤波器举例

直观来看，从观测信号  $y(k)$  中估计信号  $x(k)$  的问题可以表示为，找到一个能够最大化 SPCC  $\rho^2(x, \mathbf{h}^\top y)$  的滤波器，从而使得纯净语音信号  $x(k)$  与滤波器输出信号  $z(k)$  尽可能相关。此外，由于 SPCC  $\rho^2(x, \mathbf{h}^\top y)$  是其他两个 SPCC  $\rho^2(x, \mathbf{h}^\top x)$  和  $\rho^2(\mathbf{h}^\top x, \mathbf{h}^\top y)$  的乘积（见性质 1），因此通过最大化这两个 SPCC 中的任意一个（有约束或者无约束），还可以获得其他形式的最优滤波器。

#### 1. 语音无失真滤波器

正如 2.5.2 节所述，SPCC  $\rho^2(x, \mathbf{h}^\top x)$  是语音失真指数，且当  $\mathbf{h} = \mathbf{h}_1$  时，它取得最大值（等于 1）。因此，最大化  $\rho^2(x, \mathbf{h}^\top x)$  能够得到滤波器  $\mathbf{h}_1$ 。在这种情况下，

$$\text{SNR}(\mathbf{h}_1) = \text{SNR} \quad (2.121)$$

$$\rho^2(x, \mathbf{h}_1^\top x) = 1 \quad (2.122)$$

$$z_1(k) = y(k) \quad (2.123)$$

滤波器  $\mathbf{h}_1$  对于纯净语音和噪声都没有影响。换句话说， $\mathbf{h} = \mathbf{h}_1$  不会使纯净语音失真，但也不会改进输出信噪比。

#### 2. 最大信噪比滤波器

容易看出，最大化  $\rho^2(\mathbf{h}^\top x, \mathbf{h}^\top y)$  等价于最大化输出信噪比  $\text{SNR}(\mathbf{h})$ ，同时也等价于求解广义特征值问题：

$$\mathbf{R}_{xx} \mathbf{h} = \lambda \mathbf{R}_{yy} \mathbf{h} \quad (2.124)$$

假设  $\mathbf{R}_{yy}^{-1}$  存在，这一问题的最优解即为矩阵  $\mathbf{R}_{yy}^{-1} \mathbf{R}_{xx}$  的最大特征值  $\lambda_{\max}$  对应的特征矢量  $\mathbf{h}_{\max}$ ，因此

$$\text{SNR}(\mathbf{h}_{\max}) = \lambda_{\max} \quad (2.125)$$

$$\rho^2(\mathbf{h}_{\max}^\top x, \mathbf{h}_{\max}^\top y) = \frac{\lambda_{\max}}{1 + \lambda_{\max}} \quad (2.126)$$

$$z_{\max}(k) = \mathbf{h}_{\max}^\top y(k) \quad (2.127)$$

通过这一滤波器，还可以推导出 SPCC 的另外一个有趣的性质。

#### 性质 10

$$\rho^2(x, \mathbf{h}_{\max}^\top x) = \frac{\text{SNR}(\mathbf{h}_{\max})}{\text{SNR}} \cdot \rho^2(v, \mathbf{h}_{\max}^\top v) \quad (2.128)$$

由于  $\text{SNR}(\mathbf{h}_{\max}) \geq \text{SNR}(\mathbf{h}_1) = \text{SNR}$ ，这就意味着

$$\rho^2(x, \mathbf{h}_{\max}^T x) \geq \rho^2(v, \mathbf{h}_{\max}^T v) \quad (2.129)$$

也就是说，滤波器  $\mathbf{h}_{\max}$  对语音信号  $x(k)$  引入的失真，比其对噪声信号  $v(k)$  引入的失真小。

### 3. 维纳滤波器

下面将对 SPCC  $\rho^2(x, \mathbf{h}^T y)$  最大化。实际上，如果取  $\rho^2(x, \mathbf{h}^T y)$  关于  $\mathbf{h}$  的微分，并使结果等于 0，同时假设矩阵  $\mathbf{R}_{xx}$  和  $\mathbf{R}_{yy}$  满秩，则容易得到

$$\mathbf{R}_{xx} \mathbf{h}_1 (\mathbf{h}^T \mathbf{R}_{yy} \mathbf{h}) = (\mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h}) \mathbf{R}_{yy} \mathbf{h} \quad (2.130)$$

如果使最优滤波器  $\mathbf{h}_W$  满足以下关系：

$$\mathbf{h}_W^T \mathbf{R}_{yy} \mathbf{h}_W = \mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h}_W \quad (2.131)$$

则能够得到

$$\mathbf{h}_W = \mathbf{R}_{yy}^{-1} \mathbf{R}_{xx} \mathbf{h}_1 \quad (2.132)$$

式 (2.132) 即为式 (2.7) 给出的经典维纳滤波<sup>[125]</sup>。可以验证，式 (2.132) 给出的  $\mathbf{h}_W$  满足式 (2.131) 和式 (2.130)。维纳滤波具有下列性质：

#### 性质 11

最大化 SPCC  $\rho^2(x, \mathbf{h}^T y)$  等价于在  $\mathbf{h}^T \mathbf{R}_{yy} \mathbf{h} = \mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h}$  的约束条件下，最大化滤波器输出信号  $z(k)$  的方差  $E[z^2(k)]$ 。

#### 性质 12

$$\rho^2(x, \mathbf{h}_W^T y) = \frac{1}{\xi_{nr}(\mathbf{h}_W)} \cdot \frac{1 + \text{SNR}(\mathbf{h}_W)}{\text{SNR}} \quad (2.133)$$

这意味着

$$\xi_{nr}(\mathbf{h}_W) \geq \frac{1 + \text{SNR}(\mathbf{h}_W)}{\text{SNR}} \quad (2.134)$$

利用性质 2 和性质 12，可以推导一个更优的下限：

$$\xi_{nr}(\mathbf{h}_W) \geq \frac{[1 + \text{SNR}(\mathbf{h}_W)]^2}{\text{SNR} \cdot \text{SNR}(\mathbf{h}_W)} \geq \frac{1 + \text{SNR}(\mathbf{h}_W)}{\text{SNR}} \quad (2.135)$$

#### 性质 13（与 2.2 节给出的定理相同）

采用式 (2.132) 给出的最优维纳滤波器得到的输出信噪比始终大于或等于其输入信噪比。

证明：

给出  $y(k)$  与  $\mathbf{h}_W^T y(k)$  之间的 SPCC：

$$\rho^2(y, \mathbf{h}_W^T y) = \frac{(\mathbf{h}_1^T \mathbf{R}_{yy} \mathbf{h}_W)^2}{\sigma_y^2 (\mathbf{h}_W^T \mathbf{R}_{yy} \mathbf{h}_W)} = \frac{\sigma_x^2}{\sigma_y^2} \cdot \frac{\sigma_x^2}{\mathbf{h}_W^T \mathbf{R}_{xx} \mathbf{h}_1} = \frac{\rho^2(x, y)}{\rho^2(x, \mathbf{h}_W^T y)} \quad (2.136)$$

因此

$$\rho^2(x, y) = \rho^2(y, \mathbf{h}_W^T y) \cdot \rho^2(x, \mathbf{h}_W^T y) \leq \rho^2(x, \mathbf{h}_W^T y) \quad (2.137)$$

利用式(2.103)和2.5.2节中的性质2, 得到

$$\frac{\text{SNR}}{1+\text{SNR}} \leq \frac{\text{SNR}(\mathbf{h}_w)}{1+\text{SNR}(\mathbf{h}_w)} \quad (2.138)$$

式(2.138)重新组织后得到

$$\frac{1}{1+\frac{1}{\text{SNR}}} \leq \frac{1}{1+\frac{1}{\text{SNR}(\mathbf{h}_w)}} \quad (2.139)$$

也就意味着

$$\frac{1}{\text{SNR}} \geq \frac{1}{\text{SNR}(\mathbf{h}_w)} \quad (2.140)$$

因此

$$\text{SNR}(\mathbf{h}_w) \geq \text{SNR} \quad (2.141)$$

证毕。

上面的证明过程极其简单, 并且比2.2节中的证明过程更容易理解。

#### 性质14

$$\frac{[1+\text{SNR}(\mathbf{h}_w)]^2}{\text{SNR} \cdot \text{SNR}(\mathbf{h}_w)} \leq \xi_{\text{nr}}(\mathbf{h}_w) \leq \frac{(1+\text{SNR})[1+\text{SNR}(\mathbf{h}_w)]}{\text{SNR}^2} \quad (2.142)$$

或者

$$\frac{1}{\rho^2(\mathbf{h}_w^\top \mathbf{v}, \mathbf{h}_w^\top \mathbf{y}) \cdot \rho^2(\mathbf{h}_w^\top \mathbf{x}, \mathbf{h}_w^\top \mathbf{y})} \leq \text{SNR} \cdot \xi_{\text{nr}}(\mathbf{h}_w) \leq \frac{1}{\rho^2(\mathbf{x}, \mathbf{y}) \cdot \rho^2(\mathbf{h}_w^\top \mathbf{v}, \mathbf{h}_w^\top \mathbf{y})} \quad (2.143)$$

证明: 对于下限, 参见式(2.135)。利用性质12和式(2.137), 容易证明其上限。

#### 性质15

$$v_{\text{sd}}(\mathbf{h}_w) = 1 - \rho^2(\mathbf{x}, \mathbf{h}_w^\top \mathbf{x}) \cdot \left\{ 1 - \frac{1}{[1+\text{SNR}(\mathbf{h}_w)]^2} \right\} \quad (2.144)$$

式(2.144)表明了语音失真指数  $v_{\text{sd}}(\mathbf{h}_w)$  与SPCC  $\rho^2(\mathbf{x}, \mathbf{h}_w^\top \mathbf{x})$  之间的关系。当  $\rho^2(\mathbf{x}, \mathbf{h}_w^\top \mathbf{x})$  较大(或较小)时,  $v_{\text{sd}}(\mathbf{h}_w)$  较小(或较大), 因而纯净语音信号的失真较低(或较高)。同时,

$$\rho^2(\mathbf{x}, \mathbf{h}_w^\top \mathbf{x}) \geq \frac{\text{SNR}}{1+\text{SNR}} \cdot \frac{1+\text{SNR}(\mathbf{h}_w)}{\text{SNR}(\mathbf{h}_w)} \quad (2.145)$$

随着输出信噪比逐渐增加, SPCC  $\rho^2(\mathbf{x}, \mathbf{h}_w^\top \mathbf{x})$  的下限逐渐降低, 因而纯净语音的失真可能增加。

下面讨论最大化 SPCC 与最小化 MSE 的关系。MSE 定义为

$$\begin{aligned} J(\mathbf{h}) &= E[e^2(k)] \\ &= \sigma_x^2 + \mathbf{h}^\top \mathbf{R}_{yy} \mathbf{h} - 2\mathbf{h}_1^\top \mathbf{R}_{xx} \mathbf{h} \\ &= \sigma_x^2 \left[ 1 + \frac{1}{\xi_{nr}(\mathbf{h})} \cdot \frac{1 + \text{SNR}(\mathbf{h})}{\text{SNR}} - 2 \frac{\mathbf{h}^\top \mathbf{R}_{xx} \mathbf{h}}{\mathbf{h}_1^\top \mathbf{R}_{xx} \mathbf{h}} \cdot \rho^2(x, \mathbf{h}^\top x) \right] \end{aligned} \quad (2.146)$$

### 性质 16

$$\tilde{J}(\mathbf{h}_w) = \text{SNR}[1 - \rho^2(x, \mathbf{h}_w^\top y)] \quad (2.147)$$

式中:  $\tilde{J}(\mathbf{h}_w)$  为归一化 MMSE, 其定义如式 (2.16)。因此, 正如所预料的, 当 SPCC 最大化时, MSE 最小化。

**证明:** 在式(2.146)中利用性质 12、式(2.131)和性质 1, 容易验证式(2.147)。

### 性质 17

$$\frac{\text{SNR}}{1 + \text{SNR}(\mathbf{h}_w)} \leq \tilde{J}(\mathbf{h}_w) \leq \frac{\text{SNR}}{1 + \text{SNR}} \quad (2.148)$$

或

$$\rho^2(\mathbf{h}_w^\top v, \mathbf{h}_w^\top y) \leq \frac{\tilde{J}(\mathbf{h}_w)}{\text{SNR}} \leq \rho^2(v, y) \quad (2.149)$$

**证明:** 通过采用  $\rho^2(\mathbf{h}_w^\top v, \mathbf{h}_w^\top y)$  的上下界和式 (2.147) 可以证明这里的上下限。

### 性质 18

$$v_{sd}(\mathbf{h}_w) = \frac{1}{\text{SNR}} \left[ \tilde{J}(\mathbf{h}_w) - \frac{1}{\xi_{nr}(\mathbf{h}_w)} \right] \quad (2.150)$$

**证明:** 见文献[41]。

### 性质 19

$$\frac{1}{[1 + \text{SNR}(\mathbf{h}_w)]^2} \leq v_{sd}(\mathbf{h}_w) \leq \frac{1 + \text{SNR}(\mathbf{h}_w) - \text{SNR}}{(1 + \text{SNR})(1 + \text{SNR}(\mathbf{h}_w))} \quad (2.151)$$

或

$$\rho^4(\mathbf{h}_w^\top v, \mathbf{h}_w^\top y) \leq v_{sd}(\mathbf{h}_w) \leq \rho^2(v, y) \cdot \rho^2(\mathbf{h}_w^\top v, \mathbf{h}_w^\top y) + \rho^2(v, y) - \rho^2(\mathbf{h}_w^\top v, \mathbf{h}_w^\top y) \quad (2.152)$$

**证明:** 这里的上下限能够通过性质 14、17 和 18 证明。

### 性质 20

对于维纳滤波器, 从 MSE 的角度来看,

$$\text{SNR}(\mathbf{h}_w) \geq \text{SNR} \Leftrightarrow \xi_{nr}(\mathbf{h}_w) > 1, v_{sd}(\mathbf{h}_w) < 1 \quad (2.153)$$

因此, 除了一些特殊情况, 如  $\text{SNR}(\mathbf{h}_w) = \text{SNR}$  时,  $\xi_{nr}(\mathbf{h}_w)$  和  $v_{sd}(\mathbf{h}_w)$  这两个测度能够很好地指示维纳滤波器的性能。在  $\text{SNR}(\mathbf{h}_w) = \text{SNR}$  这种情况下,

$$\xi_{\text{nr}}(\mathbf{h}_w) = \frac{(1+\text{SNR})^2}{\text{SNR}^2} > 1 \quad (2.154)$$

$$v_{\text{sd}}(\mathbf{h}_w) = \frac{1}{(1+\text{SNR})^2} > 0 \quad (2.155)$$

$$\mathbf{h}_w = \frac{\text{SNR}}{1+\text{SNR}} \mathbf{h}_1 \quad (2.156)$$

在信号  $x(k)$  不可预测（随机白噪声信号）时，这种情况才会发生。这一特殊情况下，定义式 (2.19) 和式 (2.20) 会存在疑义，即可能存在噪声抑制和语音失真，但输出信噪比却没有改进。这是因为对于常数  $c \neq 0$  且  $c \neq 1$ ，

$$\xi_{\text{nr}}(c \cdot \mathbf{h}_w) \neq \xi_{\text{nr}}(\mathbf{h}_w) \quad (2.157)$$

$$v_{\text{sd}}(c \cdot \mathbf{h}_w) \neq v_{\text{sd}}(\mathbf{h}_w) \quad (2.158)$$

### 性质 21

对于维纳滤波器，从 SPCC 的角度来看，

$$\text{SNR}(\mathbf{h}_w) \geq \text{SNR} \Leftrightarrow \rho^2(\mathbf{h}_w^T \mathbf{x}, \mathbf{h}_w^T \mathbf{y}) \geq \rho^2(\mathbf{x}, \mathbf{y}), \rho^2(\mathbf{x}, \mathbf{h}_w^T \mathbf{x}) \leq 1 \quad (2.159)$$

当  $\text{SNR}(\mathbf{h}_w) = \text{SNR}$  时，

$$\rho^2(\mathbf{h}_w^T \mathbf{x}, \mathbf{h}_w^T \mathbf{y}) = \rho^2(\mathbf{x}, \mathbf{y}) \quad (2.160)$$

$$\rho^2(\mathbf{x}, \mathbf{h}_w^T \mathbf{x}) = 1 \quad (2.161)$$

此时，基于 SPCC  $\rho^2(\mathbf{h}_w^T \mathbf{x}, \mathbf{h}_w^T \mathbf{y})$  和  $\rho^2(\mathbf{x}, \mathbf{h}_w^T \mathbf{x})$  的测度准确反映了输出信噪比，因为当输出信噪比没有得到改进时，语音失真指数  $\rho^2(\mathbf{x}, \mathbf{h}_w^T \mathbf{x})$  表明确实不存在语音失真，而噪声抑制指数  $\rho^2(\mathbf{h}_w^T \mathbf{x}, \mathbf{h}_w^T \mathbf{y})$  也表明确实没有进行噪声抑制。采用 SPCC 时，上面所讨论的疑义不再存在，因为 SPCC 具有下列性质，即对于常数  $c \neq 0$ ，

$$\rho^2(c \cdot \mathbf{h}_w^T \mathbf{x}, c \cdot \mathbf{h}_w^T \mathbf{y}) = \rho^2(\mathbf{h}_w^T \mathbf{x}, \mathbf{h}_w^T \mathbf{y}) \quad (2.162)$$

$$\rho^2(\mathbf{x}, c \cdot \mathbf{h}_w^T \mathbf{x}) = \rho^2(\mathbf{x}, \mathbf{h}_w^T \mathbf{x}) \quad (2.163)$$

性质 20 和 21 从根本上表明，与基于 SPCC 并从 SPCC 准则推导出的等价测度相比，从 MSE 推导出的噪声抑制因子  $\xi_{\text{nr}}(\mathbf{h}_w)$  和语音失真指数  $v_{\text{sd}}(\mathbf{h}_w)$  会出现一些疑义。

### 4. 折中滤波器

除了上述滤波器以外，还能够推导出在语音失真和信噪比改进之间进行折中的其他最优滤波器。例如，找到一个最小化语音失真同时又保证一定程度的信噪比改进的滤波器是很有意义的。这一问题的数学表达式如下：

$$\max_{\mathbf{h}} \rho^2(\mathbf{x}, \mathbf{h}^T \mathbf{x}) \quad \text{s.t.} \quad \text{SNR}(\mathbf{h}) = \beta_1 \cdot \text{SNR} \quad (2.164)$$

式中： $\beta_1 > 1$ 。利用拉格朗日乘子  $\mu$  将这一约束与代价函数结合，式 (2.164)

可以重写为

$$\max_{\mathbf{h}} \mathcal{L}(\mathbf{h}, \mu) \quad (2.165)$$

式中

$$\mathcal{L}(\mathbf{h}, \mu) = \frac{(\mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h})^2}{\sigma_x^2 (\mathbf{h}^T \mathbf{R}_{vv} \mathbf{h})} + \mu \left( \frac{\mathbf{h}^T \mathbf{R}_{xx} \mathbf{h}}{\mathbf{h}^T \mathbf{R}_{vv} \mathbf{h}} - \beta_1 \cdot \text{SNR} \right) \quad (2.166)$$

取  $\mathcal{L}(\mathbf{h}, \mu)$  关于  $\mathbf{h}$  的梯度并使结果等于 0, 得到

$$\begin{aligned} & \frac{2\sigma_x^2 (\mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h})(\mathbf{h}^T \mathbf{R}_{xx} \mathbf{h}) \mathbf{R}_{xx} \mathbf{h}_1 - 2\sigma_x^2 (\mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h})^2 \mathbf{R}_{xx} \mathbf{h}}{(\sigma_x^2 \cdot \mathbf{h}^T \mathbf{R}_{xx} \mathbf{h})^2} + \\ & \mu \frac{2(\mathbf{h}^T \mathbf{R}_{vv} \mathbf{h}) \mathbf{R}_{xx} \mathbf{h} - 2(\mathbf{h}^T \mathbf{R}_{xx} \mathbf{h}) \mathbf{R}_{vv} \mathbf{h}}{(\mathbf{h}^T \mathbf{R}_{vv} \mathbf{h})^2} = 0_{L \times 1} \end{aligned} \quad (2.167)$$

寻找满足以下关系式的最优滤波器  $\mathbf{h}_T$ :

$$\mathbf{h}_1^T \mathbf{R}_{xx} \mathbf{h}_T = \mathbf{h}_T^T \mathbf{R}_{xx} \mathbf{h}_T \quad (2.168)$$

在这种情况下, 式 (2.167) 变为

$$\frac{\mathbf{R}_{xx} \mathbf{h}_1}{\sigma_x^2} - \frac{\mathbf{R}_{xx} \mathbf{h}_T}{\sigma_x^2} + \mu \frac{(\mathbf{h}_T^T \mathbf{R}_{vv} \mathbf{h}_T) \mathbf{R}_{xx} \mathbf{h}_T - (\mathbf{h}_T^T \mathbf{R}_{xx} \mathbf{h}_T) \mathbf{R}_{vv} \mathbf{h}_T}{(\mathbf{h}_T^T \mathbf{R}_{vv} \mathbf{h}_T)^2} = \mathbf{0}_{L \times 1} \quad (2.169)$$

在式 (2.169) 两边同时左乘  $\mathbf{h}_T^T$ , 则能够验证, 滤波器  $\mathbf{h}_T$  确实满足式 (2.168)。

对式 (2.169) 进行简单变换后发现

$$\mathbf{R}_{xx} \mathbf{h}_1 - \mathbf{R}_{xx} \mathbf{h}_T + \mu \text{SNR} \xi_{nr}(\mathbf{h}_T) \mathbf{R}_{xx} \mathbf{h}_T - \mu \beta_1 \text{SNR}^2 \xi_{nr}(\mathbf{h}_T) \mathbf{R}_{vv} \mathbf{h}_T = \mathbf{0}_{L \times 1} \quad (2.170)$$

定义下列参量:

$$\tilde{\mathbf{R}}_{xx} = \frac{\mathbf{R}_{xx}}{\sigma_x^2} \quad (2.171)$$

$$\tilde{\mathbf{R}}_{vv} = \frac{\mathbf{R}_{vv}}{\sigma_v^2} \quad (2.172)$$

$$\mu' = \mu \beta_1 \text{SNR}^2 \xi_{nr}(\mathbf{h}_T) \quad (2.173)$$

则能够得到最优的折中滤波器为

$$\mathbf{h}_T = \left[ \frac{\mu'}{\text{SNR}} \mathbf{I}_{L \times L} + \left( \mathbf{I}_{L \times L} - \frac{\mu'}{\beta_1 \text{SNR}} \right) \tilde{\mathbf{R}}_{vv}^{-1} \tilde{\mathbf{R}}_{xx} \right]^{-1} \tilde{\mathbf{R}}_{vv}^{-1} \tilde{\mathbf{R}}_{xx} \mathbf{h}_1 \quad (2.174)$$

式 (2.174) 可以与式 (2.11) 给出的维纳滤波进行比较。

滤波器  $\mathbf{h}_T$  与文献[59,69]中推导的滤波器目的相同。通过调整参数  $\mu'$  和  $\beta_1$ , 可以得到不同形式的折中滤波器。例如, 当  $\mu' = 0$  时, 得到语音无失真滤波器  $\mathbf{h}_T = \mathbf{h}_1$ , 当  $\mu' = 1$  且  $\beta_1 \rightarrow \infty$  时, 得到维纳滤波器  $\mathbf{h}_T = \mathbf{h}_W$ 。

通过最大化输出信噪比并使语音失真保持在一定的水平, 可以推导出另一种形式的折中滤波器。这一优化问题的数学表达式为

$$\max_{\mathbf{h}} \text{SNR}(\mathbf{h}) \quad \text{s.t.} \quad \rho^2(x, \mathbf{h}^T x) = \beta_2 \quad (2.175)$$

式中:  $\beta_2 < 1$ 。采用与式(2.164)的优化问题相同的步骤, 可以得到由式(2.175)推导出的折中滤波器:

$$\mathbf{h}_{T,2} = \left[ \frac{\mu''}{\text{SNR}} \mathbf{I}_{L \times L} + \left( \mathbf{I}_{L \times L} - \frac{\mu''}{\beta'_2 \text{SNR}} \right) \tilde{\mathbf{R}}_{vv}^{-1} \tilde{\mathbf{R}}_{xx} \right]^{-1} \tilde{\mathbf{R}}_{vv}^{-1} \tilde{\mathbf{R}}_{xx} \mathbf{h}_1 \quad (2.176)$$

式中

$$\beta'_2 = \beta_2 \xi_{nr}(\mathbf{h}_{T,2}) \quad (2.177)$$

$$\mu'' = \beta_2 \frac{[\text{SNR} \xi_{nr}(\mathbf{h}_{T,2})]^2}{\mu} \quad (2.178)$$

最优折中滤波器  $\mathbf{h}_T$  和  $\mathbf{h}_{T,2}$  具有相同的形式, 但后者在实际中很少使用, 其原因在于语音失真的程度很难控制。

## 2.6 总 结

最优滤波器在单麦克风和多麦克风噪声抑制中具有关键性作用。根据不同的应用环境, 可以推导不同的最优滤波器, 从而能够针对给定的问题获得满意的性能。

本章研究了三种重要的滤波器: 维纳滤波器、Frost 滤波器和卡尔曼滤波器。维纳滤波简单且非常实用, 但存在某些局限性。上面已经详细分析了这一最优滤波器对目标信号引入的失真。Frost 滤波是在维纳滤波的基础上增加了某些约束后得到的。本书的后面将会介绍, 如果合理地利用信号模型, Frost 滤波具有出色的性能。卡尔曼滤波可以被视为维纳滤波针对非平稳信号的一种推广, 其效果十分显著, 但必须要用到某些先验信息, 而这些先验信息在实际应用中通常难以得到。同时, 本章还引入了一种能够替代 MSE 的新准则 SPCC, 并且分析了如何利用该准则, 以及将它用于推导不同类型最优滤波器的内在原因。

# 第3章 传统波束形成技术

## 3.1 引言

波束形成技术具有悠久的发展历史，它已经在许多领域得到研究，如雷达、声纳、地震学、通信等。它可以应用于许多不同的方向，例如信号检测、波达方向估计（DOA）以及从被噪声、干扰源和混响污染的观测信号中增强目标信号。传统的波束形成可以描述为一个作用于传感器阵列输出的空间滤波器，以构造特定的波束方向图（方向特性）。这样一个空间滤波过程可以进一步分解为两个子过程：时间对齐和加权求和。时间对齐将每个传感器的输出延迟（或提前）适当的时间，使各个传感器接收的来自目标方向的信号成分能够在时间上同步。该步骤需要事先知道到达时间差（TDOA）。TDOA 能够从阵列观测信号中通过时延估计得到。加权求和就是要对时间对齐的信号进行加权，然后将加权结果加在一起形成一个输出。尽管这两个步骤在控制阵列波束模式中都具有重要作用（时间对齐控制着波束方向，加权求和控制着主瓣的波束宽度和旁瓣的特性），波束形成的重点通常集中于加权系数的确定。在许多应用中，加权系数可以根据事先确定的阵列波束模式来确定，但通常更好的方法是根据信号和噪声的特性以自适应的方式估计加权系数。

基于空间滤波的波束形成是针对单一频率就足以描述的窄带信号设计的。对于具有丰富频率成分的宽带语音而言，这些波束形成对不同频率具有不同的波束模式，并且波束带宽会随着频率的增加而减小。如果采用这样的波束形成，当波束指向与源入射角度不一致时，源信号会被低通滤波。此外，来自于与波束形成指向不同方向的噪声将在其整个频谱范围内非均匀地衰减，导致阵列输出中出现某些令人烦扰的现象。因此，必须研究具有不变响应的宽带波束形成技术。设计这样的宽带波束形成的常见方法是进行子带分解，并在不同频率上独立地设计窄带波束形成。这等价于对阵列输出进行空-时滤波，也就是众所周知的滤波-求和结构。宽带波束形成的核心问题则变为确定空-时滤波器的系数。

本章讨论在信号增强应用中传统波束形成潜在的基本思想（注意随着波束形成功能的不同，波束形成的基本原理有所变化。除了信号增强，波束形成的另一个主要应用是 DOA 估计，这将是第 9 章中的内容）。本章首先将简要讨论与采用单个传感器相比，采用阵列所具有的优势。然后将研究采用何种方法解

决窄带问题。尽管窄带技术不是针对语音处理而设计的，但这些技术构成了声学环境下更先进的宽带波束形成技术的基础，并且某些情况下能够对宽带信号取得较好结果。许多在窄带情况下发展起来的基本思想可以扩展到宽带情形。为了说明这一点，将进一步讨论麦克风阵列领域中更令人感兴趣的（响应不变）宽带波束形成的基本理论。

## 3.2 问题描述

在传感器阵列中，广泛采用的信号模型假设每个传播通道仅引入了时延和衰减。基于这一假设，在 $N$ 个传感器构成阵列的情形下，时刻 $k$ 的阵列输出表示为

$$\begin{aligned} y_n(k) &= \alpha_n s[k - t - \mathcal{F}_n(\tau)] + v_n(k) \\ &= x_n(k) + v_n(k), \quad n = 1, 2, \dots, N \end{aligned} \quad (3.1)$$

式中： $\alpha_n (n = 1, 2, \dots, N)$ 为由传播效应导致的衰减因子，其取值介于0和1之间； $s(k)$ 为未知的源信号（既可以是窄带也可以是宽带的）； $t$ 为从未知源到传感器1的传播时间； $v_n(k)$ 为第 $n$ 个传感器处的加性噪声； $\tau$ 为麦克风1和麦克风2之间的相对时延（更常见的称谓是到达时间差（TDOA））， $\mathcal{F}_n(\tau)$ 为麦克风1和麦克风 $n$ 之间的相对时延，且 $\mathcal{F}_1(\tau) = 0$ ， $\mathcal{F}_2(\tau) = \tau$ 。本章做一个关键假设，即 $\tau$ 和 $\mathcal{F}_n(\tau)$ 是已知或者可以估计得到的，并且源信号和噪声信号不相关。还假设式(3.1)中的所有信号都是零均值且平稳的。

通过对阵列观测值 $y_n(k)$ 进行处理，可以获得许多关于声源的有用信息，如其位置、频率等。然而，本章主要关注减小加性噪声 $v_n(k)$ 可能对目标信号的影响，从而改善信噪比(SNR)。不失一般性，将第一个传感器的接收信号作为参考信号。由此，本章的目标可以表述为恢复信号 $x_1(k) = \alpha_1 s(k - t)$ 。

## 3.3 延迟-求和技术

采用阵列在增强目标信号的同时抑制干扰噪声的优势可以通过延迟-求和(Delay-and-Sum, DS)波束形成来展示。这样一个波束形成由两个基本处理步骤组成<sup>[27,63,72,73,135,197,243]</sup>。第一步是根据每个传感器与参考传感器之间的TDOA，对每个传感器信号进行相应的时移。利用式(3.1)给出的信号模型，经过时移之后得到

$$\begin{aligned} y_{a,n}(k) &= y_n[k + \mathcal{F}_n(\tau)] \\ &= \alpha_n s(k - t) + v_{a,n}(k) \\ &= x_{a,n}(k) + v_{a,n}(k), \quad n = 1, 2, \dots, N \end{aligned} \quad (3.2)$$

式中

$$v_{a,n}(k) = v_n[k + \mathcal{F}_n(\tau)]$$

下标“a”表示时间对齐后的传感器信号。第二步是将时移后的信号相加，得到 DS 波束形成的输出为

$$\begin{aligned} z_{DS}(k) &= \frac{1}{N} \sum_{n=1}^N y_{a,n}(k) \\ &= \alpha_s s(k-t) + \frac{1}{N} v_s(k) \end{aligned} \quad (3.3)$$

式中

$$\begin{aligned} \alpha_s &= \frac{1}{N} \sum_{n=1}^N \alpha_n \\ v_s(k) &= \sum_{n=1}^N v_{a,n}(k) \\ &= \sum_{n=1}^N v_n[k + \mathcal{F}_n(\tau)] \end{aligned}$$

现在，可以分析 DS 波束形成的输入和输出信噪比。对于式 (3.1) 给出的信号模型，参考信号的输入信噪比为

$$\text{SNR} = \frac{\sigma_{x_1}^2}{\sigma_{v_1}^2} = \alpha_1^2 \frac{\sigma_s^2}{\sigma_{v_1}^2} \quad (3.4)$$

式中： $\sigma_{x_1}^2 = E[x_1^2(k)]$ ， $\sigma_{v_1}^2 = E[v_1^2(k)]$ ， $\sigma_s^2 = E[s^2(k)]$  分别为信号  $x_1(k)$ ， $v_1(k)$ ， $s(k)$  的方差。经过 DS 处理之后，输出信噪比可以表示为式 (3.3) 右边第一项和第二项的方差之间的比值：

$$\begin{aligned} \text{oSNR} &= N^2 \alpha_s^2 \frac{E[s^2(k-t)]}{E[v_s^2(k)]} \\ &= N^2 \alpha_s^2 \frac{\sigma_s^2}{\sigma_{v_s}^2} \\ &= \left( \sum_{n=1}^N \alpha_n \right)^2 \frac{\sigma_s^2}{\sigma_{v_s}^2} \end{aligned} \quad (3.5)$$

式中

$$\begin{aligned} \sigma_{v_s}^2 &= E \left\{ \left[ \sum_{n=1}^N v_n[k + \mathcal{F}_n(\tau)] \right]^2 \right\} \\ &= \sum_{n=1}^N \sigma_{v_n}^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \varrho_{v_i v_j} \end{aligned} \quad (3.6)$$

式中： $\sigma_{v_n}^2 = E[v_n^2(k)]$  为噪声信号  $v_n(k)$  的方差； $\varrho_{v_i v_j} = E\{v_i[k + \mathcal{F}_i(\tau)]v_j[k + \mathcal{F}_j(\tau)]\}$  为

$v_i(k)$  和  $v_j(k)$  之间的互相关函数。

只有当

$$oSNR > SNR \quad (3.7)$$

时, DS 波束形成才有意义。这意味着信号  $z_{DS}(k)$  比任意麦克风输出信号  $y_n(k)$  都具有更少的噪声, 有可能成为  $x_1(k)$  的一个不错的近似。

### 特殊情况 1:

在这一特殊情况下, 假设麦克风处的噪声信号不相关, 即  $\rho_{v_iv_j} = 0$ ,  $\forall i, j = 1, 2, \dots, N, i \neq j$ , 并且它们都具有相同的方差, 即  $\sigma_{v_1}^2 = \sigma_{v_2}^2 = \dots = \sigma_{v_N}^2$ 。还假设所有的衰减因子都等于 1 (即  $\alpha_n = 1, \forall n$ )。这样可以得到

$$oSNR = N \cdot SNR \quad (3.8)$$

从中可以看出, 在前述条件下, 对传感器输出进行简单的时移和相加操作, 使得 SNR 的提高倍数等于传感器数目的个数。

### 特殊情况 2:

这里仅假设噪声信号具有相同的能量, 并且所有的衰减因子都等于 1。在这种情况下

$$oSNR = \frac{N}{1 + \rho_s} \cdot SNR \quad (3.9)$$

式中

$$\rho_s = \frac{2}{N} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \rho_{v_iv_j}$$

$$\rho_{v_iv_j} = \frac{\rho_{v_iv_j}}{\sigma_{v_i} \sigma_{v_j}}$$

$\rho_{v_iv_j}$  是相关系数且  $|\rho_{v_iv_j}| \leq 1$ 。通常这一系数的取值范围介于 -1 和 1 之间。如果麦克风处的噪声信号完全相关, 即  $\rho_{v_iv_j} = 1$ , 则  $\sum_{i=1}^{N-1} \sum_{j=i+1}^N \rho_{v_iv_j} = N(N-1)/2$ 。在这种情况下,  $oSNR = SNR$ 。这样, 采用 DS 技术不可能获得任何增益。随着相关系数  $\rho_{v_iv_j}$  的取值从 1 降低到 0, SNR 的增益逐渐增加。在某些情况下, 相关系数  $\rho_{v_iv_j}$  可能为负数。这可能在噪声信号来自于同一个点声源时发生。在这种情况下, 可能得到高于  $N$ , 甚至无穷大的信噪比增益。

讨论 DS 波束形成性能的另一种方式是通过研究波束方向图 (波束方向图也称为方向性图或者空间图)<sup>[217]</sup>, 它给出了阵列系统输入-输出性能的完整描述。从前面的分析易于看出, DS 波束形成实际上是一个  $N$  点的空间滤波器, 其波束方向图定义为该空间滤波器方向响应的幅度。利用式 (3.2) 和式 (3.3), 可以得到空间滤波器的第  $n$  个系数为  $\frac{1}{N} e^{j2\pi f \mathcal{F}_n(\tau)}$ , 其中  $f$  表示频率。该滤波器的方向响应可以通过傅里叶变换后得到。由于  $\mathcal{F}_n(\tau)$  同时依赖于阵列几何形状和

声源位置，因此，DS 波束形成的波束模式应该是阵列几何形状和声源位置的函数。此外，波束模式还是传感器数目和信号频率的函数。现在假设有一个由  $N$  个全向传感器组成的均匀线性阵列，如图 3.1 所示。若将两个相邻传感器之间的距离表示为  $d$ ，并假设声源位于远场，波束到达阵列的入射角度为  $\theta$ ，第  $n$  个传感器与参考传感器之间的 TDOA 可以写为

$$\mathcal{F}_n(\tau) = (n-1)\tau = (n-1)d \cos(\theta)/c \quad (3.10)$$

式中： $c$  表示空气中的声速。

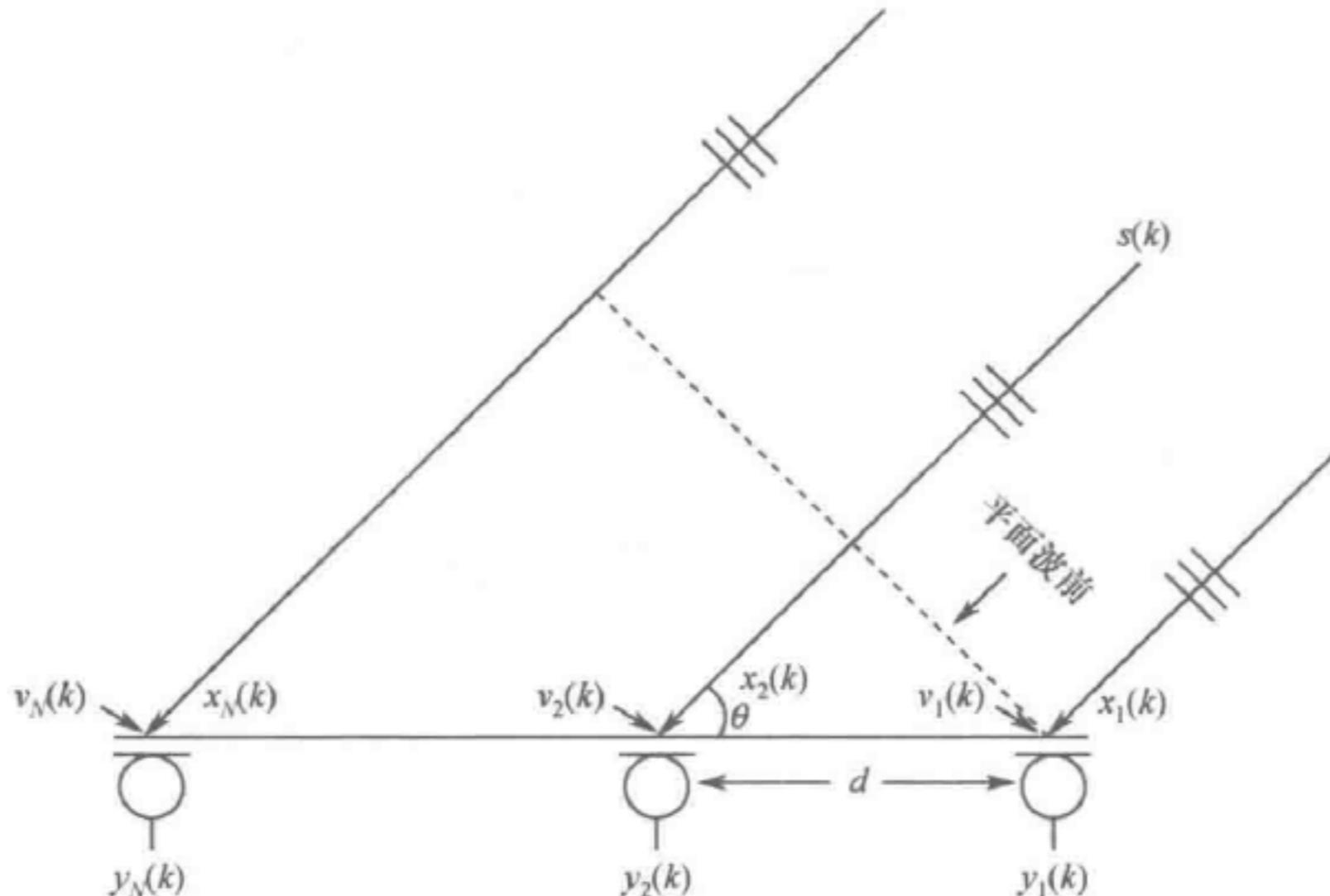


图 3.1 一个均匀线性阵列的示意图，其中声源  $s(k)$  位于远场，入射角度为  $\theta$ ，相邻两个麦克风之间的距离为  $d$

在这种情况下，DS 滤波器的方向响应，也就是滤波器的空间傅里叶变换<sup>[8,90]</sup>，可以表示为

$$\begin{aligned} S_{\text{DS}}(\psi, \theta) &= \frac{1}{N} \sum_{n=1}^N \left[ e^{j2\pi(n-1)fd \cos(\theta)/c} \right] e^{-j2\pi(n-1)fd \cos(\psi)/c} \\ &= \frac{1}{N} \sum_{n=1}^N e^{-j2\pi(n-1)fd[\cos(\psi) - \cos(\theta)]/c} \end{aligned} \quad (3.11)$$

式中： $\psi (0 \leq \psi \leq \pi)$  为方位角。波束模式表示为

$$\begin{aligned} \mathcal{A}_{\text{DS}}(\psi, \theta) &= |S_{\text{DS}}(\psi, \theta)| \\ &= \left| \frac{\sin[N\pi fd(\cos \psi - \cos \theta)/c]}{N \sin[\pi fd(\cos \psi - \cos \theta)/c]} \right| \end{aligned} \quad (3.12)$$

图 3.2 绘出了由 10 个传感器组成的均匀线性阵列的波束模式，其中  $d = 8\text{cm}$ ， $\theta = 90^\circ$ ， $f = 2\text{kHz}$ 。它总共由 9 个波束构成（一般在  $[0^\circ, 180^\circ]$  范围内的波束数目等于  $N-1$ ）。具有最大幅度的波束称为主瓣，其余的都称为旁瓣。有关主瓣的一个重要参数是波束宽度（主瓣宽度），其定义为主瓣两边的第一个过零点之间的范围。对于上面的线性阵列，容易计算得到波束宽度为

$2\arccos[c/(Ndf)]$ 。这一数值将随着传感器数目、相邻传感器之间的距离和信号频率的增加而减小。旁瓣的高度表示对来自目标方向以外的噪声和干扰声源的增益模式。在阵列和波束形成设计中，希望尽可能地使旁瓣降低，从而使得来自目标方向以外的信号能够被尽可能地衰减。此外，采用长度为  $N$  的空间滤波器，将始终存在  $N-1$  个零点。通过设计加权系数，可以使得这些零点对准干扰声源的方向。这与自适应波束形成技术有关，将在后面的章节进行更详细地讲述。

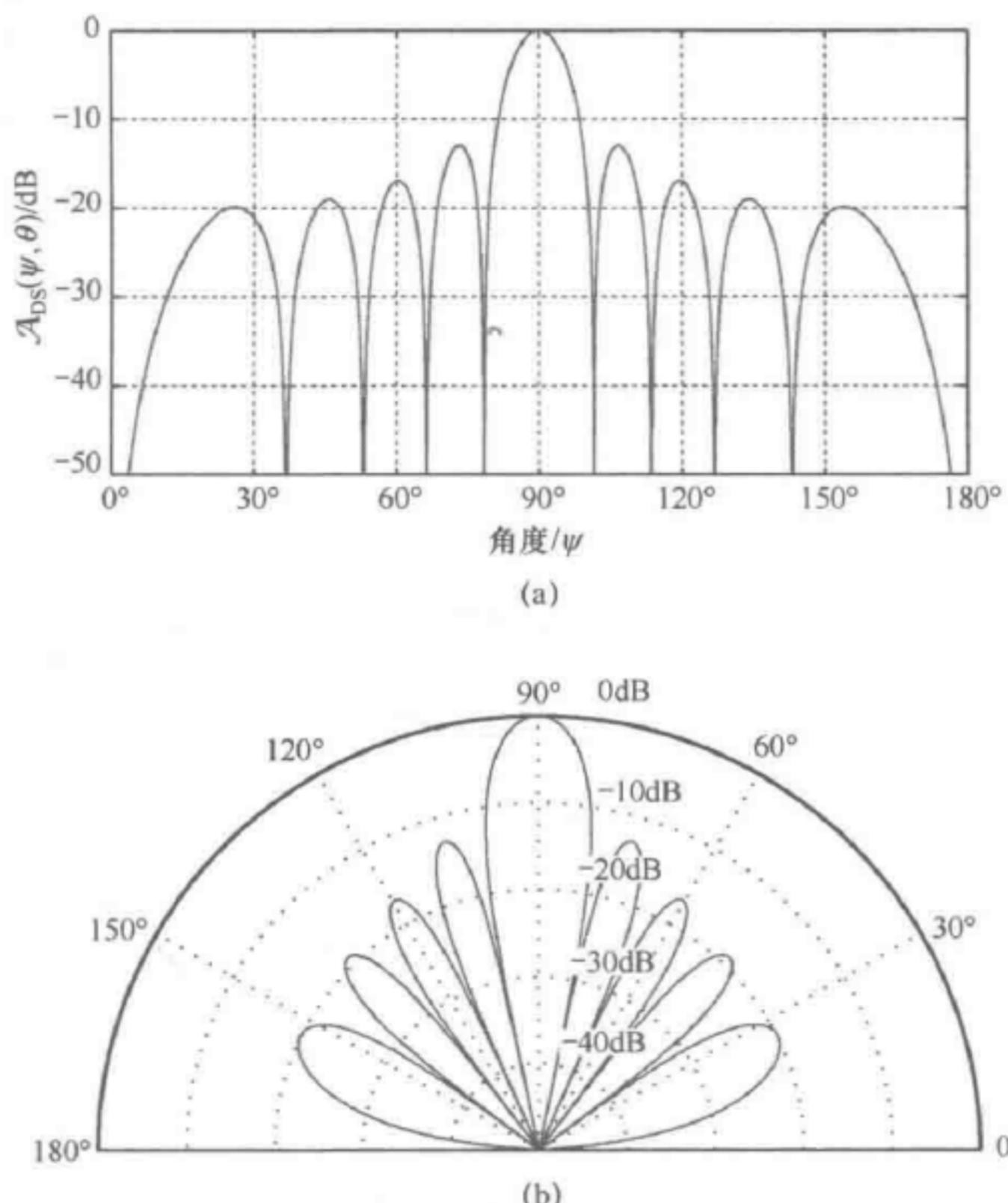


图 3.2 一个具有 10 个传感器的阵列在  $\theta = 90^\circ$ ,  $d = 8\text{cm}$ ,  $f = 2\text{kHz}$  时的波束模式  
(a) 用笛卡儿坐标表示; (b) 用极坐标表示。

在结束本节之前，将指出一个和传感器间距相关的潜在问题。从前面的分析看出，阵列波束宽度随着间距  $d$  的增加而减小。因此，如果想要一个更尖锐的波束，可以简单地增加间距  $d$ ，这将导致更大的阵列孔径。通常这将带来更多的噪声抑制。因此，在阵列设计中，期望这一间距尽可能大。然而，当  $d$  大于  $\lambda/2 = c/(2f)$ ，其中  $\lambda$  为信号的波长，会发生空间混叠。为了展示这一问题，图 3.2(b) 绘出了所采用的均匀线性阵列的波束模式。信号频率  $f$  仍然为 2kHz。但这里的阵列间距为 24cm。对应的波束模式如图 3.3 所示。这里可以看到三个具有最大幅度为 1 的波束。主瓣外的另外两个称为栅瓣。由栅瓣方向传播来的信号与主瓣方向传播来的信号将是不可分辨的。这种模糊性通常称为空间混叠。为了避免空间混叠，阵列间距必须满足  $d \leq \frac{\lambda}{2} = \frac{c}{2f}$ 。类似于奈奎斯特采样理论，

这一结果可以解释为空间采样定理。

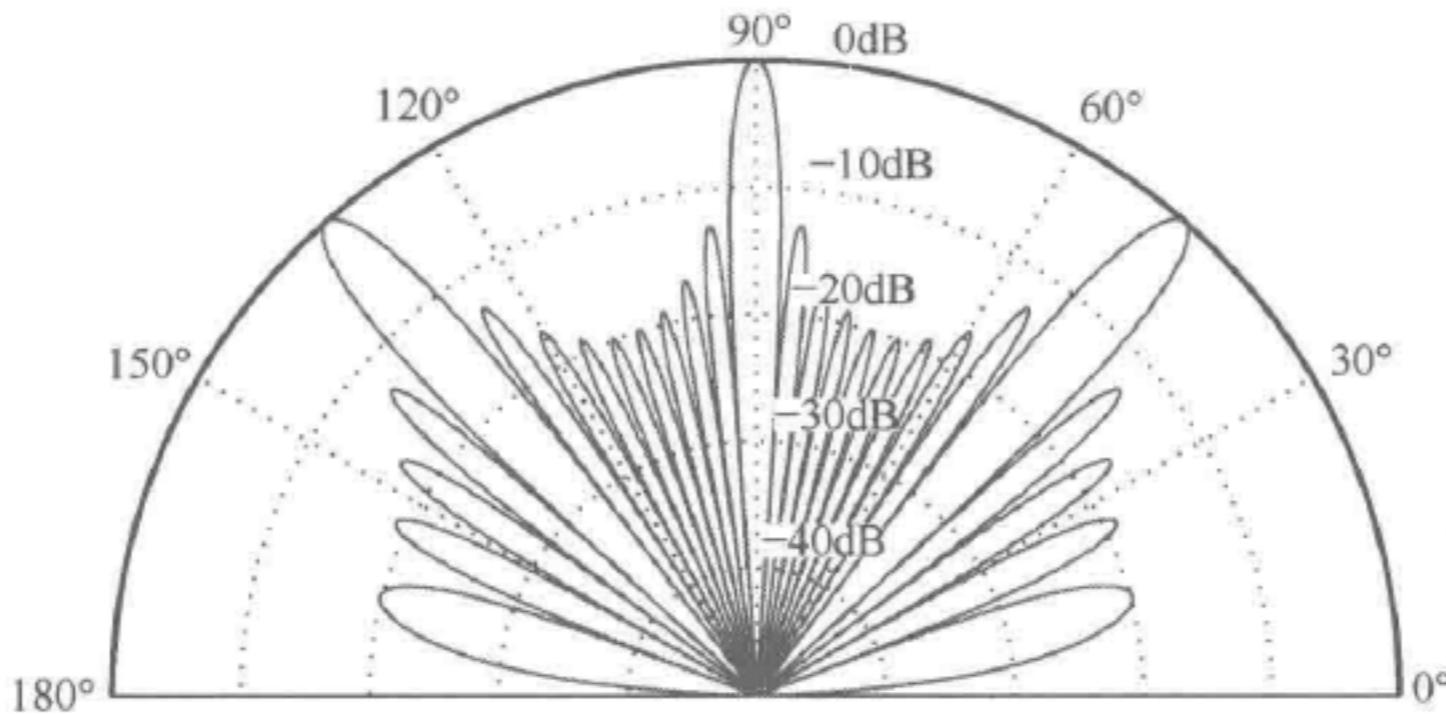


图 3.3 一个具有 10 个传感器的阵列在  $\theta = 90^\circ$ ,  $d = 24\text{cm}$ ,  $f = 2\text{kHz}$  时的波束方向图（用极坐标表示）

### 3.4 固定波束形成的设计

从前面的讨论可以看出，一旦阵列的几何形状固定，并且目标对准方向确定，DS 波束形成的波束模式特性，包括波束宽度、旁瓣幅度、零点位置都将固定。这意味着如果想要调整波束模式，必须对阵列几何形状进行物理调整，显然，对于已经固化的阵列系统这是不可能的。此时会产生一个自然的想法：能否在不改变阵列几何形状的条件下通过信号处理技术调整其波束模式从而改进其性能？本节将尝试回答这一问题，并讨论一系列被称为固定波束形成的技术，这些技术考虑了阵列几何形状，但假设无法得到任何与声源和噪声信号有关的信息。

重新回顾 DS 波束形成，容易看出，其内在的思路是将一个长度为  $N$  的空间滤波器应用于传感器输出。这与采用有限冲激响应（FIR）滤波器实现时间滤波的思路类似。因此，所有用于设计 FIR 滤波器的技术，包括加窗和最优逼近方法<sup>[177]</sup>，都可以在这里得到应用。为了说明如何设计一个波束形成以获得想要的波束模式，考虑最小二乘（Least-Squares, LS）这种广泛使用的最优逼近方法。

假设  $\mathbf{h} = [h_1 \ h_2 \ \dots \ h_N]^\top$  是一个长度为  $N$  的波束形成滤波器，对应的方向响应为

$$S(\psi) = \sum_{n=1}^N h_n e^{-j2\pi f \mathcal{F}_n[\tau(\psi)]} = \mathbf{h}^\top \boldsymbol{\varsigma}(\psi) \quad (3.13)$$

式中

$$\boldsymbol{\varsigma}(\psi) = \left[ e^{-j2\pi f \mathcal{F}_1[\tau(\psi)]} \ e^{-j2\pi f \mathcal{F}_2[\tau(\psi)]} \ \dots \ e^{-j2\pi f \mathcal{F}_N[\tau(\psi)]} \right]^\top$$

LS 方法的目标是优化滤波器系数  $h_n (n=1, 2, \dots, N)$ ，从而使得到的方向响应能够最好地逼近设定的方向响应。为了达到这一目的，首先定义 LS 的逼近准则：

$$\varepsilon^2 = \int_0^\pi \vartheta(\psi) |S(\psi) - S_d(\psi)|^2 d\psi \quad (3.14)$$

式中:  $S_d(\psi)$  表示目标方向响应;  $\vartheta(\psi)$  为正的加权实函数, 用于加重或去加重特定角度的重要性。

将式 (3.13) 代入式 (3.14), 可以将 LS 逼近准则重写为

$$\varepsilon^2 = \mathbf{h}^T \mathbf{Q} \mathbf{h} - 2 \mathbf{h}^T \mathbf{p} + \int_0^\pi \vartheta(\psi) |S_d(\psi)|^2 d\psi \quad (3.15)$$

式中

$$\begin{aligned} \mathbf{Q} &= \int_0^\pi \vartheta(\psi) \boldsymbol{\varsigma}(\psi) \boldsymbol{\varsigma}^H(\psi) d\psi \\ \mathbf{p} &= \int_0^\pi \vartheta(\psi) \operatorname{Re}[\boldsymbol{\varsigma}(\psi) S_d(\psi)] d\psi \end{aligned} \quad (3.16)$$

$\operatorname{Re}(\cdot)$  表示取实部; 上标 “H” 表示矢量或者矩阵的共轭转置。

求解  $\varepsilon^2$  关于  $\mathbf{h}$  的偏微分, 并使结果等于零, 得到

$$\mathbf{h}_{LS} = \mathbf{Q}^{-1} \mathbf{p} \quad (3.17)$$

注意到矩阵  $\mathbf{Q}$  是  $\mathcal{F}_N[\tau(\psi)]$  的函数, 矢量  $\mathbf{p}$  是  $\mathcal{F}_N[\tau(\psi)]$  和  $S_d(\psi)$  的二元函数。因此, LS 波束形成滤波器同时依赖于阵列几何形状和目标方向响应。

现在考虑与 3.3 节中所使用的均匀线性阵列相同的情况。假设已知声源位于某一特定区域(介于角度  $\psi_1$  和  $\psi_2$  之间), 但是没有声源入射角度的准确信息。因此, 想要设计一个波束形成, 使得入射角度位于  $\psi_1$  和  $\psi_2$  之间的信号能够通过, 而使来自其他方向的信号衰减。在这种情况下, 希望在数学上获得如下目标方向响应:

$$S_d(\psi) = \begin{cases} 1, & \psi_1 \leq \psi \leq \psi_2 \\ 0, & \text{其他} \end{cases} \quad (3.18)$$

如果假设所有的角度都同等重要, 即  $\vartheta(\psi)=1$ , 则

$$\begin{aligned} \mathbf{Q} &= \int_0^\pi \boldsymbol{\varsigma}(\psi) \boldsymbol{\varsigma}^H(\psi) d\psi \\ &= \begin{bmatrix} \int_0^\pi 1 d\psi & \int_0^\pi e^{j\tilde{d}_1 \cos \psi} d\psi & \dots & \int_0^\pi e^{j\tilde{d}_{N-1} \cos \psi} d\psi \\ \int_0^\pi e^{-j\tilde{d}_1 \cos \psi} d\psi & \int_0^\pi 1 d\psi & \dots & \int_0^\pi e^{j\tilde{d}_{N-2} \cos \psi} d\psi \\ \vdots & \vdots & \ddots & \vdots \\ \int_0^\pi e^{-j\tilde{d}_{N-1} \cos \psi} d\psi & \dots & \dots & \int_0^\pi 1 d\psi \end{bmatrix} \\ &= \begin{bmatrix} \pi & \int_0^\pi \cos(\tilde{d}_1 \cos \psi) d\psi & \dots & \int_0^\pi \cos(\tilde{d}_{N-1} \cos \psi) d\psi \\ \int_0^\pi \cos(\tilde{d}_1 \cos \psi) d\psi & \pi & \dots & \int_0^\pi \cos(\tilde{d}_{N-2} \cos \psi) d\psi \\ \vdots & \vdots & \ddots & \vdots \\ \int_0^\pi \cos(\tilde{d}_{N-1} \cos \psi) d\psi & \dots & \dots & \pi \end{bmatrix} \end{aligned} \quad (3.19)$$

$$p = \begin{bmatrix} \int_{\psi_1}^{\psi_2} 1 d\psi \\ \int_{\psi_1}^{\psi_2} \cos(\tilde{d}_1 \cos \psi) d\psi \\ \vdots \\ \int_{\psi_1}^{\psi_2} \cos(\tilde{d}_{N-1} \cos \psi) d\psi \end{bmatrix} \quad (3.20)$$

式中:  $\tilde{d}_n = 2\pi n f d / c, n=1, 2, \dots, N-1$ 。

式 (3.19) 和式 (3.20) 中的积分看起来难以计算, 但是肯定可以采用数值方法计算得到。这里考虑两个设计案例。在第一个案例中, 考虑声源可能随着时间在  $60^\circ \sim 120^\circ$  范围内移动。为了不使源信号失真, 希望波束形成具有大的波束宽度, 能够覆盖  $60^\circ \sim 120^\circ$  的范围。图 3.4 给出了这样一个采用 LS 技术的波束形成设计。正如图 3.4 所示, 其主瓣比 DS 波束形成的主瓣宽得多。

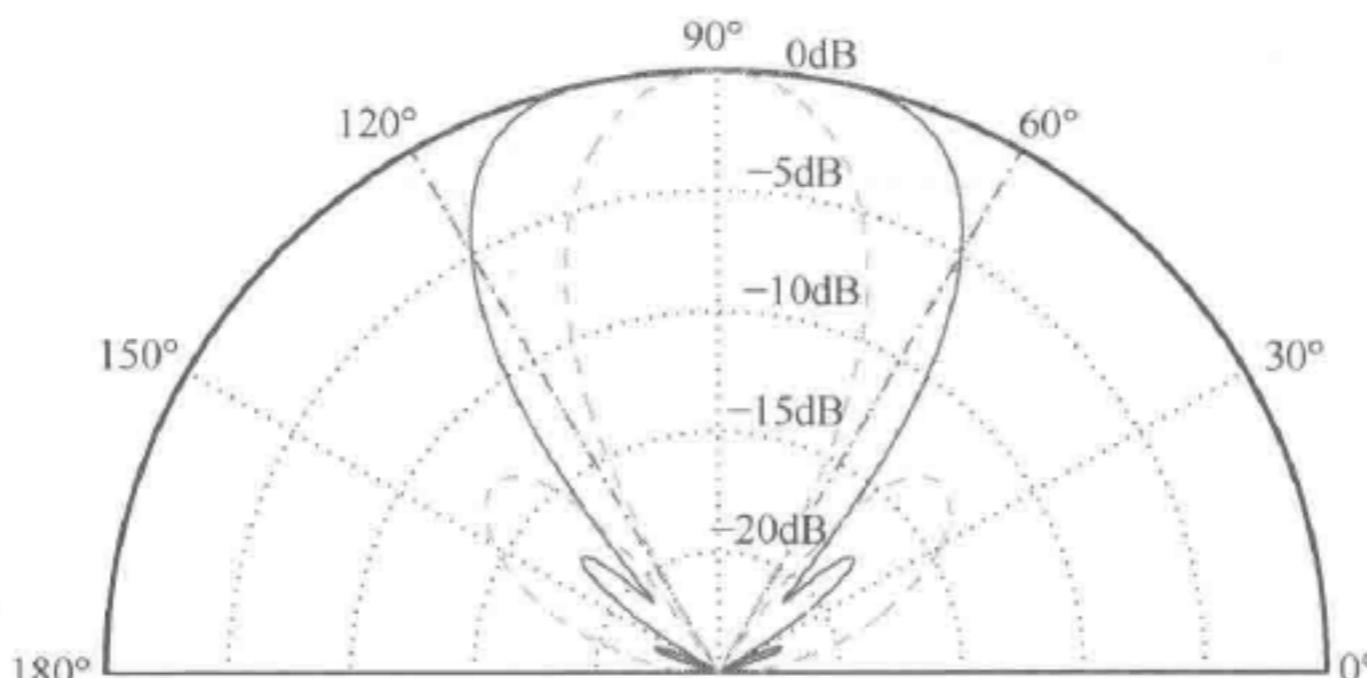


图 3.4 采用 LS 技术设计的波束方向图 (实线), 阵列为具有 10 个传感器的均匀线性阵列:  $d = 4\text{cm}$ ,  $f = 1.5\text{kHz}$ ,  $\psi_1 = 60^\circ$ ,  $\psi_2 = 120^\circ$ 。作为比较, 同时还给出了 DS 波束方向图 (虚线) 和目标波束方向图 (虚点线)

在第二个案例中, 假设已知声源位于  $90^\circ$  方向, 其误差小于  $\pm 5^\circ$ 。这里希望波束宽度更窄以抑制更多干扰。采用 LS 方法得到的对应波束模式如图 3.5 所示。可以看出, 这里的波束宽度比 DS 波束形成要窄得多。

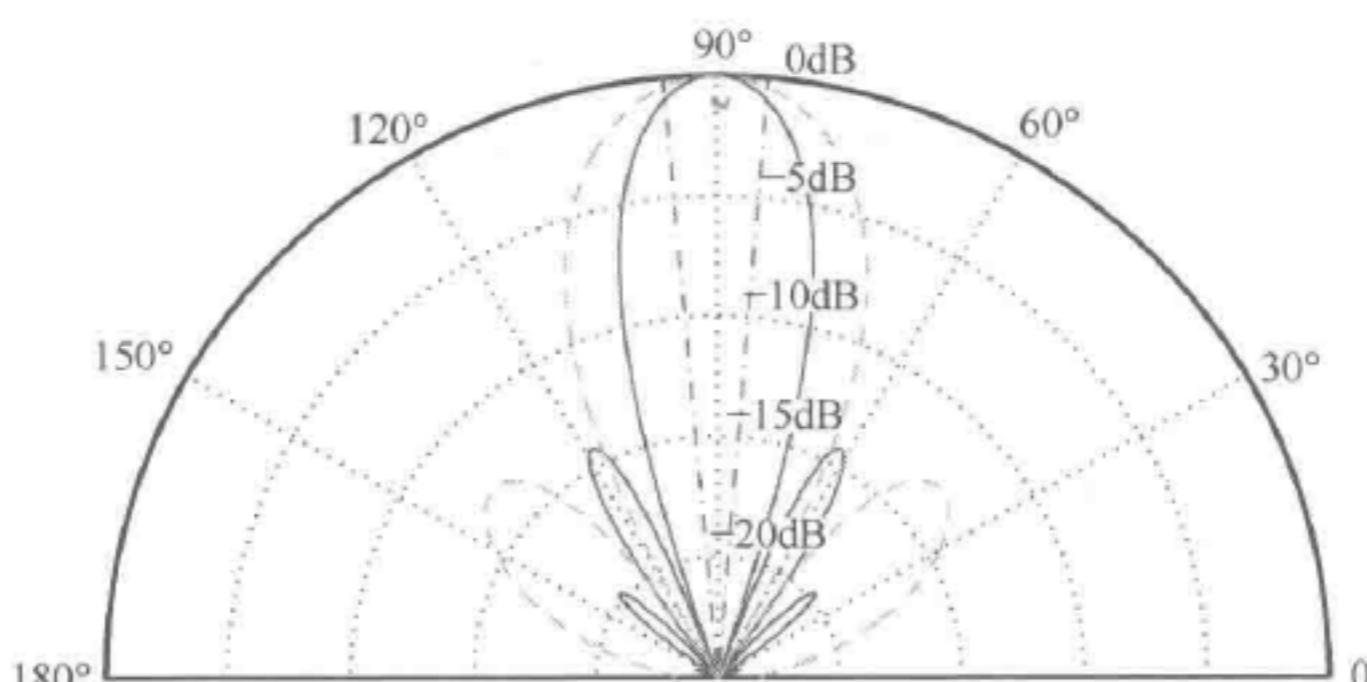


图 3.5 采用 LS 技术设计的波束方向图 (实线), 阵列为具有 10 个传感器的均匀线性阵列:  $d = 4\text{cm}$ ,  $f = 1.5\text{kHz}$ ,  $\psi_1 = 85^\circ$ ,  $\psi_2 = 95^\circ$ 。作为比较, 同时还给出了 DS 波束方向图 (虚线) 和目标波束方向图 (虚点线)