

MAIKEFENG ZHENLIE YOUHUA SHEJI ZHONG
DE SUANFA YU LILUN FENXI

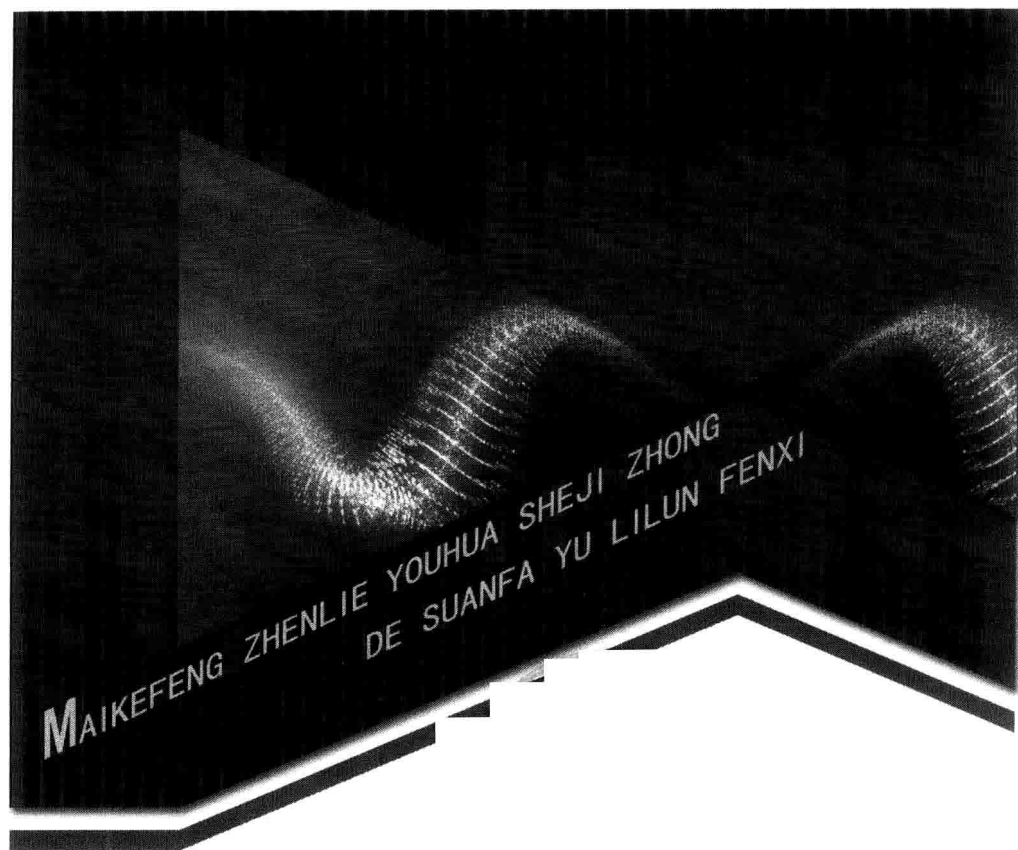
麦克风阵列优化设计 中的算法与理论分析

► 冯志国 编著



重庆大学出版社
<http://www.cqup.com.cn>

本专著受重庆市科学技术协会和重庆师范大学学术专



麦克风阵列优化设计 中的算法与理论分析

► 冯志国 编著

重庆大学出版社

内容提要

麦克风阵列在语音信号处理中发挥着非常重要的作用。本书从最优化的方法出发,系统地考虑了麦克风阵列的优化设计问题。书中分别考虑了麦克风阵列的近场设计、远程设计、实时设计、鲁棒性设计、摆放设计和离散系数设计等一系列问题,采用了各种最优化方法求解问题的解,分析了最优解的性质,同时提出了摆放设计等新概念和问题,实现了信号增强和声源定位等功能,在移动通信、视频会议和助听器等领域发挥着重要作用。

本书适用于信号处理、运筹学和应用数学等相关专业的高年级本科生和研究生作参考。

图书在版编目(CIP)数据

麦克风阵列优化设计中的算法与理论分析 / 冯志国
编著. —重庆:重庆大学出版社,2015.6
ISBN 978-7-5624-9060-9

I. ①麦… II. ①冯… III. ①传声器—设计
IV. ①TN641

中国版本图书馆 CIP 数据核字(2015)第 097985 号

麦克风阵列优化设计中的算法与理论分析

冯志国 编 著

策划编辑:鲁 黎

责任编辑:李定群 高鸿宽 版式设计:鲁 黎

责任校对:秦巴达 责任印制:赵 晟

*

重庆大学出版社出版发行

出版人:邓晓益

社址:重庆市沙坪坝区大学城西路 21 号

邮编:401331

电话:(023) 88617190 88617185(中小学)

传真:(023) 88617186 88617166

网址:<http://www.cqup.com.cn>

邮箱:fxk@cqup.com.cn(营销中心)

全国新华书店经销

POD:重庆新生代彩印技术有限公司

*

开本:787×1092 1/16 印张:11.25 字数:154千

2015 年 6 月第 1 版 2015 年 6 月第 1 次印刷

ISBN 978-7-5624-9060-9 定价:32.00 元

本书如有印刷、装订等质量问题,本社负责调换

版权所有,请勿擅自翻印和用本书

制作各类出版物及配套用书,违者必究

前 言

信号处理与最优化一直都是紧密联系的。早期可以从低通滤波器的设计就可以看到,著名的 PM 算法^[1]就采用了切比雪夫逼近问题的交换算法原理,这在滤波器设计中发挥着非常重要的作用。除此之外,在天线信号处理、语音信号处理和图像信号处理等问题中都有着非常多的数学模型,其中大部分都需要转化为优化问题处理。另外,信号处理问题中很多算法的性能都很一般,很多快速的算法并没有在信号处理问题中发挥作用。因此,我们非常有必要将最优化的一些好的算法应用在信号处理问题中。

滤波器在信号处理中的地位是非常重要的,它的功能是将原信号进行处理并且输出有用的信息。因此,关于滤波器的设计就显得十分重要。在很多问题中,滤波器的设计都转换为各种优化问题进行处理。在麦克风阵列的模型中,每个麦克风的后面都有一个滤波器在处理接收到的语音数据。因此,通过这些滤波器的设计可以建立很多优化问题。但是,这些优化问题的规模一般都比较小,需要发展快速并有效的算法去处理这些优化问题。

本书讨论的是最优化方法在麦克风阵列的各类波束形成问题中的应用。全书共分为7章,分别讨论了近场麦克风阵列的优化设计、近场麦克风阵列的鲁棒性优化设计、近场麦克风阵列的摆放设计、远场麦克风阵列的优化设计和麦克风阵列的滤波器离散参数设计等。

第1章分别介绍了信号处理的原理,包括了输入、处理和输出,麦克风阵列的模型及相应的优化设计问题的原理,麦克风阵列模型中常用的优化问题及相应的方法,如整数规划、半无限规划和正半定规划等。

第2章考虑了近场麦克风阵列优化设计问题的最优解的分析。首先是近场麦克风阵列的波束形成问题的建立,然后通过分析最优解在滤波器长度变化时的性质,推导出在哪些条件下,最优目标函数值的下确界可通过滤波器长度趋于无穷时的极限值达到。同时,本书建立了一个特定的优化问题求解最优目标函数值的下确界,即性能极限值。

第3章讨论了求解近场麦克风阵列优化设计问题的各类算法,包括了内点法、活动约束点逼近法和窗函数方法。每种方法都有其特点,它们从不同的角度解决了这个优化设计问题。本书相应地讨论了每一种方法的计算复杂度,并通过实际的数值例子验证了算法的快速性与有效性。

第4章讨论的是近场麦克风阵列的鲁棒性优化设计问题,即当麦克风的位置存在某种不确定性时的优化设计问题。本书首先建立了相应的鲁棒性优化问题,然后通过发展一个逼近的方法,有效地降低了鲁棒性优化问题求解的规模,最后采用各类算法(如内点法)去求解这个问题的数值解,并通过实际例子来说明算法的可行性。

第5章考虑了近场麦克风阵列的摆放设计问题,有效地建立了在近场情形下的麦克风阵列摆放设计问题,包括了滤波器长度固定的情形和滤波器长度不固定的情形。通过引入无限长度系数解将摆放设计问题转化成可以求解的优化问题,本书发展了两种方法,即梯度型算法和结合了梯度型算法与遗传算法的混合下降法。对于每一种方法,本书都给出了具体例子以验证摆放设计问题在性能上的提升是非常有效的。

第6章考虑了远场麦克风阵列优化设计的最优解分析。由于在远场情形下的麦克风阵列优化设计问题已经有了太多的优化算法,本书只考虑麦克风阵列优化设计问题的最优解的性能分析,如分别考虑了频率域下的性能分析

和空间域下的性能分析。对于每一种情形,本书推导了一系列重要的条件,使得在这些条件下,最优目标函数值能够趋向于最理想的性能值,即零。为了验证这些收敛性的条件,本书给出了很多实际例子,充分阐明了各种条件的必要性。

第7章考虑了麦克风阵列的滤波器离散系数设计。当滤波器的系数采取SPT表示以降低滤波器实现的硬件成本时,本书建立了相应的离散系数设计问题。本书发展了一个创新的转换方法,可将空间 $\{0,1,-1\}^n$ 很方便地转换为一个大的整数空间,从而使得变量的个数得以显著减少,这有效地降低了问题的规模。本书将问题转化为一个无约束的整数规划问题,并发展了一个基于离散填充函数的全局最优化方法求解这个问题。它可在局部最优解之间切换,从而能有效地搜索到比较好的解。

优化算法的改进对于麦克风阵列优化设计问题在性能和效率上都有着很好的提升。另外,麦克风阵列的优化设计问题还远远不止这些。但是,求解这些问题所采用的优化方法同样适用于其他情形的麦克风阵列优化设计问题,并起到指导性的作用。

本书的主要工作是我在访问香港理工大学姚嘉晖副教授期间所做,在此,对香港理工大学以及姚嘉晖副教授表示非常感谢。

编 者

2014年11月

目 录

第 1 章 麦克风阵列优化设计综述	1
1.1 数字信号处理的原理	1
1.2 麦克风阵列优化设计原理	4
1.3 常用优化问题介绍	9
第 2 章 近场麦克风阵列优化设计的性能分析	18
2.1 简介	18
2.2 近场波束形成设计问题	20
2.3 最优解的性能分析	23
2.4 数值例子	30
第 3 章 近场麦克风阵列优化设计的算法研究	32
3.1 简介	32
3.2 内点法	33
3.3 活动约束点逼近法	43
3.4 窗函数快速方法	55

第 4 章 近场麦克风阵列的鲁棒性优化设计	70
4.1 简介	70
4.2 鲁棒性设计问题	71
4.3 逼近性方法	73
4.4 数值例子	77
第 5 章 近场麦克风阵列的摆放结构设计	84
5.1 简介	84
5.2 摆放结构设计问题	85
5.3 数值例子	92
5.4 混合下降法	101
第 6 章 远场麦克风阵列优化设计的性能分析	108
6.1 简介	108
6.2 远场波束形成设计问题	109
6.3 频率域下的性能极限值	113
6.4 空间域下的性能极限值	119
6.5 数值例子	130
第 7 章 麦克风阵列的滤波器离散系数设计	138
7.1 简介	138
7.2 滤波器离散系数设计问题	140
7.3 无限精度解	143
7.4 有限精度解	147
7.5 数值例子	155
参考文献	161

第 1 章 麦克风阵列优化设计综述

1.1 数字信号处理的原理

信号处理就是把记录下来的信号进行处理,以便抽取有用信息的过程,它包括了对信号进行提取、变换、分析、综合等处理过程,从而得到人们感兴趣的信息。信号处理的目的包括了以下三个方面:第一是去伪存真,即把信号中多余的部分(如干扰性噪声)去掉;第二是信号特征的抽取,即把信号转化成易于分析、识别的形式以更好地作后续性处理;第三是调制与解调,即将信号转化为容易传输、变换和存储等形式,或者从这些形式的信号还原出原始信号。

伴随着计算机的发展,信号的处理采取了数字化的过程,即先把模拟信号转化为数字信号,然后利用高效的数字信号处理器或各种算法对其进行数字信号处理。它的原理如图 1.1 所示。

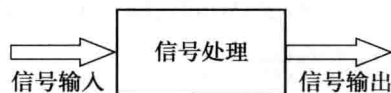


图 1.1 数字信号处理流程图

由图 1.1 可知,一个数字信号处理流程分为以下 3 个部分:

(1) 信号输入部分

在这一部分,连续型信号被接收后,通过一个模数转换器将其转化为离散形式的数字信号。这是一个对自变量和幅值同时进行离散化的过程,采样定理保证了数字信号在足够大的抽样频率下,信号的真实度会缺失。

(2) 信号处理部分

这一部分主要是对离散化的信号进行处理,使得处理后的信号符合某些给定的目标。这包括了频域变换、数字滤波、识别、合成等过程。

(3) 信号输出部分

这一部分主要是输出处理过的离散信号,并通过一个数模转换器还原为连续性信号。但是这一步并不是必需的。

因此,第一部分和第三部分对于信号没有本质的改变,即没有牵涉到设计的问题。可以设计的是第二部分,它可将信号转化成所需要的形式以满足给定的性能要求。因此,中间的处理部分是信号处理问题的关键所在。这一部分的工作就是采用滤波器来实现的,从而引申出滤波器的设计问题。

下面介绍滤波器。滤波器是信号处理的核心部分,它的设计决定了系统所具有的功能。一般而言,数字滤波器主要分为有限脉冲响应(FIR)滤波器和无限脉冲响应(IIR)滤波器。其中,FIR 滤波器是线性的,而 IIR 滤波器是非线性的。

从稳定性、易实现性的角度来看,FIR 滤波器非常有优势,因此得到了广泛的应用。本书主要采用了 FIR 滤波器,它的原理如图 1.2 所示。

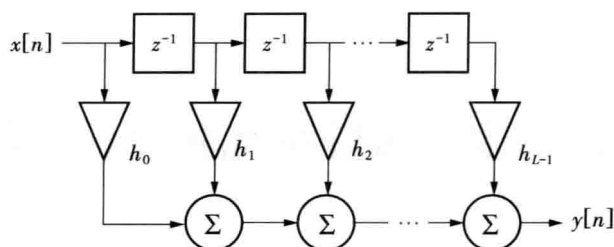


图 1.2 FIR 滤波器的流程图

对于输入的信号,它的处理方式是卷积形式,即

$$y(n) = \sum_{i=0}^{L-1} x(n-i) * h(i) \quad (1.1)$$

其中, x 是输入信号, y 是输出信号, h 是滤波器的系数。信号数据是先进先出的,具体的输出表达为

$$y(1) = x(1)h(0)$$

$$y(2) = x(2)h(0) + x(1)h(1)$$

$$\vdots$$

$$y(k) = x(k)h(0) + x(k-1)h(1) + \cdots + x(k-L+1)h(L-1)$$

$$\vdots$$

也可将上式写为矩阵表示形式,即

$$\begin{pmatrix} y(1) \\ y(2) \\ \vdots \\ y(L) \\ \vdots \\ y(k) \end{pmatrix} = \begin{pmatrix} x(1) & 0 & 0 & \cdots & 0 \\ x & x & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ x(L) & x(L-1) & x(L-2) & \cdots & x(1) \\ \vdots & \vdots & \vdots & & \vdots \\ x(k) & x(k-1) & x(k-2) & \cdots & x(k-L+1) \end{pmatrix} \begin{pmatrix} h(0) \\ h(1) \\ \vdots \\ h(L-1) \end{pmatrix}$$

需要补充一点的是,在很多模型里会采用很多个滤波器同时进行处理,如滤波器组和阵列中的滤波器等,它的原理与上面类似,只是由于存在若干个滤波器,在处理部分我们需要对这些信号进行必要的分解和合成。

1.2 麦克风阵列优化设计原理

信号的种类非常广泛,其中最常见的一种信号就是语音数据。语音作为互相交流和互相通信的最主要和最便捷的工具,相应的应用范围就非常广,对于它的处理就显得非常重要。

对于语音数据的处理,首先需要做的是接收语音数据。耳朵是接收语音数据最常见的工具,但是它无法将数据保存下来进行处理。麦克风的的功能就是接收语音数据,并且可将数据保存下来进行处理,如实现降噪和信号增强等。但是,如果麦克风只有一个,那么它可以实现的功能很少,并且效果不太好。现在开始逐渐采用多个麦克风的技术,即采用多个麦克风组成的阵列进行处理。这个技术就是麦克风阵列,目前涉及很多相关的文献,如文献[2-9]。它不仅在语音应用方面有着广泛的应用,同时在雷达、声呐、无线通信和生物医学等应用上也发挥着重要的作用。麦克风阵列技术不仅能实现降噪、信号增强等功能,同时也在视频会议、移动通信、助听器等实际应用上普遍使用^[10-12],麦克风阵列的技术甚至已经开始在一些高端智能手机上应用。

麦克风阵列的原理从天线阵列中来,不同的是天线阵列处理的是窄带信号,并且信号都是远场的,即信号从一个很远的地方发送过来。麦克风阵列处理的语音信号可以是远场的,如户外语音信号,它的反射效果很小,近似认为是远场信号。麦克风阵列还可以处理近场语音信号,如在一个房间里面的语音信号,反射的影响非常的大。

麦克风阵列并不是简单地把很多个麦克风放在一块,它们之间有着相互的联系。具体而言,它的结构如图 1.3 所示。

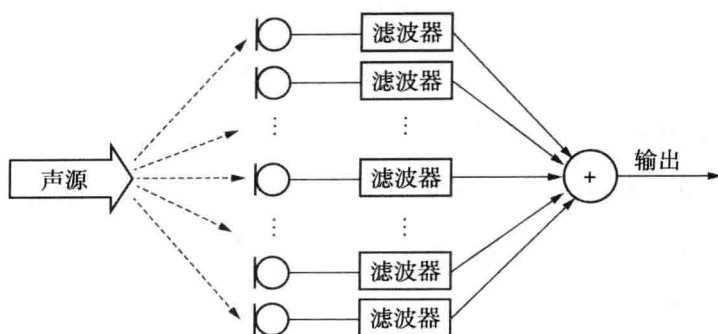


图 1.3 麦克风阵列的结构

由图 1.3 的左边可以看到,语音信号首先有它的信号源,如果有很多信号,那么就有很多信号源。例如,人的喉咙就是人声的信号源,各种环境噪声也有相应的信号源,这些信号是可以混在一起的。

由图 1.3 的中间部分可以看到,语音信号从声源位置发出后,传递到各个麦克风。麦克风接收到语音信号后,首先将它离散成数字信号。为了对这些数字信号进行处理,在每个麦克风的后面都带有一个滤波器。这些滤波器都带有一定的功能,信号经过滤波器的处理后会实现一定的功能。

由图 1.3 的右边可以看到,信号从滤波器处理出来后会得到很多信号,需要将这些信号合成为一个信号,这是一个求和操作,然后就得到输出的信号。这个信号就会实现所需的功能,如降噪、消除回声等。

对于这个模型,还需要考虑以下因素:

1.2.1 信号模型

信号模型决定了信号处理的方式。不同的信号模型下,采取的信号处理方法会有很大不同。关于信号模型,有以下几点考虑:

首先是信号的类型。信号从物理形式来看有很多种表现形式,如光、电、声等。从维数来看,信号可分为一维信号(如语音)、二维信号(如图像)、多维信号(如彩色图像)。本书处理的是语音这种一维信号,相比起二维或多维信

号,复杂度并不高。但是,由于它是宽带信号,处理起来比较复杂。

其次是传递函数。在某些信号模型里,源信号的位置与接收设备的位置一般不同,这导致源信号与接收端的输入信号会有所不同,这两边的信号由一个传递函数决定,这个传递函数决定了输入的信号相较于源信号所发生的变化。不同环境下,这个传递函数是不一样的。一般来说,按照环境的范围可分为远场和近场两种情况。

(1) 远场环境

这种情形一般指在户外的空旷环境里,语音信号距离麦克风足够的远,并且语音的反射可以忽略。具体而言,就是声源与麦克风的距离大于 $2L^2/\lambda$ 的环境,其中, L 是麦克风阵列的长度, λ 是语音的波长。对于这种情形,传递函数由输入信号的方向决定,即

$$A(f, \theta) = e^{-j2\pi fr \cos \theta}$$

式中 f ——信号的频率;

θ ——信号方向与某个给定的参考方向的夹角;

r ——麦克风位置与某个给定的参考位置的距离。

在远场这种环境下,处理的是信号的频率和入射方向的信息。

(2) 近场环境

这种情形指源信号与麦克风的距离不大的环境,即声源与麦克风的距离小于 $2L^2/\lambda$ 的环境。在这种环境下,不同位置的声源到麦克风的传递函数的区别是明显的。它一般分为无反射和有反射两种情况考虑。在一些隔音效果很好的环境里,声音的反射近似可以忽略,那么,无反射的情况就可以考虑。它的传递函数主要由源信号与麦克风的距离决定,即

$$A(f, r) = e^{-j2\pi f \|r - r'\|}$$

式中 f ——信号的频率;

r ——信号源的位置;

r' ——麦克风的位置。

在一般的近场环境里都存在着反射回音的情况,它的传递函数不仅由源信号与麦克风的距离决定,同时也被各种反射的信号所影响。因此,它的传递函数非常复杂。目前并没有一个通用的公式,但有一些给定模型下的公式,如映射模型^[13],它是由上式经过无限次反射所得到的和函数的模型。

在近场这种环境下,处理的是源信号的频率和位置的信息。

1.2.2 傅里叶变换

傅里叶变换在信号处理中的应用非常广泛。一般来说,在时域里是不需要考虑傅里叶变换的,但是如果考虑信号的频率信息,就需要对信号做傅里叶变换,将信号转为相应的频率表达形式。

对于输入的离散信号 x , 它的傅里叶变换为

$$X(f) = \sum_{k=1}^n x(i) e^{j2\pi kf} \quad (1.2)$$

对于 FIR 滤波器, 它的频率响应定义为

$$H(f) = \sum_{k=0}^{L-1} h(i) e^{j2\pi kf/f_s} \quad (1.3)$$

式中 f_s ——抽样频率。

通过傅里叶变换, 得出卷积(1.1) 的频率形式为

$$Y(f) = X(f) H(f) \quad (1.4)$$

式中 X ——输入信号的频率响应;

Y ——输出信号的频率响应;

H ——滤波器的频率响应。

卷积的频域表达形式比起时域的表达形式简单很多, 处理起来也比较容易。

1.2.3 性能指标

对于最优化问题而言,需要有一个目标函数。因此,滤波器系数的设计需要对应一个性能指标,即输出的信号需要达到的性能要求。在现实中有很多种性能指标,最常见的是信号增强,如降噪、消除回声等。对于不同的信号模型,相应的性能指标也会不一样。例如,在无线信号或语音信号的定位模型中,性能指标可以是源信号的定位误差。

一般情况下,性能指标会采取降噪的形式。对于这种形式,相应的具体定义也可以有很多种,例如:

① 最小误差形式,即输出信号与理想信号的误差值达到最小。它的形式可以表示为

$$\sum_{i=1}^n \|y(i) - y_d(i)\|^2 \quad (1.5)$$

这是误差的最小方差即 l^2 形式。除了这种形式,还有 l^∞ 形式,即

$$\max_{i=1, \dots, n} |y(i) - y_d(i)| \quad (1.6)$$

另外,为了寻找稀疏系数解,需要加上系数的 l^1 形式,即

$$\sum_{i=1}^n |h(i)| \quad (1.7)$$

上面的形式都是从时域去考虑的。在实际问题中,由于信号带有频率性质,经常需要从频率域去考虑问题,上面的形式就转化为频率域下相应的形式,如式(1.5)变为

$$\sum_{f \in \Omega} \|Y(f) - Y_d(f)\|^2 \quad (1.8)$$

其中, Ω 是某个给定的频率区域。

② 最大信噪比形式,即信号与噪声的比值。它的形式可表示为

$$\frac{\mathbf{h}^H \mathbf{R}_s \mathbf{h}}{\mathbf{h}^H \mathbf{R}_n \mathbf{h}} \quad (1.9)$$

式中 R_s, R_n ——源信号和噪声的协方差矩阵。

③ 最小信号扭曲形式和最大噪声抑制形式。这两种形式是新的定义方式,将在第7章详细讨论。

一般来说,最优化过程中性能指标只能有一个。如果需要同时考虑很多性能指标,那么只能选定某个性能指标作为目标函数,其他的性能指标则转化为限制条件。

基于上面这些基础,我们就可以考虑滤波器的优化设计问题了。它的一般形式可叙述如下:

问题 1.1 给定某个信号模型,设计滤波器的系数,使得某个给定的性能指标达到最优,同时满足某些限制条件。

上述限制条件不仅包括了某些性能要求,也包括了其他条件,如关于滤波器系数的限制、鲁棒性设计中的不确定参数的限制等。

其实,这个优化设计问题的表述并不是很具体。对于每个具体的信号模型,优化问题的具体表现形式不一样,相应的处理方式也会不一样。从第2章开始,本书将针对于不同的滤波器设计问题进行处理。

1.3 常用优化问题介绍

信号处理中经常涉及优化问题,它的一般形式为

$$\begin{aligned} \min_{\mathbf{x} \in \mathbf{R}^n} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g_i(\mathbf{x}) \leq 0 \quad i = 1, \dots, m \end{aligned} \quad (1.10)$$

式中 $\mathbf{x}$ ——需要优化的变量;

$f(\mathbf{x})$ ——目标函数;