目录

[1绪论 3](#_Toc130287644)

[1.1课题来源与背景 3](#_Toc130287645)

[1.2选题意义 3](#_Toc130287646)

[1.3研究内容与目标 4](#_Toc130287647)

[1.4 论文章节安排 4](#_Toc130287648)

[2光学自动设计概述 5](#_Toc130287649)

[2.1光学自动设计概念 5](#_Toc130287650)

[2.1.1 光学设计 5](#_Toc130287651)

[2.1.2 光学自动设计 6](#_Toc130287652)

[2.2 光学自动设计问题与挑战 6](#_Toc130287653)

[3 常用的光学自动设计算法 10](#_Toc130287654)

[3.1适应法 10](#_Toc130287655)

[3.2正交法 11](#_Toc130287656)

[3.3最小二乘法 11](#_Toc130287657)

[3.4阻尼最小二乘法 12](#_Toc130287658)

[4算法设计 14](#_Toc130287659)

[‍4.1 数学模型 14](#_Toc130287660)

[‍4.2 一般信赖域方法 15](#_Toc130287661)

[‍4.3 基于信赖域方法的改进算法 16](#_Toc130287662)

[4.3.1 停止准则 16](#_Toc130287663)

[4.3.2 半径更新策略 17](#_Toc130287664)

[4.3.2 半径放缩方式 19](#_Toc130287665)

[‍4.4 改进算法总结与分析 20](#_Toc130287666)

[5算法验证与分析 21](#_Toc130287667)

[5.1 数据集与实验运行环境 21](#_Toc130287668)

[5.1.1 实验数据集介绍 21](#_Toc130287669)

[5.1.2 实验运行环境介绍 21](#_Toc130287670)

[5.2 验证xxx问题的影响实验 21](#_Toc130287671)

[5.3 xxx算法性能比较实验 21](#_Toc130287672)

[6总结与展望 21](#_Toc130287673)

[6.1论文总结 21](#_Toc130287674)

[6.2工作展望 21](#_Toc130287675)

[‍参考文献 22](#_Toc130287676)

# 1绪论

## 1.1课题来源与背景

本论文选题来自国家重点研发项目激光材料及器件在线测试与自动化设计技术，项目牵头单位为中国科学院软件研究所。该项目将激光和计算机技术交叉融合，使光学设计软件到检测仪器均实现自主可控。其中的智能光学设计系统可以实现复杂面型建模、折反光线追迹、全局优化、异构并行加速等光学自动设计技术。光学自动设计算法是决定该智能光学设计系统性能的重要内容。本论文的工作是改进在光学自动设计中常用的阻尼最小二乘法，提出一种新的光学自动设计算法。

## 1.2选题意义

随着中国科技和制造水平的不断提升，激光系统研发模式的自动化和智能化需求日益增长，但目前我国尚无自主可控的光学设计软件，国外商用软件智能化程度仍显不足，无法满足全参数、高精度的智能化光学设计的需求。此外，在该领域，国外商用软件垄断了大部分中国市场，核心技术依赖外部输入。因此，设计一款自主控制的光学软件具有重大现实意义，而本论文计划研究的光学自动设计算法是该软件自主创新的重要一环。此外，目前对于阻尼最小二乘法的几种改进算法存在不足，尚有优化空间。本论文引入信赖域方法的思想，对阻尼最小二乘法进行改进，使其能够更好的应用在光学自动设计问题中，具有较强的理论意义。

## 1.3研究内容与目标

## 1.4 论文章节安排

本文深入研究了光学自动设计算法，并用信赖域思想对其进行改进和完善，从满足像差目标出发，对优化质量进行提升，同时考虑计算花费，以获取更高性能的优化方案。本文将分为六个章节对其进行阐述，其主要结构安排如下：

第一章作为本文的绪论内容，对课题来源以及研究背景及意义做了简单的介绍，在提出本文的研究内容及目标以后给出本文内容的大致安排。

第二章作为对后续章节的引入，具体阐述了光学自动设计问题以及面临的问题与挑战。

第三章针对光学自动设计对几种常用的算法进行研究，给出了它们的优缺点分析。

第四章将信赖域思想引入阻尼最小二乘法，提出了一种基于信赖域分解的光学自动设计算法。该方法改进了原有阻尼最小二乘法，保证了xxxx。

第五章对本文算法进行验证和分析，xxx

第六章对本文的主要研究内容进行了总结，针对存在的不足，从算法的效率和xxx指出了本文的后续研究方向。

# 2光学自动设计概述

## 2.1光学自动设计概念

### 2.1.1 光学设计

光学设计是一门利用光学原理和相关学科知识，以优化和改进光学系统的结构和性能的学科。例如，常见的望远镜、显微镜、相机等均为不同类型和数量的透镜、反射镜、棱镜等光学元件组成的复杂系统，它们能够对入射到系统中的光线进行传输和变换，进而实现放大、缩小、聚焦等功能。这些功能取决于各个元件对于不同波长、不同方向、不同位置上入射光线所产生的影响。这些影响可以用数值来表示，称为像差。像差越小，说明系具有更高的清晰度、分辨率、色彩还原度等指标。

在实际应用中，理想的像差水平并不是所有方案都能达到的。要寻找一个既符合功能要求又具有低像差的优良方案，需要光学设计者进行多次计算和调整，这是一个复杂而耗时的过程。在此过程中，人类设计者虽然可以发挥自己的经验和判断力，但也难以避免遇到困难和瓶颈。因此，传统光学设计存在着人为干预过多、效率低下等问题。

一般来说，一个完整的光学设计流程可描述为：首先根据所要实现的功能选定可能合适或者已经存在过类似案作为参考，并确定各个元件的结构参数（如厚度、曲率半径、材质等），这些结构参数组成了一个方案中所有元件的特征。然后利用数值方法或者近似公式计算出该方案下系统产生各种像差函数值。如果像差函数值超过了预设范围，则依靠设计者的经验或者直观感受对结构参数进行修改，并重新计算像差函数值；如果像差函数值在预设范围内，则说明该方案已经达到了目标要求，并将其作为最终方案。

在传统光学设计中，光学设计过程和效果都依赖设计者的经验，人类扮演着至关重要但也非常耗时耗力的角色。

### 2.1.2 光学自动设计

随着科技的发展和社会的进步，对高效率、高质量、高复杂度的光学系统的需求日益增加，传统的人工或半自动化的光学设计方法已经难以满足这些需求。因此，光学自动设计应运而生。

光学自动设计是利用计算机科学的方法来进行光学设计的方法。该方法通过运行计算机程序，根据预设的目标函数和约束条件，自动地搜索和调整光学系统的结构参数，从而使光学系统达到最佳或接近最佳的成像质量、色差校正、畸变控制等性能指标。光学自动设计不仅可以提高光学设计的效率和准确性，减小人为的错误和偏差，还可以探索一些传统方法难以发现的新颖的光学系统。

光学自动设计可被理解为一个数学上的最优化问题。在光学自动设计中，自变量为光学系统的各个结构参数，包括镜头各面的曲率半径，透镜的厚度以及透镜之间的间隔等。镜头的焦距、横向放大率，以及各类几何像差、波像差都称为像差，它们都是结构参数的函数。光学自动设计的目标是寻找一组结构参数，可以满足像差函数值达到或接近目标值。这可以被理解为舒徐中的最优化问题，可采用一些优化算法进行求解。

## 2.2 光学自动设计问题与挑战

在理想情况下，求解光学自动设计最优化问题，可以通过对解析式求导，转化为一个求驻点的问题。但在实际情况中，光学自动设计的复杂特性为其带来了以下问题与挑战。

1. 多模态

光学自动设计问题中可能存在多个全局最优解或多个有意义的局部极值解。这意味着在目标空间中，有多个点或区域都能达到最优或次优的目标函数值，而在决策空间中，这些点或区域可能相距很远，具有不同的形态。因此光学自动设计问题是一个多模态优化问题。多模态问题的难点在于需要在搜索过程避免陷入局部最优或丢失潜在的最优解。

1. 多约束

多约束优化问题是指在求解一个或多个目标函数的最优值时，需要满足一些限制条件的问题。由于多约束优化问题涉及多约束条件，通常很难找到一个解能够同时使所有目标函数达到最优，并且满足所有约束条件。

在光学自动设计过程中，需要考虑光学系统受到各个参数特性带来的多个约束的影响。多种约束条件为寻找最优解带来了困难。

1. 高维性

在光学自动设计过程中，光学系统的结构参数是影响其性能的重要因素。这些参数包括透镜的曲率半径、厚度、间隙、折射率等，它们都需要根据设计目标进行调整。然而，光学系统中可独立变化的结构参数数量较大，通常有几十个甚至上百个。这导致参数空间具有高维性，即每个参数都是一个维度，组成了一个高维空间。

高维性带来了高维诅咒（Curse of dimensionality）问题，即在高维空间中，低维空间成立的方法不再适用或有效的现象。光学自动设计中高维诅咒主要体现在两个方面：一是距离度量的失效；二是优化问题求解的困难。

距离度量的失效是指在高维空间中两个点之间的距离会趋于相同，即距离失效现象。这是因为高维空间中体积和表面积随着维度增加而指数增长，导致空间中大部分点都集中在边界附近，而边界附近的点之间的距离相差不大。这就导致了距离度量在高维空间中失效或不稳定，在计算不同结构参数点之间的距离时难以比较其大小，这对于基于距离度量的优化算法有很大影响，可能导致结果不准确或无效。

优化问题求解的困难是指在高维参数空间中寻找最优解或满足约束条件的可行解需要消耗巨大的计算资源和时间。这是因为高维参数空间由多个参数构成，每个参数有其取值范围，形成一个维度。不同维度的参数相互组合，构成整个空间。然而，随着维度的增加，参数空间的复杂度和规模呈指数级增长。在如此复杂和广阔的空间中寻找最优解或可行解会导致计算量过大、效率低下且难以实现。因此需要有效的搜索算法或利用启发式信息来缩小搜索范围。

1. 多目标

光学自动设计所要解决的核心问题是像差设计问题，即根据初始设计给出的对系统中每个透镜组的光学特性和成像质量的要求。为了同时考虑多种不同类型或不同视场位置处的像差函数值，并使它们都在限定范围内，需要建立一个多目标优化模型。

多目标优化问题相比单目标优化存在困难，主要有以下几点：

多目标优化问题往往没有一个唯一最优解，而是存在一组非劣解。因此，在多目标优化问题中，需要引入一些方法来评价和选择非劣解。

多目标优化问题中，各种目标函数可能存在复杂且相互冲突的关系，不都是能任意控制的独立的量，即满足一个目标函数可能会导致其他目标函数无法满足。因此，在多目标优化问题中，需要在不同目标函数之间进行协调和权衡，以达到总体上的最优化。

多目标优化问题中，由于各个目标函数可能具有不同的物理意义、量纲或数量级，在进行比较或聚合时需要进行适当的归一化或缩放处理。此外，在进行搜索时还需要考虑各个目标函数之间可能存在的耦合性或非线性。

1. 非线性

光学自动设计的非线性体现在像差函数的非线性。由于光线折射定律是非线性的，像差与结构参数的关系是一个极为复杂的非线性函数关系，甚至不可能给出确切的函数表达式。

由于光学自动设计问题具有高度非线性，存在不可导、不连续、多峰等特点，导致求解过程难以保证收敛和精度，像差函数的梯度信息难以获取。因此无法直接使用解析方法或者基于梯度计算的方法来求解最优化问题。

为了实现光学自动设计，必须将真实像差模型理解成一个黑盒模型，使用数值计算方法求出确定的光学结构参数对应的像差函数值，通过调整光学结构参的方式进行优化。

1. 经验依赖性

光学自动设计并非一种全自动化的过程，它仍然依赖设计者在初始阶段进参数设置，并在设计阶段进行干预，以及在最终阶段进行评价。

初始阶段，设计者需要根据光学理论和经验给出合理的初始参数，如镜头数目、形状、材料等，这些参数对后续的优化过程和结果有着重要的影响。若初始参数设置不恰当，则可能导致优化无法收敛甚至生成不可行的方案。因此，设计者需要具备足够的专业知识和创造力来设定和调整初始参数，并根据具体需求进行优化目标和约束条件的制定。

设计过程中，通常使用迭代法来求解问题。然而，迭代过程受初始值、误差、算法本身等因素制约，可能导致收敛速度缓慢、震荡或发散等现象。针对这些问题，设计者需要对设计过程进行人工干预，在保持目标函数不变的前提下人为调整结构参数，以促进迭代收敛并获得较优结果。

最终阶段，设计者需要对光学自动设计生成的镜头方案进行评价和筛选，考虑其可制造性、可靠性和经济性等因素。这些因素取决于工艺条件、市场需求和成本预算等多方面的因素，并不一定与优化目标完全吻合。因此，需要设计者有足够的判断力和决策力，在此阶段选择最适合的镜头方案，并进行必要的修改和验证。

综上所述，光学自动设计难以完全替代人工设计，仍然需要设计者与程序相配合，发挥人类经验的优势。

# 3 常用的光学自动设计算法

现今较为有效而实用的光学自动设计算法有适应法、正交法、阻尼最小二乘法等多种优化方法。

## 3.1适应法

适应法的特点是采取单个像差函数分别达到目标值的方法进行像差校正，利用各种像差对每个自变量的偏导数，近似地把像差和结构参数之间的关系线性化，把每个像差变成一个线性函数。

使用适应法的条件是像差数目必须小于或等于结构参数总数目，若像差数目大于结构参数总数，则适应法不能应用。

假设有n个自变量，m个像差函数。自变量用表示，像差函数用表示，即

将像差函数与结构参数的关系近似为线性关系，得到下列像差方程组

其中，为每个结构参数相应的改变量，是初始结构参数对应的像差函数值，是初始结构参数加上修正向量时的像差函数值，是像差函数对每个结构参数的偏微商，各偏导数可以通过差分方法获得。

当时，结构参数与像差函数的数量相同。像差方程组可得到唯一解，接着这个解构造出新的像差函数方程再次求解，重复这个过程直到到达最小像差。

当时，意味着结构参数比像差函数数量多。这样就导致了像差方程组有无穷多组解，其中每一组都能满足像差方程组。但是我们希望找到一个最优的解，使得光学系统成像质量达到最高或者接近理想情况。为了从无穷多组解中选择一个确定的解，我们可以再添加一个附加条件来约束结构参数。一种常用的限制条件是要求解向量长度最小，即要求结构参数不出现大幅度变化。数学上就是要求值最小。选择该限制条件的原因是满足该要求的解能够使光学结构更加紧凑和稳定，并且与参数和像差函数之间保持良好的线性关系。同时未被优化或者主要考虑以外的其他像差也会相对较小或者不显著影响成像质量。为了在给定限制条件下找到使值最小的解，可以用拉格朗日乘子法来求解这个最优化问题。这样就得到了在第一步中的最小化解的值和对应的内部结构参数 。但是这个解并不一定是最优的，因为它只满足了第一步中的限制条件。为了进一步优化光学系统，我们需要根据第一步中得到的解构造新的像差函数方程。这个过程可以重复进行，直到达到预期的优化效果。

适应法让单项像差分别达到目标值的特点与人工校正的过程相近，因此可以充分发挥设计人员的经验和像差理论知识的作用。此外，适应法不存在陷入局部极值的问题，因此对初始设定结构参数值依赖性较小。这些特点使得适应法能够给出明确的校正步骤，并对系统效果产生显著的提升。

## 3.2正交法

正交法优化是一种优化方法，它通过将变量分解成互相垂直的方向，并在每个方向上进行离散的采样，来寻找使评价函数值最小的解。 正交下降优化不需要计算评价函数的导数，因此适用于评价函数有噪声或不光滑的情况，比如非序列光学系统。

## 3.3最小二乘法

最小二乘法（least squares method）是一种数学优化技术，其目标是寻找使误差平方和最小的参数匹配。误差平方和衡量了观测值与理论值之间的偏离程度，反映了数据与模型的拟合质量。最小二乘法可以通过求解误差平方和函数关于参数的一阶条件来得到解析解。

最小二乘法可以应用在光学设计领域中。在光学设计中，最小二乘法可以通过最小化像差函数值与目标值之差的平方和来优化结构参数，从而提高光学系统性能。 然而，由于像差函数具有高度非线性，该问题属于复杂非线性最小二乘问题。对于非线性最小二乘问题，基于导数标准方法不适用，因为它们需要计算复杂且不稳定的高阶导数，无法直接求得解析解。在这种情况下需要采用迭代方法求解。迭代方法通过逐步搜索参数空间来进行估计，在每一步都采用函数线性近似来简化计算，并逐次修正该近似以增加精度和收敛速度。

非线性最小二乘问题中逐步修正的方式一般包括梯度下降法和高斯-牛顿法。梯度下降法根据目标函数的负梯度方向更新参数向量，每次更新的幅度由预设的学习率控制。该方法的优势在于实现简单，适用范围广泛。该方法的劣势在于收敛速率较低，易受局部最优解或鞍点的影响。高斯-牛顿法利用一阶泰勒展开将非线性函数近似为线性函数，然后求解新近似误差函数平方和的零梯度条件得到参数向量的更新量。该方法的优势在于收敛速率较高，能够有效利用目标函数的曲率信息。该方法的劣势在于计算复杂度较高。

## 3.4阻尼最小二乘法

1.阻尼最小二乘法是xxxx（算法）

2阻尼最小二乘法的思想（gn+gd）有优势

3核心难点：阻尼因子xxx调整方式

4存在不足

阻尼最小二乘法是一种用于参数估计和函数逼近优化等问题的算法。阻尼最小二乘法的基本思想是在每一步迭代中，求解一个带有正则化项的线性最小二乘问题，从而得到变量的修正量。阻尼最小二乘法的优点是可以避免奇异点和不稳定解的出现，同时具有较高的收敛速度和计算效率。

阻尼最小二乘法对高斯-牛顿法和梯度下降法求解方式进行插值得新的修正向量计算方法。插值系数被称为阻尼因子。

阻尼最小二乘法作为高斯-牛顿法和梯度下降法的结合方法，可综合两种方法的优势。该方法既有梯度下降法从远离极值点收敛的能力，也有高斯-牛顿法从近极值点快速收敛的能力，具有较好的稳定性。

该方法将非线性最小二乘问题转换为迭代求解非线性方程组问题，计算方式比较简单。此外，阻尼因子的引入有效避免了步长过长的问题。但是在迭代过程中，调整阻尼因子的策略只考虑了修正向量的正确性，直接放缩阻尼因子，没有进一步细分，因此该策略存在粗略、不够合理的问题。

# 4算法设计

## ‍4.1 数学模型

光学设计问题的数学模型可表述为[1–3]：

考虑一组结构参数以及个关于的非线性像差函数，用维向量表示光学系统的目标像差。

期望得到的光学系统对不同像差达到设定目标的要求程度不同，因此采取加权的方式对不同像差函数加权，对应像差函数有加权向量。

由此得到方程组

方程组可简写为

定义评价函数

根据最小二乘法的思想，问题转化为求解使最小的解。

该问题是非线性最小二乘问题，无法直接求解，因此需要在迭代中不断修改结构参数，逐渐降低评价函数的值。

假设中参数线性变化，可给出的线性估计。当第次迭代时，假定修正向量足够小，近似函数为

其中下标表示迭代次数， 是的雅可比矩阵。

预计增加后，的估计值为

令偏导数为可得解得修正向量。为防止下降的步长过长，为添加阻尼项，是阻尼因子[4]。

令偏导数为使新的达到最小值，得到

则是阻尼最小二乘法在每次迭代中计算修正向量的方法。

阻尼最小二乘法中，如何在迭代过程中调整阻尼因子会影响到算法的性能。调整的一般策略是检验当前解得的修正向量是否使评价函数下降。如果减小，则接受新的并为下一次迭代减少；如果评价函数没有减小，说明当前太小，则不更新，并增大。[5]

该方法将非线性最小二乘问题转换为迭代求解非线性方程组问题，计算方式比较简单。此外，阻尼因子的引入有效避免了步长过长的问题。但是在迭代过程中，调整阻尼因子的策略只考虑了修正向量的正确性，直接放缩阻尼因子，没有进一步细分，因此该策略存在粗略、不够合理的问题。

## ‍4.2 一般信赖域方法

信赖域算法是一种数值优化算法[6]。信赖域算法通过在当前迭代点附近区域搜索的方法获得新的迭代点。

原非线性方程组可写作一个一般优化问题

定义作为函数的线性近似函数

该函数在足够小时才能认为是的近似，因此只在当前附近区域使用该近似函数。在信赖域算法中，假定在一个以当前为中心，半径长度为内的球内，该线性估计是准确的。该球形区域被称为信赖域，被称为信赖域半径。

计算修正向量的方法是在信赖域中寻找使估计函数值最低的，即求解问题，该问题称为信赖域子问题。

定义为真实下降值与预计下降值之比，用于决定是否接受该修正向量以及如何调整信赖域半径。

较大的值反映在相同的预计下降值条件下达到了更低的评价函数值。因此一般的信赖域半径更新策略为在值大于预设值时接受该修正向量[7]并扩大信赖域半径，反之缩小信赖域半径。

## ‍4.3 基于信赖域方法的改进算法

### 4.3.1 停止准则

为了保证算法能够有效地终止，需要设定两个停止准则[1]：一是迭代次数达到预先规定的最大值，即；二是算法收敛到一个可接受的解。根据最小二乘法的理论，算法收敛的理想条件是评价函数等于，即残差函数平方和为零。然而，在光学自动设计中，由于问题本身的复杂性，很难实现残差函数全部为。因此，只能采取次优策略，寻找评价函数的极小值点。通常情况下，当梯度向量的范数小于某个给定的阈值时，可以认为评价函数已经达到局部最优点。该停止条件可以表示为，其中是一个较小的正数。

为了提高光学自动设计问题中评价函数的下降效率，本文在已有的停止准则基础上，增加了一种基于步长大小的停止准则。该准则主要针对计算得到的修正向量总是无法有效地降低评价函数的现象。当出现这种情况时，阻尼因子会逐渐增大，从而导致步长过小。由于计算机数值运算存在精度误差和舍入误差等问题，可能导致预期下降值和实际下降值均被舍入为，使得无法满足继续迭代或者更新阻尼因子的条件，并且失去了优化目标函数的意义。此外，过小的步长也不利于评价函数的快速下降，并且会增加计算时间。因此，本文认为，在步长小于一定阈值时，应该终止迭代。该停止条件表示为，其中是一个较小的正数。

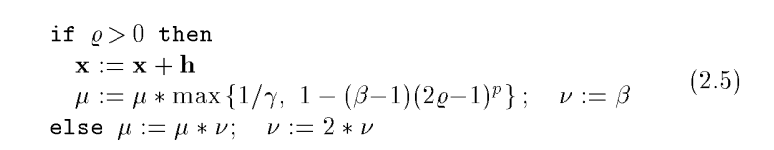
### 4.3.2 半径更新策略

根据信赖域半径更新方式，可得到一种新的更新阻尼因子的方式。

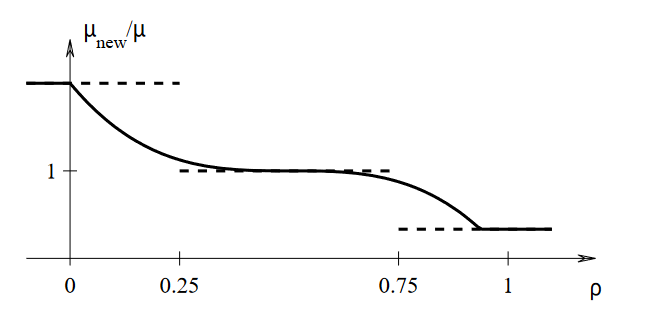
xxx

ρ判断优化结果。

但是该方法存在粗略性，没有对各种情况进行细致的分析和处理。分类方式比较粗糙，只是简单地将的值域直接分为几个大区间，而没有考虑更多的细节和特征。为改进算法，Nielsen提出了一种新的策略[8]，对阻尼因子的更新进行细化。



如图是两种方法更新阻尼因子策略的对比。



一般方法中，阻尼因子的变化比率随值变化而不连续，导致在和阈值附近出现阻尼因子的急剧波动。为了克服这一缺陷，Nielsen提出了一个三次函数模型，对原有的不连续曲线进行拟合，在阈值前后保持较为平滑的变化趋势。实验结果显示，与传统方法相比，该策略使得阻尼因子在迭代过程中变化更加平滑，避免了抖动现象的发生。这样不仅提高了算法的稳定性，而且减少了迭代次数，提升了计算效率。

此外，Nielsen在更新策略中引入了一个变量，以增强阻尼因子的调节能力。当连续迭代中出现的情况时，变量的指数增长使得阻尼因子可以迅速达到一个较大的值。这样可以有效地缩小步长，提高线性模型对真实模型的拟合精度，进而获得一个更优的下降方向。

Nielsen策略应用在光学自动设计问题中会经常出现算法在未达到最优解时就停止迭代的现象。本文认为，这是由于Nielsen策略存在的局限性导致的。

首先，Nielsen策略使用了给定的的值阈值（0.25和0.75）拟合阻尼因子调节比例，而这些阈值并不适合所有类型的光学设计问题。由于光学自动设计问题通常具有高度非线性和复杂性，线性模型往往不能很好地拟合真实模型，导致预测改变量与实际改变量相差较大，因此在优化过程中值通常较小。如果仍然采用Nielsen策略，阻尼因子更容易增长而不是下降，阻尼因子在迭代过程中总体呈升高趋势，导致步长过小而退出算法。

此外，Nielsen策略在连续出现的情况总使阻尼因子乘一个指数增长的系数。由于光学自动设计问题极高的非线性，值常常小于零，在这个过程中阻尼因子以极高的速度增长，在少量几次的迭代后阻尼因子会达到一个足够高的值使步长过小导致算法过早退出。在这个急速提升阻尼因子的过程中可能会错过使真实函数下降的方向。这也是Nielsen策略在光学自动设计问题中存在的局限性之一。

为了减轻Nielsen策略在光学自动设计中的局限性带来的影响，可对Nielsen策略进行以下改进。

为了保证减少阻尼因子抖动的前提下避免算法过早退出，可以在Nielsen策略思想的基础上，基于较小的阈值拟合一条新的更新曲线，并用该曲线作为更新阻尼因子的依据。经验表明xxx曲线效果较好。经验表明，xxx曲线的效果较好。

由于Nielsen策略的阻尼因子会随着的连续出现而指数级地放大，这对于光学自动设计来说是不合适的。因此，本文提出了一种改进的方法，即在时，只将阻尼因子乘以一个固定的常数，而不是指数增长。这样可以减缓阻尼因子的增长速度，从而增加探索更多潜在下降方向的机会。

最后，本文还考虑了阻尼因子的重置问题。由于在的连续出现的过程中，阻尼因子会不断增大，这可能导致后续迭代过程中步长过于小，下降速度变慢，从而影响算法的收敛效率。为了避免这种情况，本文采取一种阻尼因子的重置策略，即在的迭代过程中，如果出现一次的迭代，则说明当前搜索方向是有效的，此时将阻尼因子恢复到初始设置值，以便于在接下来的迭代中仍能在较为广泛的范围内搜索有效下降步骤。

### 4.3.2 半径放缩方式

缩放（Scaling）是一种利用线性变换处理优化问题的技术。缩放最初源于几何学，指的是按照一定的比例因子在不同方向上对图形进行放大或缩小。后来，缩放也被应用于优化问题，以解决数值误差和病态问题。在优化问题中，缩放意味着对目标函数或约束条件进行适当的变换，使其数值范围更加合理，从而提高优化算法的效率和稳定性。

在优化问题中，如果目标函数对某一特定变量的微小变化非常敏感，则称该问题存在缩放不良（poorly scaled）现象。缩放不良可能导致不同变量下降速度不均衡，造成收敛速度缓慢。此外，目标函数的数值误差也可能被放大，影响优化结果的精度。

为了减轻缩放不良的影响，可以对变量进行合适的缩放处理。当变量的变化具有不同速度时，在建模过程中可以采用对角线缩放（diagonal scaling），即将变量乘以一个对角矩阵，使其数值范围更加均衡。这样做可以使问题更加平衡和稳定，并且可以用稳健的方式处理问题，提高算法效率和精度。

本文主要讨论信赖域算法中的缩放技术。信赖域算法基于线性模型对真实模型进行近似，因此在真实模型变化率较低的方向上具有较高的近似精度。根据信赖域方法的理论，信赖域边界处的点应具有相同的可信度。然而，当真实模型存在缩放不良现象时，球形信赖域内的搜索可能不合理。因此本文考虑使用对信赖域进行缩放处理后的椭球形信赖域。

为了平衡信赖域各个方向的变化，采取的缩放方式需要满足在真实模型变化率较高的方向上减小信赖域半径，在真实模型变化率较低的方向上增大信赖域半径。使用数学方式表示信赖域限制条件如下

其中，以对角阵对变量进行对角缩放处理。的对角元素分别表示对应变量对所有像差函数的偏导数平方和，可以用于衡量该变量在像差函数下的敏感度。在式（xx）的限制下，敏感度越高的变量，其对应的信赖域半径越小。

以新的信赖域方式重写新的信赖域子问题为

Xxx&xxx

该问题同样可采用拉格朗日乘子法进行处理得到

Xxx

在新问题中，求解该子问题的方式为

Xxx

## ‍4.4 改进算法总结与分析

# 5算法验证与分析

本章通过对xxx的实验，验证了算法对xxx问题产生的影响。在4.3中通过实验分析了xxx的性能，并验证了xxx算法对xxx的改进效果。

## 5.1 数据集与实验运行环境

### 5.1.1 实验数据集介绍

### 5.1.2 实验运行环境介绍

## 5.2 验证xxx问题的影响实验

## 5.3 xxx算法性能比较实验

# 6总结与展望

## 6.1论文总结

## 6.2工作展望

# ‍参考文献

[1] KIDGER M J. Use of the Levenberg-Marquardt (damped least-squares) optimization method in lens design[J/OL]. Optical Engineering, 1993, 32(8): 1731. DOI:10.1117/12.145076.

[2] BUCHELE D R. Damping Factor for the Least-Squares Method of Optical Design[J/OL]. Applied Optics, 1968, 7(12): 2433. DOI:10.1364/AO.7.002433.

[3] Encyclopedia of Optical Engineering[M/OL]. Taylor & Francis, 2011[2023-01-30]. http://www.tandfonline.com/doi/book/10.1081/E-EOE. DOI:10.1081/E-EOE.

[4] MARQUARDT D W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters[J/OL]. Journal of the Society for Industrial and Applied Mathematics, 1963, 11(2): 431-441. DOI:10.1137/0111030.

[5] BOYD S, VANDENBERGHE L. Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares[M/OL]. 1 版. Cambridge University Press, 2018[2023-01-09]. https://www.cambridge.org/core/product/identifier/9781108583664/type/book. DOI:10.1017/9781108583664.

[6] MORÉ J J. Recent Developments in Algorithms and Software for Trust Region Methods[M/OL]//BACHEM A, KORTE B, GRÖTSCHEL M. Mathematical Programming The State of the Art. Berlin, Heidelberg: Springer Berlin Heidelberg, 1983: 258-287[2023-02-03]. http://link.springer.com/10.1007/978-3-642-68874-4\_11. DOI:10.1007/978-3-642-68874-4\_11.

[7] YUAN Y xiang. Recent advances in trust region algorithms[J/OL]. Mathematical Programming, 2015, 151(1): 249-281. DOI:10.1007/s10107-015-0893-2.

[8] NIELSEN H B. DAMPING PARAMETER IN MARQUARDT’S METHOD[J]. . Introduction.