

附件5B：

毕业设计（论文）开题报告

R语言corrplot包开发及在多元控制图中的应用

1、课题的目的及意义

1.1 课题的目的

1.1.1 掌握R packages的开发

通过学习掌握，如何设计，开发一个R packages，并参与到R packages 的后续维护，升级当中。提高学生的编程能力，团队意识，项目开发经验。

1.1.2. 结合相关矩阵，多元控制图，对质量检测的成本优化

了解当前工业工程领域中控制图的应用，了解多元控制图原理。在质量检测中，经常使用多元控制图对拥有多个检测属性的产品进行质量稳定性监控。可以利用相关分析减少检测变量的数量，达到减少质量检测的成本。

1.2 课题意义

1.2.1 开发R packages的意义

随着人工智能，大数据时代的到来，各个学科和数据科学都有不同程度的交叉，工业工程亦是如此。工业工程师在企业管理中会经常接触数据，有处理数据，分析数据的需求。R语言作为一门专门面向数据分析，挖掘的编程语言对工业工程师来说是一件解决问题的利器。

学习使用，制作和维护R packages，参与到R packages的生命周期当中可以提升编程能力，问题解决能力。更重要的是了解深入到数据科学这个领域当中去。

R作为一个开源的平台，又一大批爱好者组成的社区来支撑，越多的使用者，开发者加入对于R的生态有很好的提升。

1.2.2 结合相关分析研究多元控制图应用的意义

工业工程师经常使用控制图来进行质量控制等，使用控制图本身就是利用数据可视化的方式进行管理。但是当被检查对象的拥有多维变量时，无法使用单一变量进行衡量。为了使用更多维度的数据对被检查对象的描述，同时又能够在同一个图中观察，需

要对数据进行压缩。对数据的压缩会导致丧失一部分准确性，但是可以减少质量检测的成本。所以探讨该问题是一个非常有意义的。

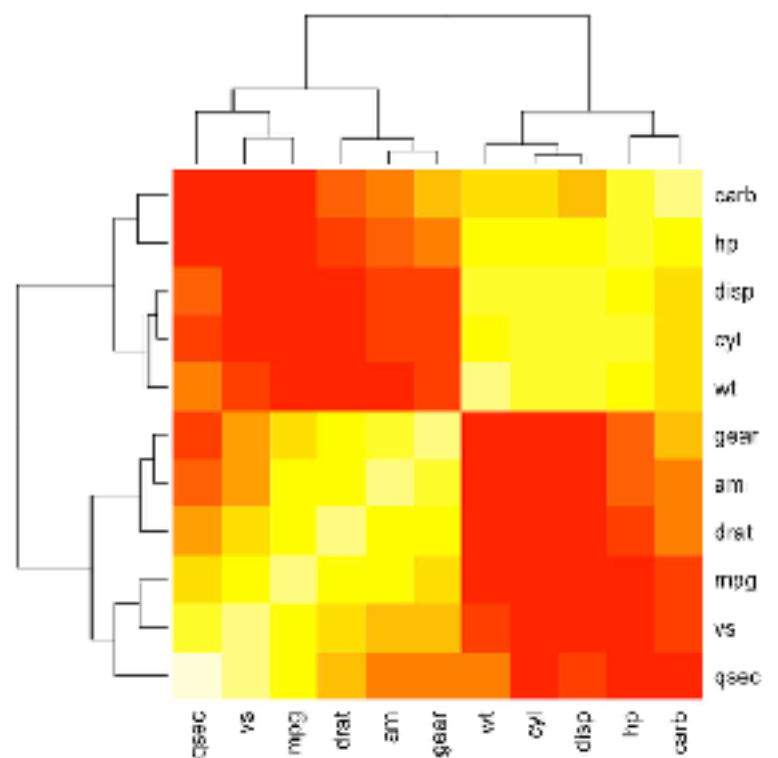
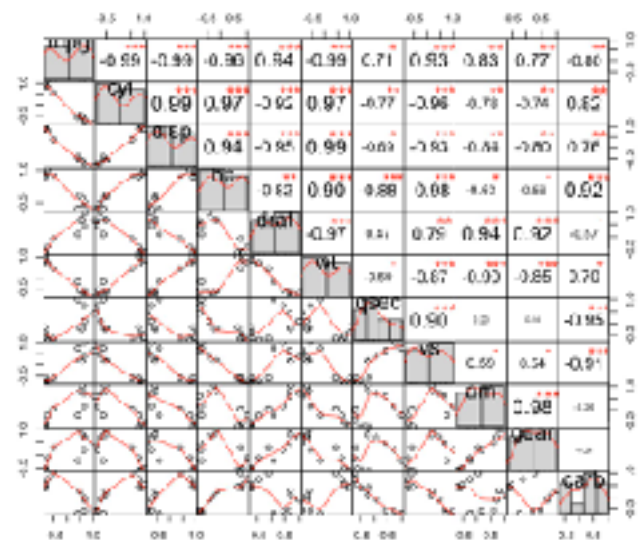
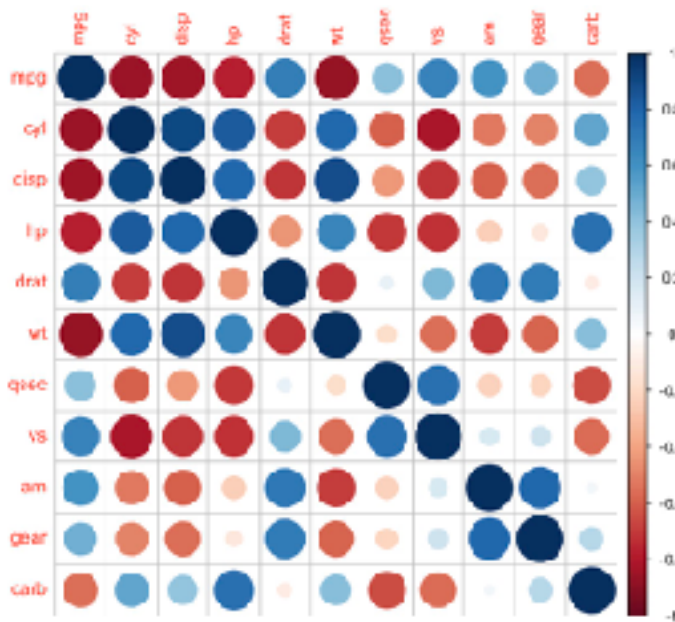
1.3 国内外研究现状

1.3.1 相关矩阵绘制现状

相关矩阵的绘制是矩阵绘制问题的一种。目前矩阵的绘制主要有两种方式：1.网络法（network） 2.网格法（grid）（参考文献）。网络法主要适用于矩阵稀疏时，用节点表示变量，边表示变量之间的关系，该方法适用于社交网络分析。网格法适合稠密的矩阵，适合相关矩阵的可视化。

	网格式方法	网络式方法
优点	<ul style="list-style-type: none">• 适合于所有数字矩阵• 不会出现节点重合、边交叉现象• 可以展示节点间较为复杂的关系	<ul style="list-style-type: none">• 仅限于关系矩阵• 直观性强、更紧密，更易读• 矩阵稀疏时更有效
缺点	<ul style="list-style-type: none">• 矩阵维数较高时，占用空间较多• 不如网络图直观易懂• 路径追踪难度大	<ul style="list-style-type: none">• 矩阵密集时容易出现节点、边大量重合• 需要较好布局算法防止节点、边重合• 节点间关系复杂时效果较差
共同点	<ul style="list-style-type: none">• 都依赖具体统计、数学模型为基础来探求变量、个体之间的关系，且这些关系发掘算法有很多是可以互通的	

目前相关矩阵的可视化R语言包有corrplot包，PerformanceAnalytics包，heatmap函数等。这些包各有特点，但是corrplot是专注相关矩阵可视化的R语言包，同时在相关矩阵可视化中有着较高认可度。



corrplot的功能分为两个层面，统计模型和展示层面。统计模型主要是指是对矩阵变量的顺序进行重排，以便发现规律。展示层面是指利用图形，颜色等计算机图形来清晰地表达矩阵的复杂信息。

corrplot诞生于数据可视化渐渐升起热潮的时代，在数据可视化快速发展的今天，corrplot绘制的图形需要具有可交互性便于使用者更好，更方便地使用。所以开发corrplot包的新版本是一件非常有趣，有意义的事情。

1.3.2 多元控制图现状

控制图在质量控制中用来衡量产品生产的稳定性。显然，如果对一件产品检查越多的属性，那么对这件产品度量准确性就越高，将多个变量数据用来衡量产品生产的稳定性就需要多元控制图。

引用参考文献（质量管理，检测方面；多元控制图方面）

参考文献：

2、课题任务、重点研究内容、实现途径

2.1课题任务

本课题有两个任务 1) 开发新版本的corrplot包，2) 结合相关分析使用多元控制图，提高控制图的应用效果。

1) 开发新版本的corrplot包

学习如何制作R packages，学习d3.js,html,css等前端知识，在github上和其他人一起合作完成。

2) 结合控制图，优化质量检测成本

利用corrplot的新版本成果，结合相关分析，在保证控制图中数据一定稳定性时，对生产过程中的质量检测进行一定优化。

2.2重点研究内容

由于该课题的内容主要分为两大块，R packages的开发和SPC领域

中的应用。由于时间和能力的限制，本科毕业设计不可能在这两个领域中同时作出非常成果，因此该课题将重点放在corrplot新版本的开发上。

2.3 实现途径

为实现需要的交互效果，corrplot新版本需要增加一些新的函数。新的函数需要建立在一个R packages htmlwidgets的基础之上，在htmlwidgets框架基础之上，通过d3.js对html的svg元素实现操作，实现矩阵可视化的绘图，布局，动态等。

3. 进度计划

序号	起止周次	工 作 内 容
1	1周至2周	学习文献检索方法，调研了解当前云制造、工艺路线优化的基本原理和方法，完成文献检索、综述及外文资料翻译
2	3周至4周	调研云制造模式下的生产制造等信息，并查阅文献掌握工艺路线优化的最新研究成果
3	5周至6周	参考当前工艺路线选择的方式，依据云制造的特点，设计云制造模式下工艺路线选择模式
4	7周至8周	通过历史数据，分析计算各制造资源制造能力
5	9周至13周	依据设计的工艺路线优化模式，结合案例进行分析
6	14周至15周	测试、修改该云制造模式下多工艺路线优选模式

7	16周至17周	完成毕业设计论文
8	18周	答辩

学生签名：

年 月 日

4、指导教师意见

指导教师签名：

年 月 日