

金融科技实践教育丛书

量化 投资 实践

(MATLAB版)

点宽金融科技实践教育丛书编委会 组编



科学出版社

金融科技实践教育丛书

量化投资实践(MATLAB 版)

点宽金融科技实践教育丛书编委会 组编

科学出版社

北京

内 容 简 介

本书中的案例来源于作者的实际工作,充分体现案例的实用性和程序的可模仿性,案例程序中附有详细的注释.例如 β , VaR 和 CVaR 的计算, Dual Thrust, R-break 和海龟交易策略等程序,读者可以直接使用或根据需要在源代码的基础上进行修改、完善.

本书分 5 篇,共计 23 章.第一篇介绍量化投资中常用的 MATLAB 编程知识和技巧;第二篇介绍几个基本的数理统计知识,包括随机变量、统计矩、极大似然估计、置信区间、假设检验、 p 值和 Spearman 秩相关性等;第三篇主要对线性回归、Kalman 滤波和 ARMA 等梳理模型进行概述;第四篇主要介绍量化投资中几个经典的策略模型,包括期货中的 R-breaker, Dual Thrust 和 Aberration 等经典策略,以及股票中的多因子模型;第五篇介绍关于风险管理的实践意义,主要实现了 VaR 和 CVaR 的计算.

本书可供具有金融数学或统计背景的科研工作者、经济金融机构的研究人员和从业人员、高等院校相关专业的教师和学生及对量化投资和 MATLAB 感兴趣的人士阅读.

图书在版编目(CIP)数据

量化投资实践: MATLAB 版/点宽金融科技实践教育丛书编委会组编. —北京: 科学出版社, 2019.6

金融科技实践教育丛书

ISBN 978-7-03-060607-5

I. ①量… II. ①点… III. ①MATLAB 软件-应用-投资学 IV. ①F830.5939

中国版本图书馆 CIP 数据核字(2019) 第 033491 号

责任编辑: 张中兴 梁 清 孙翠勤 / 责任校对: 杨聪敏

责任印制: 徐晓晨 / 封面设计: 迷底书装

科学出版社出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

北京建宏印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2019 年 6 月第 一 版 开本: 787 × 1092 1/16

2019 年 8 月第二次印刷 印张: 18 1/2

字数: 439 000

定价: 69.00 元

(如有印装质量问题, 我社负责调换)

PREFACE

前言

为了解决金融行业在定价和风险计量方面的量化处理问题,“金融数学”应运而生。从以研究理论为主的“金融数学”,到理论与实践结合的“金融工程”,再到今天最热门的学科“金融科技 (FinTech)”,人类从来都没有停止处理融合量化金融和高新尖科技问题的步伐。

随着再度升级的 AlphaGo 战胜了围棋世界冠军柯洁,“人工智能”又一次进入大众的视野。那么“人工智能”在金融市场中能否被应用呢?答案是毋庸置疑的。其实我们可以认为,量化投资就是金融市场的“AlphaGo”,它能够通过数量化的技术方式支持投资,已成为资本市场的前沿领域和重要的发展方向。在西方发达国家,量化投资经过 40 多年的发展,已经逐渐成熟。但是在中国,这个领域才刚刚起步。随着我国汇率、利率日益市场化,以及股指期货、融资融券、股票期权业务的相继推出,我国证券市场结构正在发生深刻变化,金融衍生品的定价、风险管理的手段都面临着巨大变革,投资理念和投资技术都趋向于精确化和数量化,我国的量化投资正处在不断发展变革的重要阶段。然而,国内在金融行业量化投资方面的专业人才极度匮乏,这已经成为我国量化投资发展的瓶颈。

近年来,国内高校与业界已经认识到量化投资人才培养的重要性和紧迫性,如何制订高校量化投资人才培养方案和实施方法已成为当前高校金融类相关专业教学改革的热点问题。

培养金融计量人才的一个最直接的方法就是让学生掌握一门可以量化的计算机语言,比如 MATLAB 或者 C++ 等,能够帮助学生实现金融定价和进行风险模型量化处理。我们认为,基于 MATLAB 的量化投资实践教程可依托高校金融数学专业课程建设和专业实验室丰富的软硬件资源,借助案例教学的优势以及教学考核方式的创新,显著地提升金融数学或金融工程专业教学的教学效果。依据此教学思路,我们推出基于 MATLAB 语言为工具的量化投资教材,用于培养能够解决实际金融问题的量化人才的教学或参考。

本书可以为学生提供模拟教学环境,让学生对量化投资和策略交易具有真实的体验。在编写本书的逻辑中,我们希望量化投资教学在培养学生理论联系实际的能力,塑造学生解决投资实际问题的思维方式,加强学生投资策略程序实现能力和工程化理念方面能起到本质的强化作用。另外,我们认为通过整合金融数学/金融工程类专业实验资源,开展量化投资实验课程的研究设计,其在提升学生的综合能力,强化学生对量化投资实务工作的适应能力,以及培养满足金融业实务需求的创新型和实践型人才方面起到推进作用。本书以 Auto-Trader 软件和 MATLAB 工具为依托,重点在于指导学生尽快掌握计量实现技能,因此本书不会把描述金融模型中各种理论基础的讲述作为教程的核心(但是为了方便本书的使用者,本书用到的模型理论基础部分全部基于下面的五本参考书),在每一篇,我们先陈述需要的金

融核心知识和模型, 重点在于利用 MATLAB 工具, 以最直接的方式培养学生掌握每章介绍的金融核心知识和模型, 特别是金融定价和风险模型, 并掌握落地执行的执行能力. 具体方式为, 结合证券市场实际案例向读者呈现金融数据分析的典型模型和量化投资的常用策略, 通过一系列的金融数据分析和量化策略编写实验, 帮助学生掌握金融建模的程序化建立和投资策略的执行过程, 并在此基础上进一步提高学生的量化分析和投资实战技能.

本书内容由五大部分组成, 包括:

- (1) 量化投资——基础研究工具: MATLAB 必备知识;
- (2) 量化投资——数理统计;
- (3) 量化投资——金融建模;
- (4) 量化投资——策略交易模型;
- (5) 量化投资——风险管理模型.

全书共有 23 章和 1 个附录, 教师可以根据授课内容和培养目标, 选择书中的部分章节或者全部章节进行实践指导. 由于 Auto-Trader 软件的升级可能会对书中代码和结果部分有所影响, 最新内容可通过扫描封底的“本书资源”二维码获取.

本书由深圳点宽网络科技有限公司 (以下简称点宽) 的量化研究部组织编写, 点宽量化研究部总监乐斌先生和同济大学袁先智教授担任点宽《金融科技实践教育丛书》主编.

最后, 感谢深圳数字动能信息技术公司 (以下简称数字动能) 量化研究部的倾情协助和科学出版社的大力支持, 同时感谢数字动能总经理黄嵩先生和点宽总经理徐辣女士对本书的支持. 有了他们的协助和支持, 本书才得以顺利完成. 本书在金融专业知识的提炼和融合方面, 可能存在理解不足和可商榷之处, 希望读者能提出宝贵建议, 以便加以改进和完善.

本书的五本核心参考文献:

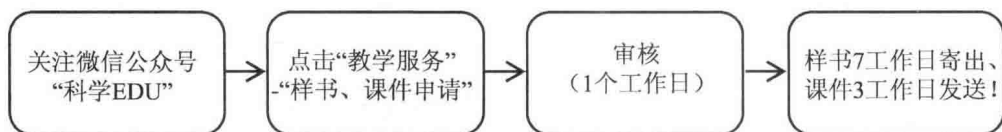
1. 原思聪. MATLAB 语言及应用. 北京: 国防工业出版社. 2011.
2. 王岩, 王爱青. 数理统计与 MATLAB 工程数据分析. 北京: 清华大学出版社. 2007.
3. Studenmund A H. Using Econometrics: A Practical Guide. 5th ed. Pearson. 2006. 中文译本为: 应用计量经济学. 5 版. 王少平, 杨继生, 刘汉中, 译. 北京: 机械工业出版社, 2007.
4. Hull J. Options, Futures and Other Derivatives. 9th ed. New York: Pearson. 2015. 中文译本为: 期权, 期货及其他衍生产品. 9 版. 王勇, 索吾林, 译. 北京: 机械工业出版社, 2016.
5. Hull J. Risk management and financial institutions. 3rd ed. New York: Pearson. 2011. 中文译本为: 风险管理和金融机构. 3 版. 王勇, 董方鹏, 译. 北京: 机械工业出版社, 2013.

科学出版社

教师教学服务指南

为了更好地服务于广大教师的教学工作，科学出版社打造了“科学 EDU”教学服务公众号，教师可通过扫描下方二维码，享受样书、课件、会议信息等服务。

样书、电子课件仅为任课教师获得，并保证只能用于教学，不得复制传播用于商业用途。否则，科学出版社保留诉诸法律的权利。



科学EDU

关注科学EDU，获取教学样书、课件资源
面向高校教师，提供优质教学、会议信息
分享行业动态，关注最新教育、科研资讯

学生学习服务指南

为了更好地服务于广大学生的学习，科学出版社打造了“学子参考”公众号，学生可通过扫描下方二维码，了解海量经典教材、教辅、考研信息，轻松面对考试。



学子参考

面向高校学子，提供优秀教材、教辅信息
分享热点资讯，解读专业前景、学科现状
为大家提供海量学习指导，轻松面对考试

教师咨询: 010-64033787 QQ: 2405112526 yuyuanchun@mail.sciencep.com

学生咨询: 010-64014701 QQ: 2862000482 zhangjianpeng@mail.sciencep.com

C CONTENTS

目 录

前言

第一篇 量化投资——基础研究工具：MATLAB 必备知识

第 1 章	MATLAB 必备知识	3
1.1	基础数据类型	3
1.2	常用数据类型	12
1.3	绘图	15
1.4	函数	17
1.5	平均值	19
1.6	离差分析	21

第二篇 量化投资——数理统计

第 2 章	离散和连续随机变量	27
2.1	离散随机变量	27
2.2	连续随机变量	31
2.3	分布拟合	39
第 3 章	统计矩——偏度和峰度	41
3.1	偏度	41
3.2	峰度	43
3.3	其他的标准矩	44
3.4	Jarque-Bera 正态检验	45
3.5	测试校准	45
第 4 章	极大似然估计	46
4.1	极大似然估计	46
4.2	正态分布的 MLE	47
4.3	正态分布的 MATLAB 实现	48
4.4	指数分布的 MLE 及 MATLAB 实现	49
4.5	案例研究——日回报的正态分布拟合	50

4.6	极大似然估计注意事项	50
第 5 章	置信区间与假设检验	52
5.1	置信区间	52
5.2	原假设和备择假设	53
5.3	假设检验的步骤	53
5.4	案例研究 1——工商银行的日回报率的假设检验	54
5.5	案例研究 2——假设检验用于平均值	58
5.6	案例研究 3——假设检验用于方差研究	60
第 6 章	p 值和多重比较偏差	63
6.1	p 值	63
6.2	案例研究——多次测试	64
6.3	敏感性和专一性的权衡	66
第 7 章	Spearman 秩相关性	67
7.1	Spearman 秩相关性	67
7.2	Spearman 案例	68
7.3	Spearman 秩相关系数用于公募基金夏普率研究	70

第三篇 量化投资——金融建模

第 8 章	期货和期货交易策略简介	75
8.1	远期合约	75
8.2	期货合约	76
8.3	期货与现货的关系	78
8.4	杠杆	81
第 9 章	线性回归	85
9.1	线性回归模型的概念	85
9.2	MATLAB 实现线性回归模型	85
9.3	线性回归与相关性	87
9.4	相关系数	88
第 10 章	多元线性回归	90
10.1	多元线性回归的概念	90
10.2	多元线性回归模型示例	90
10.3	多元线性回归预测建设银行股票价格	91
10.4	模型选择	93
第 11 章	单整、协整和平稳性	95
11.1	平稳性和非平稳性	95
11.2	单整	99
11.3	协整	108
11.4	总结	114

第 12 章	违背回归模型	115
12.1	残差	115
12.2	异方差	116
12.3	残差的序列相关性	123
12.4	Newey-West	125
12.5	多重共线性	126
第 13 章	Kalman 滤波	129
13.1	理论基础	129
13.2	Kalman 滤波模型	133
13.3	Kalman 滤波与迭代线性回归	135
13.4	Kalman 滤波发散	138
13.5	Sage-Husa 自适应滤波算法	144
第 14 章	ARMA 模型	148
14.1	基础概念及原理介绍	148
14.2	建立 ARMA 模型的一般步骤	150
14.3	案例研究——沪深 300 股指期货日收益率的 ARMA 模型拟合及 MATLAB 实现	151
第 15 章	过拟合的风险	155
15.1	什么是过拟合	155
15.2	例：选取过多参数	155
15.3	例：曲线拟合	156
15.4	例：回归参数	157
15.5	例：滚动窗口	162
15.6	避免过度拟合	170

第四篇 量化投资——策略交易模型

第 16 章	β 对冲	173
16.1	因子模型	173
16.2	风险暴露	173
16.3	风险管理	174
16.4	β 对冲的 MATLAB 实现	174
16.5	交叉对冲	175
第 17 章	配对交易	177
17.1	配对交易流程	177
17.2	配对交易策略	187
第 18 章	方向性交易策略构造	191
18.1	波动率突破策略	191

18.2	日内趋势反转策略之 R-breaker 策略	194
18.3	Dual Thrust 策略	198
18.4	Aberration 策略	206
18.5	海龟交易策略	210
第 19 章	多因子研究	219
19.1	常见因子	219
19.2	多因子模型的意义——不单单寻找表现最好的股票	220
19.3	数据处理	221
19.4	模型有效性检验	225
19.5	因子合成与降维	227
19.6	多因子研究案例	228
第 20 章	条件异方差模型	231
20.1	几种基础模型介绍	231
20.2	示例应用	234

第五篇 量化投资——风险管理模型

第 21 章	仓位集中风险	251
21.1	模拟“21 点”游戏	251
21.2	投资组合理论	252
21.3	资金约束	257
21.4	数学方法解释	257
21.5	额外的好处	258
第 22 章	最小线性相关算法	261
22.1	分散化投资	261
22.2	一些公式	261
22.3	最小线性相关算法用于投资权重分配	262
第 23 章	VaR 和 CVaR	264
23.1	VaR 和 CVaR 的定义	264
23.2	VaR 和 CVaR 的计算	266
23.3	VaR 和 CVaR 的计算演示	268
23.4	VaR 和 CVaR 运用于资产组合管理	271
参考文献		279
附录 Auto-Trader 交易软件使用手册		280

第一篇

量化投资——基础研究工具：MATLAB必备知识

本篇的目的是汇总学生在进行金融量化分析方面需要的基本工具：MATLAB语言的基本必备知识，它由一章组成。本篇包含MATLAB语言的基础数据类型、绘图功能、函数功能、基本的计量功能（包含平均值、离差分析）等相关内容。

如果读者希望对 MATLAB 语言有一个全面的了解，可以参见下面的经典文献：

1. 原思聪. MATLAB语言及应用. 北京: 国防工业出版社. 2011.
2. 周建兴, 岂兴明, 矫津毅. MATLAB 从入门到精通. 北京: 人民邮电出版社. 2008.

C HAPTER 1

第 1 章 MATLAB 必备知识

MATLAB 是一门实用性很强的编程语言, 广泛应用于许多科学实践中. 尤其在解决线性代数问题上, 处理数组和运算矩阵也十分便利. 本书的内容主要是基于 MATLAB 编写的, 因此本书的第 1 章主要介绍 MATLAB 的基础知识以及量化分析方面的几个核心概念.

1.1 基础数据类型

首先介绍一些基础的数据类型.

1.1.1 实数和虚数

实数直接用数字表示. 可用类型有

```
uint64/double/single/int8/int16/int32/int64/unit8/uint16/uint32
```

类型之间直接转换:

```
int32(1.1)
```

判断是否为某种类型:

```
isa(temp, 'double')
```

虚数的表示, 如将 $2i$ 赋值给 a :

```
>> a = 2i;
```

或

```
>> a = 2j;
```

单个数字的共轭加:

```
>> b = a';
```

查找相关方法:

```
>> methods('int32');
```

```
>> methods(11);
```

注意:

1. i 或 j 前面直接加数字表示虚数, 但是 i 和 j 前面不加数字也可以用来表示其他变量. 虚数表达不需要加 “*”;
2. `isa(variable,type)` 表示判断某个变量是否为某种类型;
3. `methods('int32')` 表示查找某种数据类型的所有可调用的函数. 注意直接写类型名字的时候要加引号, 或者直接改为某种类型下的一个实例.

1.1.2 向量 (数组)

一维数值数组. MATLAB 允许创建列向量和行向量, 列向量通过在方括号 `[]` 内把数值用分号 (行向量用逗号或空格) 隔开来创建, 对元素的个数没有限制.

列向量:

```
>> a = [2; 1; 4];
```

行向量:

```
>> a = [2, 1, 4];
```

数量乘法: 把一个向量的每个元素乘上一个数. 方法: 直接在 `[]` 外面乘上一个数, 或变量直接乘上一个数.

```
>> new_vector = vector*const;
```

转置:

```
>> new_vector = transpose(vector);
```

共轭转置: 用 “`'`” 表示, 如果向量元素均为实数, 共轭转置和 `transpose` 的效果一样.

```
>> new_vector = vector';
```

共轭不转置: `conj`

```
>> new_vector = conj(vector);
```

向量加减: 只有相同长度的行之间或列之间可以加减, 行与列不能直接加减 (可转置成同一类型后操作).

```
>> new_vector = vector1 - vector2;
```

向量元素乘除: 同上要求, 但需要在运算符前加 “`.`”

```
>> new_vector = vector1 .* vector2;
```

向量平方: 注意前面加 “`.`”

```
>> new_vector = vector1.^2;
```

向量点乘: `dot`

```
>> new_vector = dot(vector1,vector2);
```

向量合并: (同为行, 则用 “, ”; 同为列, 则用 “; ”)

```
>> new_vector = [vector1, vector2];
```

等差序列:

(1) 从 1 开始每隔 0.1 生成一个元素, 元素的最大值为 2.

```
>> new_vector = [1:.1:2];
```

(2) 从 1 开始到 2 输出 10 个等差序列, 使得最小值为 1, 最大值为 2.

```
>> new_vector = linspace(1,2,10);
```

(3) 对数值从 1 开始到 2 输出 10 个对数值等差序列, 使得最小对数值为 1, 最大对数值为 2.

```
>> new_vector = logspace(1,2,10);
```

或者

```
>> new_vector = 10.^(linspace(1,2,10));
```

长度:

```
>> len = length(vector);
```

最大值: max, 类似还有 min, sum, sqrt, std, mean, abs 等.

```
>> max_value = max(vector);
```

是否包含某元素:

```
>> ismember(element, vector);
```

数组交集:

```
>> intersect(vector1, vector2);
```

数组并集:

```
>> union(vector1, vector2);
```

数组相邻求差值, 结果与原数组少一个元素:

```
>> diff(vector);
```

复制:

```
>> repmat(a, [2,3]); % 将变量a行方向复制2次, 列方向复制3次
```

查找元素的个数:

```
>> a = [1,2,3,3];
```

```
>> histc(a,3) % 寻找a中3的个数
```

根据一个数组的条件修改另一个数组:

```
>> a = [2, 4, 5];
>> b = [4, 2, 3];
>> b(a>2 & a<5) =1;% b = [4 1 3].
% 括号内并条件只能用&不能用&&
```

注意:

1. MATLAB 识别中英文输入法, 分号等需要在英文输入法下键入, 否则报错;
2. 每个执行命令后面加入分号的目的是不输出结果. 如果不写分号, 在每次执行命令之后会把执行结果在执行界面显示.

1.1.3 特殊数组

```
a = zeros(3,4) %3行4列
b = zeros(4) %只有一个参数时, 生成二维数组
c = zeros(4,1)
d = zeros(3,4,5) %多维数组
a = ones(2) %全1
b = eye(2) %对角1
c = rand(2) %[0,1] 均匀分布随机数
d = randn(2) %正态分布随机数
e = false(2) %生成值全为1(true)或0(false)的逻辑数组
    %MATLAB中, 对逻辑数组的判断只有0和非0两种结果, 非0的数一律认为是true
f = nan(2) %生成值全为NaN的数组
```

1.1.4 矩阵的复制

```
a = [1 2;3 4]
disp('a a a;a a a')
disp([a a a;a a a])
disp('repmat(a,2,3) = ')
disp(repmat(a,2,3)) %按行复制2次, 按列复制3次
```

注意: disp 表示输出元素.

1.1.5 Kronecker 乘积

```
%C = kron(A,B), 用A的每个元素乘以B矩阵, 然后按A的元素顺序排列
a = [1 2;3 4]
b = ones(2,3)
disp('kron(a,b) = ')
disp(kron(a,b))
disp('kron(b,a) = ')
disp(kron(b,a)) %repmat 可以用kron 代替
```

1.1.6 获取元素

一维数组:

```
a = rand(1,8)
disp('a(1:2:end) = ') %end 表示最后一个元素，对于步长不为1的情况，如果不能整除，也不会报错
```

高维数组:

```
a = rand(4,5)
disp('a(3,4) = ')
disp(a(3,4))
disp('a(3,:) = ')
disp(a(3,:))
disp('a(:, [2,4]) = ')
disp(a(:, [2,4]))
```

1.1.7 删除数组元素

```
a = zeros(4,5);
for i = 1:numel(a)
a(i) = i;
end
a
a(3,:) = [];
disp('a(3,:) = [];')
a
a(1) = []; %对多维数组，如果a(1) = [], 将破坏多维结构
disp('a(1) = [];') a
```

1.1.8 排序

```
 %[Y,I] = sort(X,dim,mode)
 %mode: 'ascend' 升序'descend' 降序，默认前者
 %dim: 按第几维排序，默认是1，注意对矩阵，第1维是列，第2维是行，但对一维数组来说，默认只对那一维排序
 %如果X是一维数组，那么X(I) = Y，即Y为对X重新排列后的数组
 %如果X是多维数组，那么对于第dim维的每个子数组，X对应应在I上的取值就是Y
```

```
a = rand(1,8)
disp(['s,i] = sort(a)')
[s,i] = sort(a) %这里的排序其实按第2维
disp('a(i) = ')
disp(a(i))
b = rand(4,5)
disp(['~,i1] = sort(b(:,2),'descend');')
 [~,i1] = sort(b(:,2),'descend');
```