



CHINA
DATA
ANALYST
SUMMIT

CDA 数据分析师
www.cda.cn

R语言实现量化交易策略

用R语言解读股利贴现模型

演讲人：张丹

跨界互联
数聚未来

第四届中国数据分析师行业峰会
CHINA DATA ANALYST SUMMIT

北京 中国大饭店 2017. 07

前言

我们在分析证券市场的时候，总是试图找到价值**被低估的证券**。买进被低估的证券，等价值被修正时再卖出这些证券从而获利。

如何能才能找到，这些被低估的证券呢？

目录

1. 发现错误的定价
2. 股利贴现模型
3. 投资机会
4. A股市场案例分析
5. 用R语言实现

发现错误的定价

股票价格，合理不合理？



发现错误的定价

财务报表分析。

		2018-12-31	2017-12-31	2017-03-31	2016-12-31	2015-12-31	2014-12-31	2013-12-31	2012-12-31	2011-12-31	2010-12-31	2009-12-31	2008-12-31
报告期		年报	年报	一季报	年报	年报	年报	年报	年报	年报	年报	年报	年报
数据来源		盈利预测	盈利预测	合并报表	合并报表	合并报表	合并报表	合并报表	合并报表	合并报表	合并报表	合并报表	合并报表
日 利润表摘要													
营业总收入		14,639,370.71	13,148,087.24	3,003,526.46	11,011,310.19	10,056,445.36	14,000,539.40	12,004,307.00	10,011,010.88	8,351,725.25	6,080,724.25	4,263,729.11	4,219,971.57
同比(%)		11.34	19.41	19.67	9.50	-28.17	16.63	19.91	19.87	37.35	42.62	1.04	10.93
营业总成本				2,570,423.81	9,152,937.98	8,613,460.91	12,325,897.96	10,948,792.61	9,231,019.15	7,900,847.55	5,819,180.99	3,967,184.11	3,986,984.12
营业利润		2,456,798.60	2,143,286.07	468,764.69	1,745,569.78	1,351,617.70	1,608,922.73	1,226,301.01	802,630.79	454,230.64	274,678.18	297,452.53	233,918.65
同比(%)													
利润总额	报告期		2017-03-31	2016-12-31	2016-09-30	2016-06-30	2016-03-31	2015-12-31	2015-09-30	2015-06-30	2015-03-31	2014-12-31	2014-09-30
同比(%)	盈利能力		一季报	年报	三季报	中报	一季报	年报	三季报	中报	一季报	年报	三季报
净利润													
归属母公司股东	净资产收益率(摊薄)		6.94	28.63	22.57	14.26	6.23	26.37	22.15	13.91	5.91	32.06	24.67
同比(%)	净资产收益率(加权)		7.19	30.41	21.13	12.62	6.43	27.31	21.65	12.10	5.55	35.23	25.91
非经常性损益	净资产收益率(平均)		7.19	30.42	23.09	13.85	6.43	27.34	22.34	13.42	6.09	35.96	26.41
扣非后归属母公	净资产收益率(扣除／摊薄)		6.24	28.96	23.63	15.71	6.18	25.91	21.76	12.68	5.66	32.04	25.49
同比(%)	净资产收益率(扣除／加权)			30.77		13.91		26.84		10.99		35.20	
研发费用	净资产收益率(扣除／平均)		6.47	30.77	24.17	15.26	6.38	26.86	21.95	12.23	5.83	35.93	27.29
EBIT	净资产收益率-增发条件			30.41		12.62		26.84		10.99		35.20	
EBITDA													
日 资产负债表摘要	净资产收益率(年化)		28.74	30.42	30.79	27.71	25.74	27.34	29.79	26.84	24.38	35.96	35.21
流动资产	总资产报酬率		2.45	10.09	5.82	4.26	1.62	8.95	6.47	4.02	1.97	10.85	7.53
固定资产	总资产报酬率(年化)		9.81	10.09	7.76	8.52	6.50	8.95	8.62	8.03	7.87	10.85	10.04
长期股权投资	总资产净利率		2.12	9.02	6.58	3.86	1.90	7.94	6.14	3.60	1.75	9.83	6.76
资产总计	总资产净利率(年化)		8.47	9.02	8.78	7.73	7.61	7.94	8.19	7.21	7.00	9.83	9.01
	投入资本回报率		5.01	22.06	16.69	9.29	4.73	19.89	15.67	8.70	4.41	25.71	18.39
	人力投入回报率(ROP)		842.10	241.68	313.13	403.71	811.64	226.25	380.67	589.15	1,041.20	280.29	372.19

发现错误的定价



发现错误的定价

基本面分析：价值投资。

通过 现金流贴现，判断股价的高低。

派息：很多股票，每年会派发**股息**的，股价每股20块钱，每年会发1元红利给你。

基金派发红利，也是同理。

发现错误的定价

例如：股票A，假设持有1年，预期每股股利 $E(D1)=1$ 元，当前的股价为 $P0=20$ 元，1年后预期的股价为 $E(P1)=22$ 元，那么这只股票有没有吸引力呢？

$$\begin{aligned}\text{预期收益率 } E1(r) &= (\text{期初资产} - \text{期末资产}) / \text{期初资产} \\ &= \text{股利收益率} + \text{资本利得收益率} \\ &= (\text{股利} + \text{期末股价} - \text{期初股价}) / \text{期初资产} \\ &= (E(D1) + [E(P1) - P0]) / P0 = (1 + (22 - 20)) / 20 = 15\%\end{aligned}$$

当股价为， $E(P1)=19$ 元，

$$\text{预期收益率 } E2(r) = (E(D1) + [E(P1) - P0]) / P0 = (1 + (19 - 20)) / 19 = 0$$

发现错误的定价

资本资产定价模型(CAPM)：通过 **均值-方差**模型，计算市场任意投资组合的收益与某个共同因素之间是线性关系。

核心思想：**资产价格取决于其获得的风险价格补偿。**

当风险一样时，会选择预期收益最高的资产；而预期收益一样时，会选择风险最低的资产

参考文章：

R语言解读资本资产定价模型CAPM <http://blog.fens.me/finance-capm/>

发现错误的定价

$$\begin{aligned}\text{必要收益率为} k &= \text{无风险收益率} + \text{资产的市场风险} * [\text{投资组合收益率} - \text{无风险收益率}] \\ &= R_f + \beta * [E(R_m) - R_f] \\ &= 3.85\% + 1 * [10\% - 3.85\%] \\ &= 10\%\end{aligned}$$

判断： $E_1(r) > k > R_f > E_2(r)$, $15\% > 10\% > 3.85\% > 0$

所以：当股价1年后，为22元时，就更愿意**持有**这个股票，否则我们更愿意买 市场风险的投资组合。

发现错误的定价

当满足必要收益率，现在股价多少是合理的呢？

$$V_0 = (E(D_1) + E(P_1)) / (1+k) = (1 + 22) / (1 + 10\%) = 20.9099$$

实际现在的股价 $P_0=20$ 元， $P_0 < V_0$ ，则说明股票被低估了！

V_0 ：就是股票的内在价值， $P_0 < V_0$ ，股票价格 小于 股票内在价值。

目录

1. 发现错误的定价
2. 股利贴现模型
3. 投资机会
4. A股市场案例分析
5. 用R语言实现

股利贴现模型

DDM(Dividend discount model)，是股票估值中最基础的模型之一，基于公司给股东的分红来进行计算的。

分红代表的是实际上公司付给股东的现金流，这部分现金流的估值，能让投资者们了解到该公司股票的实际价值。

前提条件：该公司必须有分红。

股利贴现模型

根据分红发放的不同，可分成三种模型：

- 零增长股利贴现模型
- 固定增长率股利贴现模型
- 多段增长股利贴现模型

股利贴现模型

零增长股利贴现模型，公司每年派发固定的股息给股东，分红永久不变。

例如，公司股价20元，每股每年派息2元，固定每年都是2元。

用公式表达：

$$P_0 = \frac{Div_1}{(1+r)^1} + \frac{Div_2}{(1+r)^2} + \dots + \frac{Div_t}{(1+r)^t} + \dots$$

$$\text{Zero Growth: } P_0 = \frac{Div}{r}$$

股利贴现模型

固定增长率股利贴现模型：每年的股利都是增加的。

例如，公司股价20元，今年每股股息2元，明年股息2.1元，第3年股息2.205元，第4年股息2.31525，.....

股息每年按0.05进行增长，今年每股股息2元。

第2年： $2 \times (1 + 0.05)^1 = 2.1$

第3年： $2 \times (1 + 0.05)^2 = 2.205$

第4年： $2 \times (1 + 0.05)^3 = 2.31525$

定义g为股息一个固定的增长率， $g=0.05$

股利贴现模型

用公式表达：

$$\text{Constant Growth: } P_0 = \frac{\text{Div}}{r - g}$$
$$P_0 = \frac{\text{Div}}{1 + r} + \frac{\text{Div}(1 + g)}{(1 + r)^2} + \frac{\text{Div}(1 + g)^2}{(1 + r)^3} + \dots = \frac{\text{Div}}{r - g}$$

股利贴现模型

多段增长率股利贴现模型：每年的股利是不同的。

例如，公司股价20元，

- 第1年每股股息2元，第2年股息2.1元，第3年股息2.205元， $g1=0.05$ 。
- 第4年开始公司业务巨大增长，每年按 $g2=0.08$ 进行股息的增长。
- 第4年股息2.3815，第5年股息2.5719，.....

股利贴现模型

例子：AT&T

Year	Value	DPS _t or Terminal value (TV _t)	Calculation	Present value at 6.90%
0	DPS ₀ ¹	1.92		
1	DPS ₁	1.91	= 1.92 × (1 + -0.46%)	1.79
2	DPS ₂	1.91	= 1.91 × (1 + 0.10%)	1.67
3	DPS ₃	1.93	= 1.91 × (1 + 0.65%)	1.58
4	DPS ₄	1.95	= 1.93 × (1 + 1.21%)	1.49
5	DPS ₅	1.98	= 1.95 × (1 + 1.76%)	1.42
5	Terminal value (TV ₅)	39.28	= 1.98 × (1 + 1.76%) ÷ (6.90% - 1.76%)	28.14
Intrinsic value of AT&T's common stock (per share)				\$36.09
Current share price				\$38.03

$V_0 = 36.09$, $P_0 = 38.03$, 所以，内在价值 < 市场价格，股价被高估了。

目录

1. 发现错误的定价
2. 股利贴现模型
3. 投资机会
4. A股市场案例分析
5. 用R语言实现

投资机会

把未来的收益贴现到今天，核心是一个**贴现率**，也就是明天、后天的一块钱今天值多少钱。

反过来看，这个问题等价于今天投下去多少钱明天可以拿到一块钱。

因此，贴现率其实就是资产的预期收益率。

有效市场假说

弱式有效的情况下，**市场价格已充分反应出所有过去历史的证券价格信息**，包括股票的成交价、成交量、卖空金额，融资融券等；

推论一：如果弱式有效市场假说成立，则股票价格的**技术分析失去作用**，基本分析还可能帮助投资者获得超额利润。

有效市场假说

半强式有效市场假说，认为**价格已充分反应出所有已公开的有关公司营运前景的信息**。这些信息有成交价、成交量、盈利资料、盈利预测值，公司管理状况及其它公开披露的财务信息等。假如投资者能迅速获得这些信息，股价应迅速作出反应。

推论二：如果半强式有效假说成立，则在市场中利用**技术分析和基本分析都失去作用**，内幕消息可能获得超额利润。

有效市场假说

强式有效市场假说，认为价格已充分地反应了所有关于公司营运的信息，这些信息**包括已公开的或内部未公开的信息**。

推论三：在强式有效市场中，**任何方法都不能帮助投资者获得超额利润**，即使基金和有内幕消息者也一样。

投资机会

现实中很多人，一边盯着**K线图**一边收集公司的**基本资料**，然后研读宏观经济形势、产业资讯、政策变化，自学经济学、金融学教科书，不就是为了找到一只好股票，获取高一点的回报率？然而这是可能的吗？

- 人是感性的
- 收到消息有先有后
- 大家对市场的看法也不是完全一致

投资机会

实际上，是股票价格的过度波动，股价波动的程度远远超过基本面波动的程度。

股价过高或者过低的时候，未来的股价会回调或者反弹，导致未来回报率很低或者很高。

目录

1. 发现错误的定价
2. 股利贴现模型
3. 投资机会
4. A股市场案例分析
5. 用R语言实现

A股市场案例分析



A股市场案例分析

字段	2017	2016	2015	2014	2013	2012
预案公告日	2017/3/25	2016/3/31	2015/3/19	2014/3/29	2013/3/29	2012/3/29
除权除息日	2017/6/14	2016/7/13	2015/7/3	2014/7/11	2013/6/13	2012/6/7
基本情况						
方案进度	实施	实施	实施	实施	实施	实施
是否分红	是	是	是	是	是	是
分红对象	普通股股东	普通股股东	普通股股东	普通股股东	普通股股东	普通股股东
每股股利(税前)	0.74	0.69	0.67	0.62	0.63	0.42
每股股利(税后)	0.74	0.69	0.64	0.59	0.6	0.38
每股红股	--	--	--	--	--	--
每股转增	--	--	--	--	--	--
日期明细						
股本基准日期	2017/6/13	2016/7/12	2015/7/2	2013/12/31	2013/6/7	2011/12/31
股东大会公告日	2017/5/27	2016/6/29	2015/6/20	2014/7/1	2013/6/1	2012/5/31
股权登记日	2017/6/13	2016/7/12	2015/7/2	2014/7/10	2013/6/7	2012/6/6
实施公告日	2017/6/6	2016/7/5	2015/6/25	2014/7/3	2013/6/4	2012/6/1
派息日	2017/6/14	2016/7/13	2015/7/3	2014/7/11	2013/6/19	2012/6/13

A股市场案例分析

以 2016-07-13 除权日计算：

原始数据：

- 股利： $D0 = 0.69$
- 收盘价： $P0 = 17.05$
- 增长率： $g = (0.69 - 0.64)/0.64 = 0.078125$

计算2017年：

- 股利： $D1 = D0 \cdot (g + 1) = 0.69 \cdot (0.078125 + 1) = 0.7439062$
- 派息日收盘价： $P1 = 21.3$ (假设)
- 预期收益率： $R = (D1 + (P1 - P0))/P0 = (0.7439062 + (21.3 - 17.05))/17.05 = 0.2929$
- 派息日收盘价： $P1 = P0$ (假设)
- 预期收益率： $R = (D1 + (P0 - P0))/P0 = (0.7439062 + 0)/17.05 = 0.04363$

A股市场案例分析

以 2017-07-14 除权日计算：

原始数据：

- 股利： $D0 = 0.74$
- 收盘价： $P0 = 21.3$
- 增长率： $g = (0.74 - 0.69)/0.69 = 0.07246337$

预测2018年：

- 股利： $D1 = D0 \cdot (g + 1) = 0.74 \cdot (0.07246337 + 1) = 0.7936229$
- 派息日收盘价： $P1 = 24.85$ (假设, 2017-07-20)
- 预期收益率： $R = (D1 + (P1 - P0))/P0 = (0.7936229 + (24.85 - 21.3))/21.3 = 0.201578$
- 派息日收盘价： $P1 = P0$ (假设)
- 预期收益率： $R = (D1 + (P0 - P0))/P0 = (0.7936229 + 0)/21.3 = 0.037259$

A股市场案例分析

合并计算报表。

日期	股利	收盘价	增长率	预测-股利	预测-收盘价	预测-收益率
2014-07-11	0.59	9.69	-0.01667	0.58016	17.61	87.72%
2015-07-03	0.64	17.61	0.084745	0.69423	17.05	0.76%
2016-07-13	0.69	17.05	0.07812	0.7439	21.3	29.29%
2017-07-14	0.74	21.3	0.07246	0.7936	24.85	20.16%

目录

1. 发现错误的定价
2. 股利贴现模型
3. 投资机会
4. A股市场案例分析
5. 用R语言实现

用R语言实现

数据集：

- 股价数据 600036.csv
- 指数数据 000300sh.csv
- 分红数据 dividends.csv

```
> stock<-read.csv(file="600036.csv",header=FALSE,sep=";",fileEncoding="UTF-8")
> names(stock)<-c("date","close","adjclose")
>
>
> div<-read.csv(file="dividends.csv",header=FALSE,sep=";",fileEncoding="UTF-8")# 分红
> names(div)<-c("stock","date","divPreTax","divAftTax","divstock","increase")
>
> hs300<-read.csv(file="000300sh.csv",header=TRUE,sep=";",fileEncoding="UTF-8")
```

用R语言实现

股价数据 600036.csv

包括3列：

- date：交易日志
- close: 收盘价
- adjclose: 向前复权的收盘价

```
> tail(stock,15)
```

	date	close	adjclose
3624	2017-06-30	23.91	23.91
3625	2017-07-03	23.52	23.52
3626	2017-07-04	23.00	23.00
3627	2017-07-05	23.90	23.90
3628	2017-07-06	23.96	23.96
3629	2017-07-07	23.26	23.26
3630	2017-07-10	23.26	23.26
3631	2017-07-11	24.08	24.08
3632	2017-07-12	24.34	24.34
3633	2017-07-13	24.90	24.90
3634	2017-07-14	25.18	25.18
3635	2017-07-17	25.11	25.11
3636	2017-07-18	24.52	24.52
3637	2017-07-19	25.22	25.22
3638	2017-07-20	24.87	24.85

用R语言实现

分红数据 dividends.csv

包括6列：

- stock : 证券代码
- date : 除权日
- divPreTax : 税前股利
- divAftTax : 税后股利
- divStock : 红股
- increase : 转增

```
> div
```

	stock	date	divPreTax	divAftTax	divStock	increase
1	600036.SH	2017-06-14	0.74	0.74	NA	NA
2	600036.SH	2016-07-13	0.69	0.69	NA	NA
3	600036.SH	2015-07-03	0.67	0.64	NA	NA
4	600036.SH	2014-07-11	0.62	0.59	NA	NA
5	600036.SH	2013-06-13	0.63	0.60	NA	NA
6	600036.SH	2012-06-07	0.42	0.38	NA	NA
7	600036.SH	2011-06-10	0.29	0.26	NA	NA
8	600036.SH	2010-07-01	0.21	0.19	NA	NA
9	600036.SH	2009-07-03	0.10	0.06	0.3	NA
10	600036.SH	2008-07-28	0.28	0.25	NA	NA
11	600036.SH	2007-07-04	0.12	0.11	NA	NA
12	600036.SH	2006-09-21	0.18	0.16	NA	NA
13	600036.SH	2006-02-24	NA	NA	NA	0.09
14	600036.SH	2006-06-16	0.08	0.07	NA	NA
15	600036.SH	2005-06-20	0.11	0.10	NA	0.50
16	600036.SH	2004-05-11	0.09	0.07	NA	0.20
17	600036.SH	2003-07-16	0.12	0.10	NA	NA

用R语言实现

1. 股价与分红数据的整合
2. 计算指标
 1. 股利增长率
 2. 股价的收益率
 3. 预测-下期股利
 4. 预测-下期收盘价
 5. 预测-下期收益率
3. 根据资本资产定价模型，计算必要收益率
4. 根据股利贴现模型，计算股票内在价值
5. 比较内在价格与股票价格
6. 获得结论

用R语言实现

根据资本资产定价模型，计算必要收益率，持有1年。

- 无风险收益率： $R_f=0.038$
- 投资组合收益率： $R_m=0.128$ （沪深300指数）
- 招商银行市场风险： $\beta=0.55$

计算，必要回报率：

$$k = R_f + \beta(R_m - R_f) = 0.038 + 0.55 * (0.1159 - 0.038) = 0.080845$$

沪深300指数, $R_m = (3703.09 - 3282.87) / 3282.87 = 0.1280039$

- 000300.SH, 2016-07-13, 3282.87
- 000300.SH, 2017-07-14, 3703.09

用R语言实现

结果集：日期，收盘价，复权价，税后股利，股利增长率，复权收益率，下一年股利，下一年预期收益率，股利收益率，内在价值

```
> df
```

	date	close	adjclose	divAftTax	divInc	ret	D1	R	R2	V0
1	2003-07-16	12.10	2.64	0.10	NA	NA	NA	NA	NA	NA
2	2004-05-11	8.94	2.36	0.07	-0.30000000	-0.106060606	0.0490000	-0.302125280	0.005480984	5.737000
3	2005-06-20	6.19	2.48	0.10	0.42857143	0.050847458	0.1428571	0.179783060	0.023078698	6.715258
4	2006-06-16	7.16	3.66	0.07	-0.30000000	0.475806452	0.0490000	0.351815642	0.006843575	8.900213
5	2006-09-21	9.63	5.02	0.16	1.28571429	0.371584699	0.3657143	1.522919448	0.037976561	22.340843
6	2007-07-04	23.93	12.53	0.11	-0.31250000	1.496015936	0.0756250	0.037008985	0.003160259	22.818921
7	2008-07-28	24.74	13.10	0.25	1.27272727	0.045490822	0.5681818	-0.225619167	0.022966120	17.616685
8	2009-07-03	18.59	12.85	0.06	-0.76000000	-0.019083969	0.0144000	-0.310145239	0.000774610	11.792529
9	2010-07-01	12.81	9.50	0.19	2.16666667	-0.260700389	0.6016667	0.044626594	0.046968514	12.304957
10	2011-06-10	12.78	9.69	0.26	0.36842105	0.020000000	0.3557895	-0.091878758	0.027839552	10.671970
11	2012-06-07	11.25	8.85	0.38	0.46153846	-0.086687307	0.5553846	0.144478632	0.049367521	11.839411
12	2013-06-13	12.32	10.18	0.60	0.57894737	0.150282486	0.9473684	-0.136577239	0.076896787	9.781469
13	2014-07-11	9.69	8.68	0.59	-0.01666667	-0.147347741	0.5801667	0.877210182	0.059872721	16.726557
14	2015-07-03	17.61	16.37	0.64	0.08474576	0.885944700	0.6942373	0.007622788	0.039422901	16.316508
15	2016-07-13	17.05	16.49	0.69	0.07812500	0.007330483	0.7439062	0.292897727	0.043630865	20.270219
16	2017-06-14	21.30	21.30	0.74	0.07246377	0.291691935	0.7936232	0.201578553	0.037259305	23.580297

用R语言实现

结果集：日期，收盘价，复权价，税后股利，股利增长率，复权收益率，下一年股利，下一年预期收益率，股利收益率，内在价值

```
> df[15,]
      date close adjclose divAftTax  divInc      ret      D1      R      R2      V0
15 2016-07-13 17.05   16.49     0.69 0.078125 0.007330483 0.7439062 0.2928977 0.04363087 20.27022
>
>
>
>
> df[16,]
      date close adjclose divAftTax  divInc      ret      D1      R      R2      V0
16 2017-06-14  21.3    21.3     0.74 0.07246377 0.2916919 0.7936232 0.2015786 0.0372593 23.5803
```

用R语言实现

找到 2010-01-01之后的，招商银行股价 被低估的时期，即 $V0 > \text{close}$

```
> df %>%  
+   subset(date>as.Date('2010-01-01')) %>%  
+   subset(V0-close>0)
```

	date	close	adjclose	divAftTax	divInc	ret	D1	R	R2	V0
11	2012-06-07	11.25	8.85	0.38	0.46153846	-0.086687307	0.5553846	0.1444786	0.04936752	11.83941
13	2014-07-11	9.69	8.68	0.59	-0.01666667	-0.147347741	0.5801667	0.8772102	0.05987272	16.72656
15	2016-07-13	17.05	16.49	0.69	0.07812500	0.007330483	0.7439062	0.2928977	0.04363087	20.27022
16	2017-06-14	21.30	21.30	0.74	0.07246377	0.291691935	0.7936232	0.2015786	0.03725930	23.58030

结论

寻找被低估的证券，即使极小的错误定价，也能让金融分析师获利。中国的股票，已经越来越多的开始参与分红，与股民分享企业成长的成果。

作为数据分析师，我们**懂技术、懂数据、懂统计、懂金融**！那么拿起手中的武器，去金融市场**抢钱**吧。

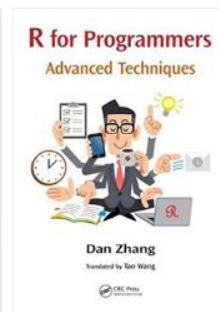
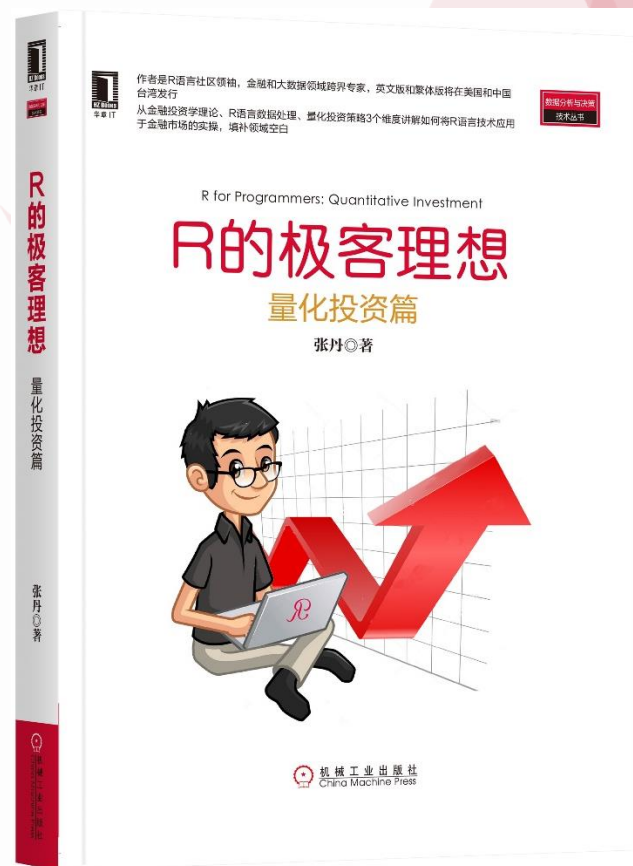
希望大家跟着我的思路，通过R语言入手，体会R语言建模的乐趣，并真正地把技术变现。我认为R语言是非常好的一门语言，从量化投资的角度，我会一直推动R语言在中国的发展，当成一项事业来做。

作者介绍

《R的极客理想》系列图书作者

联系方式:

- Blog: <http://fens.me>
- Email: bsspirit@gmail.com
- QQ群: 383275651





CDA 数据分析师
www.cda.cn

THANKS

跨界互联 数聚未来

第四届中国数据分析师行业峰会
CHINA DATA ANALYST SUMMIT