BARRA风险因子构建及数据提取

--by LBL

一、主旨

本文档主要介绍了BARRA风险因子载荷数据的计算方法，包括1.因子计算所需基础数据。2.因子计算方法。3.标准化处理方法。部分基础数据给出了所在的WIND数据库名，其他未标明的因子可以通过数据库检索文档查询。

二、因子计算所需数据及算法

PS.所取数据均按照括号内标注英文名存储

1.SIZE:

1.1.所需数据：总市值（totalMarketCap）

1.2.计算方法：SIZE=log(totalMarketCap)

2.BETA:

2.1.所需数据：

a.股票收益率（returnStock,按照复权收盘价对复权收盘价计算）

b.中证全指收益率（000985）（returnZZQZ, 按照复权收盘价对复权收盘价计算）

2.2.计算方法

计算每只股票(3000只?)股价对中证全指收益序列的BETA值，即计算股价序列对指数序列的回归系数，

*(r\_market?)*

计算时回归回溯期T=252天，回归半衰期lag=63天。若有效日期小于60天则该因子值为空。

3.GROWTH：

3.1.所需数据

a.EGRLF:长期净利润增速，未来三年净利润增速（分析师一致预测）

b.EGRSF:短期净利润增速，未来一年净利润增速（分析师一致预测）

c.EGRO:过去五年净利润增速（转化到年化增速）

d.SGRO:过去五年主营业务收入增加（转化到年化增速）

3.2.计算方法

GROWTH=0.18\* EGRLF+0.11\*EGRSF + 0.24\*EGRO + 0.47\*SGRO

4.LIQUIDITY:

4.1.所需数据：

a.股票每天成交量（volStock）

b.股票流通股本(freeFloatStock)

4.2.计算方法：

STOM=

STOQ=,其中T=3

STOQ=,其中T=12

LIQUIDITY=0.35\*STOM+0.35\*STOQ+0.30\*STOA

5.BP

5.1.所需数据：

a.净资产（股东权益合计（不含），netAsset）

b.总市值（totalMarketCap）

5.2.计算方法：

BP=

6.EARNINGYIED

6.1.所需数据

a.当前净利润（netProfit,在wind数据库ASHARECASHFLOW里，列名为net\_Profit）

b.下一年净利润预测（estNetProfit, 在wind数据库ASHARECASHFLOW里，列名为est\_net\_Profit）

c.过去12个月净利润（pastNetProfit）

d.个股现金收益（cashProfit即经营活动产生的现金流，在wind数据库ASHARECASHFLOW里，列名为net\_cash\_flows\_opera\_act）

e.总市值（totalMarketCap）

6.2.计算方法

EPFWD=( netProfit+ estNetProfit)/(2\* totalMarketCap)

CETOP=cashProfit/ totalMarketCap

ETOP= pastNetProfit/ totalMarketCap

EARNINGYIED=0.68\* EPFWD+0.21\* CETOP +0.11\* ETOP

7.LEVERAGE

需要因子：

1. 总市值（totalMarketCap）

以下b,c,d,e,f,g数据均在数据库资产负债表里。

1. 长期借款（longBorrow,wind列名为lt\_borrow）
2. 长期应付款（longTermPayAble,wind列名为lt\_payable）
3. 应付债券（bondsPayable，wind列名为bonds\_payable）
4. 总资产（totalAsset,wind列名为tot\_assets）
5. 总负债（totalBond，列名为tot\_liab）
6. 股东权益合计（netAsset,即净资产）

计算方法:

LD= longBorrow+ longTermPayAble+ bondsPayable

ME= totalMarketCap

TD= totalBond

TA= totalAsset

BE= netAsset

MLEV=(ME+LD)/ME

DTOA=TD/TA

BLEV=(BE+LD)/BE

LEVERAGE=0.38\*MLEV + 0.35\*DTOA + 0.27\*BLEV

8. Momentum

8.1.所需数据

a.股票收益率（returnStock,按照复权收盘对复权收盘计算）

8.2.计算方法

若回溯计算有效数据不足40天则为空。

Momentum=RSTR

9.NONLINEARSIZE

9.1.需要因子

a.总市值（totalMarketCap）

9.2.计算方法

a.标准化总市值（标准化方法同第三部分）

b.对市值做三次方，将该数据对同样标准化处理后的总市值做根号市值加权的回归，得到残差。

c.对残差再次标准化，得到因子值。

10.RESDUALVOL

10.1.需要因子

a.股票收益率（returnStock）

10.2.计算方法

DASTD：日频收益波动率

其中收益率序列长度取 252 个交易日，半衰期设定为 42 日。

CMRA: 月收益率极差

先计算股票过去T=12个月的累计收益Z（T）：

其中是t月的股票收益（由过去21个交易日累计得到），是无风险利率。于是极差可以表示为

其中

HSIGMA:残差波动率

其中为计算BETA中所得，即回归完市场收益后的股票收益。

RESDUALVOL =0.74\*DASTD + 0.16\*CMRA + 0.10\*HSIGMA

三、标准化方法

对风格因子计算时的数据处理主要包括去极值，空缺值填充与标准化三部分。

（1）去极值：将三倍标准差以外的数据拉回到三倍标准差的位置。

（2）空缺值填充：对因为数据原因无法计算的因子值采用能计算的同行业股票的因子值的中位数进行填充。

（3）标准化：采用市值加权的标准正态法即：