**练习题部分不需要上交，正式作业2人一组，2017年12月13日晚上11点前提交助教**

**必做题都要做，选做题至少选2题**

# 第一部分 练习题

**数据说明**

**Date：日期，从2005/01/04-2016/12/29**

**ClosePrice000001.mat：上证综指2005/01/04-2016/12/29的每天收盘价，但是日期为字符格式**

**ClosePriceADJ.mat：包含3172支股票每天日期、收盘价和股票代码、名称**

**HighPriceADJ.mat：包含3172支股票每天日期、最高价和股票代码、名称**

**LowPriceADJ.mat：包含3172支股票每天日期、最低价和股票代码、名称**

**Industry.mat：包含3080支股票的代码、名称和其所属行业**

**week\_end\_date.mat：2005/01/04-2016/12/29期间每周最后一个交易日的日期**

**month\_end\_date.mat：2005/01/04-2016/12/29期间每月最后一个交易日的日期**

**year\_end\_date.mat：2005/01/04-2016/12/29期间每年最后一个交易日的日期**

**profit.mat：3053只股票的每个季度的净利率、股票代码、名称和每个季度最后一个交易日的日期**

**Ret000001.mat：上证综指2005/01/04-2016/12/29期间每天的收益率**

**Rf.mat：2005/01/04-2016/12/29期间每天的无风险利率**

**Stk\_hs300.mat：沪深300指数成分股的股票代码和名称**

**一、基础题**

**A、熟悉常用命令（find，sort，isnan，isempty，isfinite，isinf，is\*，ismember，setdiff，setxor等）**

**1、利用ClosePriceADJ.mat数据，计算各个股票每天的收益率和最近3天（包括当天）的累积收益率；**

**2、利用ClosePriceADJ.mat、HighPriceADJ.mat、LowPriceADJ.mat数据，找出2016/07/28那天涨停、跌停、停牌的股票；（命令：find，停牌的股票当天的最高价等于最低价）**

**3、利用Stk\_hs300.mat和Industry.mat，找出HS300指数成分股中金融类和非金融的公司；（命令：ismember）**

**4、利用ClosePriceADJ.mat数据**

**（1）将ClosePrice数据中的NaN值变为0；（命令：find）**

**（2）将ClosePrice数据中含有NaN值的股票删除；（命令：find和unique）**

**5、利用ClosePriceADJ.mat数据，找出2016/07/29那天最近3天（包括2016/07/29）出现过涨停的股票；（命令：find和union）**

**7、利用ClosePriceADJ.mat、HighPriceADJ.mat、LowPriceADJ.mat数据，找出2016/07/29涨幅前10的非涨停的股票；（命令：find和sort）**

**8、利用ClosePrice000001.mat数据，将里面日期数据格式变成数字格式：（命令：datestr和str2num）**

**9、利用Ret000001.mat数据，计算其每天波动率，每天的资金净值曲线，并画图；（命令：std，ret2tick，plot）**

**B、常见循环**

**10、利用ClosePriceADJ.mat、week\_end\_date.mat、month\_end\_date.mat、**

**year\_end\_date.mat：计算各个股票每周、每月、每年的累积收益率。**

**11、利用Ret000001.mat、Rf.mat数据，计算每年的夏普比率，年化收益。**

**12、利用ClosePriceADJ.mat数据，根据每月股票收益率排名，将其分成10组，得出每组的股票的名称和代码。**

# 第二部分 正式作业

**数据说明**

**(1)ClosePrice.mat:包含2935支股票2006.12.29~2016.5.31的收盘价格数据，交易日期数据和股票代码数据**

**(2)PE.mat:包含2935支股票2006.12.29~2016.5.31的pe数据，交易日期数据和股票代码数据**

**(3)PB.mat:包含2935支股票2006.12.29~2016.5.31的pb数据，交易日期数据和股票代码数据**

**(4)ROA.mat:包含2935支股票2006.12.29~2016.5.31的roa数据，交易日期数据和股票代码数据**

**(5)size.mat:包含2935支股票2006.12.29~2016.5.31的size数据，交易日期数据和股票代码数据**

**(6)DateMon.mat：为2006.12.29~2016.5.31期间每个月的最后一个交易日的日期。**

**(7)Hat.mat：包含2935支股票2006.12.29~2016.5.31的摘牌数据，交易日期数据和股票代码数据**

**(8)RmtRf为从1994.4~2016.10月度和的数据**

**(9)** **fivefactor\_monthly.csv:包含中国A股市场1994.2-2017.3的五因子月度数据**

# 指数基金复制策略

**1、（必做）增强型指数基金的目标函数需要在跟踪误差和超额收益之间进行权衡，达到在跟踪指数的基础上尽可能的战胜市场。**

**J. E. Beasley、N. Meade和T. J. Chang（2001）等人对这个问题进行了求解，他们在其构建的模型优化目标包括两个部分：**

**（1）跟踪误差**

****

**（2）超额收益：**

****

**其中为追踪组合收益率，为标的指数收益率（和都是单期收益率，不是累积收益率）。**

**假设指数基金的目标函数如下：**

****

**其中，代表追踪误差和超额收益之间存在一种相互取舍的关系。代表只关心追踪误差，代表只关心超额收益。**

**约束条件为：**

**组合内单只股票的持有数量不得超过一定的比例，也不能小于一定的比例，其中为追踪组合中个股i的最小持有比例， 为追踪组合中个股i的最大持有比例，则有，限制了追踪组合中个股i的风险敞口。**

**请编写增强型指数基金代码。输入为被跟踪指数的收益率、跟踪指数的所有股票的收益率矩阵、参数lamda、不同股票的权重的下限和上限（采用矩阵形式），输出为增强型指数基金的权重。**

**参考完全复制型指数基金策略：**





s.t. 

 

因为是固定的，所以





令，其中 ，，

则



 为一个数，所以，因

为常数，与无关

 

对应于x=quadprog(H,f,A,b,Aeq,beq)

H=2\*R’\*R; f=2\*IR’\*R; A=-eye(N); b=zeros(N,1); Aeq=ones(1,N); beq=1;

**动量和反转策略**

**1、（必做）采用日度数据，检验A股市场惯性策略在短期是否可行？找出不同(j,k)策略的收益情况？找出收益最高的(j,k)策略的收益？（要求在J+1时刻排序，J+2时刻买入，J+K+2时刻卖出，涨停不能买入，跌停不能卖出，不考虑交易费用，投资组合持有股票个数自定，涨停不能买入的股票需要从后续的股票递补）**

**分成三个函数：**

**第一个函数：输入为J+K+2\*N的收盘价价格矩阵，此矩阵元素不存在NaN，得到一个赢家的收益率和输家的收益率**

**第二个函数：输入为T\*N的收盘价价格矩阵，矩阵元素存在NaN，需要在给定时刻删去有NaN的股票。调用第一个函数，第1行到J+K+2行得到一个赢家的收益率和输家的收益率，第2行到J+K+3行得到一个赢家的收益率和输家的收益率，以此类推，得到多个赢家的收益率和输家的收益率，进一步得到收益率的均值**

**第三个函数：输入为T\*N的收盘价价格矩阵，让排序期J和持有期K都发生变化，得到一个（J，K）赢家收益率矩阵，得到一个（J，K）输家收益率矩阵，得到一个（J，K）赢家-输家收益率矩阵。**

**提示：ST股票也在投资范围之内，ST股票的涨跌停幅度为5%**

**2、（选做）采用一维排序(C/P，E/P)验证价值在A股市场是否可行？找出价值投资的最佳持有期？计算持有期内的累积收益？（不允许卖空）**

**CAPM模型检验**

**（1、2选做一题）**

**1、（选做）采用中国股票市场的数据，编写程序采用一维排序法检验CAPM在横截面是否成立？**

**2、（选做）采用Fama Macbeth方法检验CAPM模型在横截面上是否成立？**

**3. （必做）查找周期性行业的定义，选择一个完整的周期的数据，比较周期性行业指数的beta与非周期行业指数的beta的大小，是否符合预期，请给出经济解释。**

**4.（必做）请搜集中国A股市场的各个行业股票收益率的数据，估计各个行业的beta值。使用这些行业的数据，进行GRS（Gibbons, Ross and Shanken）检验**

**给出F统计量与其P值，拒绝还是接受CAPM模型？**

# 三因子模型

**1、（选做）将中国A股的股票按照流通市值（size）分成2组，在每个市值组里面，再按照PB分成3组，构造SMB和HML因子，频度为月度。需要提供原始数据和程序，程序要一键运行。给出SMB随着时间变化的图形，给出HML随着时间变化的图形，并进行文字描述。**

**2、（必做）统计SMB和HML因子的平均值、标准差、偏度、峰度，构造t检验，看SMB和HML是否平均意义上为正。统计这两个因子在每个月份的平均值，是否存在规律？（采用fivefactor\_monthly.csv数据）**

**3、（选做）找到市场风险溢价（rm-rf）的月度时间序列数据，给出rm-rf，SMB，HML的方差协方差矩阵，看看从单个因子看，哪个因子的波动性更大，不同因子的相关性如何？（采用fivefactor\_monthly.csv数据）**

**4、（必做）给定一个股票投资组合，随意选定三只股票。**

**（1）请编写一个函数，函数输入：股票价格矩阵，权重向量，置信水平。函数输出：持有1块钱时投资组合的VaR。**

**（2）设定置信水平95%，利用提供的日频数据，分别用delta normal和历史模拟法，检验三个资产收益率序列是否为正态分布，并画出经验分布图。（正态分布检验方法为J-B统计量，请自行查找资料）**

**（3）用公式推导计算三个资产的边际VaR以及成分VaR，计算方法参照课件。**

**5、（必做）选取多个行业指数，采用FF三因子模型回归，给出回归系数？计算每个行业在三个因子上的风险，及三个因子占总风险的比例，风险指标为VaR(Value at Risk)。（行业指数数据自行下载）**

**6、（必做）给定一个size规模因子矩阵，数据中含有缺失值，请计算对应的Z值并输出一个相同大小的Z值矩阵，原数据中的缺失值保留。**

**7、（必做）在打分相加法、Z值排序法任选一种方法，采用PE、ROA、Size进行选股，每月调仓一次，涨停不能买入、跌停不能卖出、停牌不能买卖，不能做空，假设交易成本为0，每次买入50只股票，权重为等权重，请画出两条策略的净值曲线。给出收益率、标准差、夏普比率、最大回撤等指标，并说明可能的策略优化方向。**