期末实验报告

注意事项:

1. 本题限时90分钟完成.

2. 按要求提交MATLAB程序文件和结论图表, 不要提交试题附件的原始数据.

3. 必须使用MATLAB软件作为分析工具.

4. 提交每大题的程序文件为Q1\_171480101.m 依题号和学号类推.

5. 提交题目中提到的结论图表为 fig\_171480101.png 依学号类推, fig替换为题目中要求的图表名称.

**1. 附件QTTN.xlsx是上证50股票2016~2019年的行情数据, 附件IDX.xlsx是上证50指数2016~2019年的行情数据, 根据材料完成本题. (合计7分)**

(1) 在Microsoft Excel中适当整理数据, 方便导入MATLAB. (不要将表格作为答案文件回传)

(2) 将收盘价读入MATLAB, 共有975个交易日, 50个股票和1个指数, 无缺失值. 指数读成975\*1的向量, 股票数据读成975\*50的矩阵, 写出读取xlsx电子表格并将数据整理成指定形式的程序. (5分)

(3) 求指数和股票的对数收益率, 整理成974\*1, 974\*50的向量形式. (2分)

**2. 求50个股票的相关统计指标, 完成本题. (合计9分)**

(1) 求50个股票的期望收率. (2分)

(2) 求50个股票的风险. (用标准差表示, 2分)

(3) 从IDX.xlsx电子表格中统计2016~2019年每年的交易日数量, 以备题目使用.

(4) 如果2016-1-4买入股票, 2019-12-31卖出股票, 求50个股票4年的累计收益率和2016~2019年的每年累计收益率. (5分)

**3. 建立CAPM模型, 完成本题. (合计14分)**

(1) 根据2018年的数据, 用经典的CAPM模型估计50个股票的值. 无风险资产的收益率以1.5%计算. (4分)

(2) 用所求的值预测2019年的年收益率, 并与真实值比较. 本小题需另外提交预测值与真实值之间差距的直方图, 图片名称为histA. 横坐标为收益率预测误差, 纵坐标为区间中股票数量. (10分)

**4. 计算投资组合的有效前沿, 完成本题. (合计25分)**

(1) 计算50个股票的协方差. (2分)

(2) 用已经计算的2016~2019年的总收益率和协方差, 计算用50个股票构建的有效前沿. 求100种方案对应的收益率, 风险和50个股票在方案中的权重. (10分)

(3) 绘制有效前沿的图像, 图片名称为riskA. (2分)

(4) 若每个股票都必须购买, 每个股票购买不超过投资组合的10%, 重新计算有效前沿, 并绘制图像, 图片名称为riskB. (11分)

**5. 用统计机器学习方法分析股票收益率和指数收益率, 完成本题. (合计10分)**

(1) 用线性回归2018年的数据, 估计上证50指数中成分股的权重. 输出柱状图, 图片名称为weightA. 每个柱形是1个股票的权重. (5分)

(2) 用随机森林算法, 通过2016~2018年的数据训练随机森林回归器, 自变量为50个股票的收益率, 因变量为指数的收益率.

(3) 预测2019年上证50指数的每日收益率, 与真实值比较并作出直方图, 图片名称为histB. 横坐标为收益率预测误差, 纵坐标为区间中股票数量. (5分, 若误差较小的另加5~15分)

(4) 输出随机森林模型中, 各股票收益率的特征重要性, 作为成分股的权重. 输出柱状图, 图片名称为weightB. 每个柱形是1个股票的权重. (完成另加5分, 若误差较小的另加5~15分)

**6. 建立金融风险VaR模型, 完成本题. (合计35分)**

(1) 用2016~2018年的数据训练GARCH模型, 在程序中写出AIC, BIC循环定阶的过程, 定阶和训练模型的程序命名为garch, 不放在主要代码中. (训练GARCH模型10分, 定阶过程5分)

(2) 将GARCH模型导出为mat文件, 以备使用.

(3) 用2016~2018年的数据, 假设收益率服从正态分布, 用移动平均法计算金融风险VaR. 窗口长度为250. 绘制VaR折线图, 图片名称为varA. 一条折线为收益率, 另一条折线为1% VaR风险下限. (10分)

(4) 用训练好的GARCH模型预测2019年的收益率, 用Monte-Carlo模拟预测10000条路径. 假设预测的收益率服从正态分布, 计算VaR并绘制折线图, 图片名称为varB. (10分)