浦发银行和中国联通股票分析

一、 利用 quantmod 包下载同一市场上两只股票五年价格数据，得到两只股票的日收益率、月收益率数据。

利用 quantmod 下载两只沪市 A 股股票，分别为浦发银行（PF）和中国联通（ZGLT）。选取数据开始时间为 2014 年 4 月 19 日，结束时间为 2019 年 4 月 19 日，共约 5 年；分别计算两只股票的对数日收益率、月收益率。

绘制时序图如图 1 和图 2：

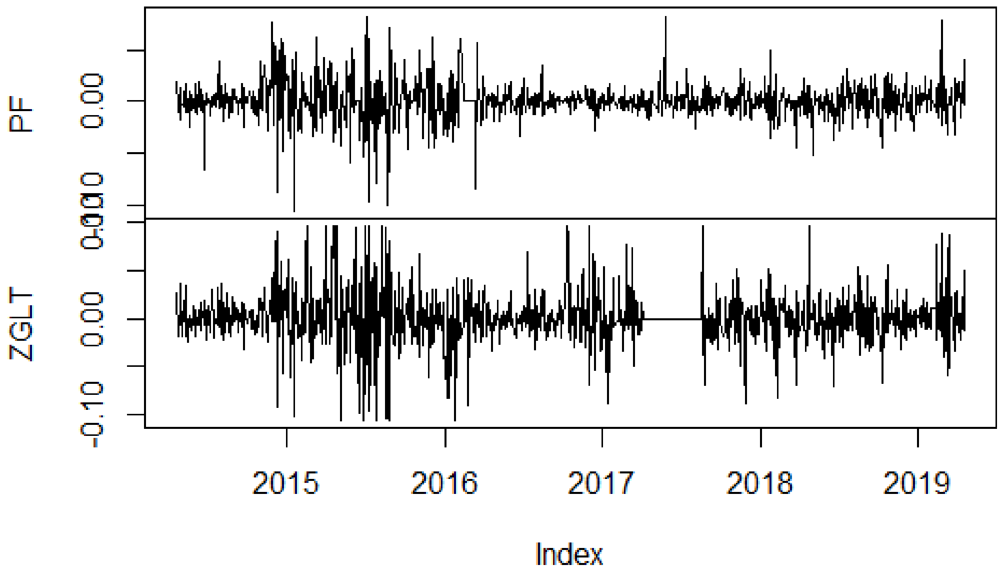

图 1：两只股票日收益率时序图

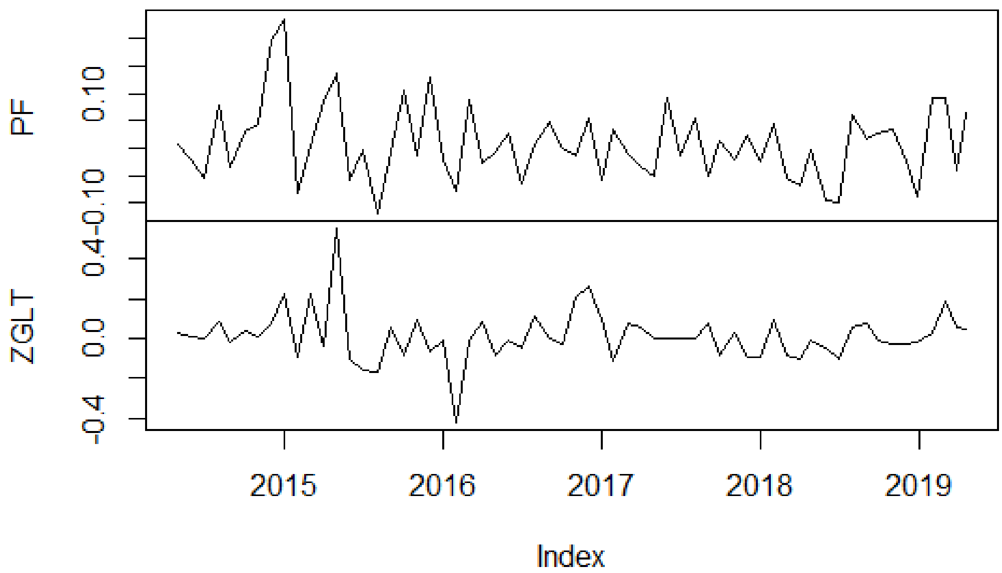

图 2：两只股票月收益率时序图

感谢中国人民大学统计学院杨昊宇同学提供的代码、数据与报告

二、 多元正态性检验

2.1 实现 PPT（p.238）中表格：基于 b_2, k_2 统计量检验两只股票收益率的联合正态性。

检验的原假设是两序列服从二元正态分布，即 $X \sim N_2(\mu, \Sigma)$。由 squared Mahalanobis distances：$D_i = (X_i - \bar{X})' S^{-1} (X_i - \bar{X})$，以及 Mahalanobis angles：$D_{ij} = (X_i - \bar{X})' S^{-1} (X_j - \bar{X})$ 计算 b_2 和 k_2 进行检验：

$$b_2 = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij}^3, \quad k_2 = \frac{1}{n} \sum_{i=1}^{n} D_i^4$$

在原假设成立的条件下当样本量 n 趋向于无穷时，$\frac{n}{6} b_2 \sim \chi_4^2, \frac{k_2 - 8}{\sqrt{64/n}} \sim N(0,1)$。

计算的 b_2, k_2 统计量结果如表 1：

表 1：统计量和相应的 p 值

	Daily	Monthly
n	1220	61
b_2	0.64	1.27
p-value	0.00	0.01
k_2	20.62	12.97
p-value	0.00	0.00

上述结果 p 值均小于 0.05，因此有理由拒绝原假设，即两只股票不服从二元正态分布。

2.2 Q-Q 图：分别绘制两只股票收益率（正态分布）、Mahalanobis distances（卡方分布）的 Q-Q 图，进行说明。

边际正态性检验：

首先绘制两只股票日收益率（正态分布）的 Q-Q 图，图 3：

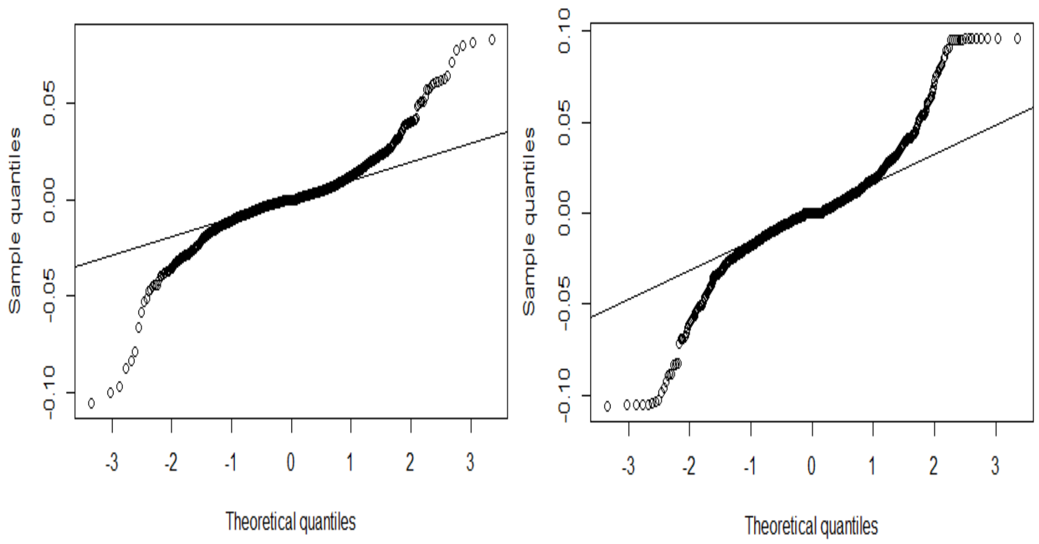

图 3：两只股票日收益率（正态分布）Q-Q 图（PF 左，ZGLT 右）

绘制两只股票月收益率（正态分布）的 Q-Q 图，图 4：

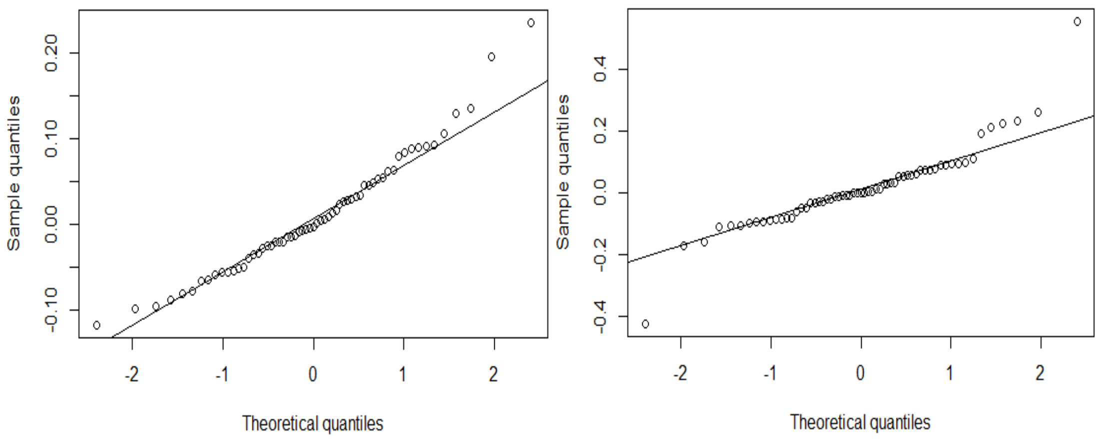

图 4：两只股票月收益率（正态分布）Q-Q 图（PF 左，ZGLT 右）

由图 3 可以明显看到两只股票的日收益率均不服从正态分布，在尾部差异较大；由图 4 可以基本判断浦发银行股票的月收益率基本服从正态分布（尾部差异相对较小），中国联通股票的月收益率数据（尾部差异相对较大）不服从正态分布。

分别对上述两只股票的日收益率、月收益率进行 Shapiro-Wilk 检验，检验结果如表 2：

表 2：Shapiro-Wilk 检验

	PF		ZGLT	
returns	Daily	Monthly	Daily	Monthly
p-value	1.98e-28	5.38e-02	1.76e-26	3.67e-05

由 Shapiro-Wilk 检验，浦发银行股票的月收益率数据不能拒绝原假设，即认

为该序列服从正态分布。

联合正态性检验：

绘制两只股票的日收益率和月收益率（卡方分布）的 Q-Q 图，图 5：

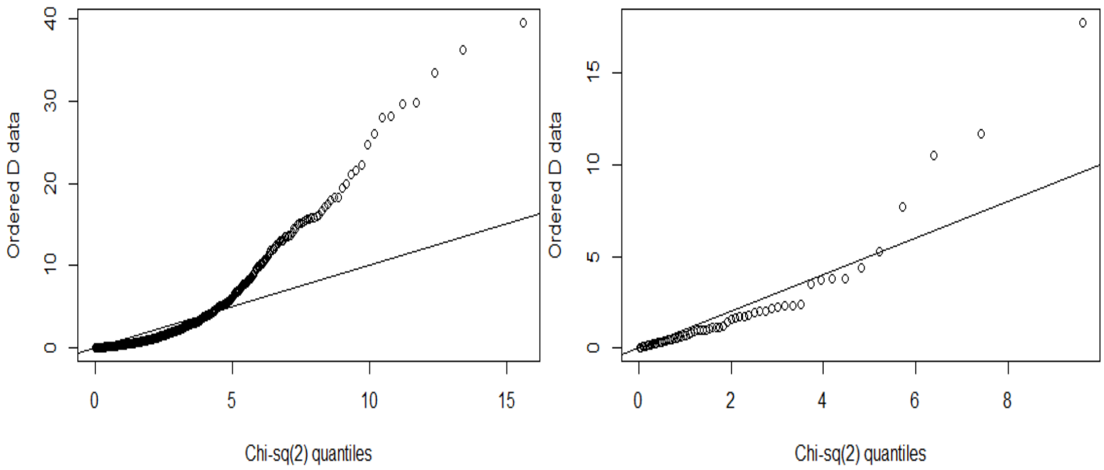

图 5：两只股票的日收益率和月收益率的 Q-Q 图（日左，月右）

由上述 Q-Q 图可以看出两只股票的日收益率明显不服从卡方分布，但两只股票的月收益率基本服从卡方分布，计算 KS 统计量的 p 值如表 3：

表 3：KS 检验

	Daily	Monthly
p-value	0.00	0.11

在上述分析过程中发现月收益率数据的正态性与日收益率数据的正态性有一定的偏差，可能的原因是 5 年的月收益率数据样本量相对较少。因此考虑增加样本量，即将时间长度设置为 2008 年 4 月 19 日至 2019 年 4 月 19 日，对两只股票的月收益率数据分别进行边际正态性检验（Shapiro-Wilk 检验）以及联合正态性检验（KS 检验），检验结果如表 4：

表 4：增加样本量的月收益率序列正态性检验结果

	marginal		Joint
returns	Monthly（PF）	Monthly（ZGLT）	Monthly
p-value	3.04e-05	4.57e-07	1.46e-06

表 4 中的结果表明当样本量增加时，有理由拒绝两只股票的月收益率数据的正态性假设。

感谢中国人民大学统计学院杨昊宇同学提供的代码、数据与报告

三、 二元分布拟合（日收益率）

3.1 分别利用二维正态分布和二维 t 分布进行模型拟合。

首先利用二维正态分布进行拟合，拟合结果为：

均值 $\mu=(4.67e-04,6.90e-04)^T$，协方差矩阵 $\Sigma=\begin{pmatrix}2.98e-04 & 1.98e-04\\ 1.98e-04 & 7.23e-04\end{pmatrix}$，

相关系数 $\rho_{12}=\rho_{21}=0.43$。

利用二维 t 分布进行拟合，拟合结果为：

均值 $\mu=(4.911e-06,-4.89e-04)^T$，协方差矩阵 $\Sigma=\begin{pmatrix}7.87e-05 & 4.51e-05\\ 4.51e-05 & 1.86e-04\end{pmatrix}$，

自由度为 2.05。

3.2 实现 PPT（p.252）中图：模拟生成同样数目的二维正态分布和二维 t 分布样本，绘制散点图，对尾部和相依性进行说明。

生成同样数目的二维正态分布样本绘制散点图如图 6：

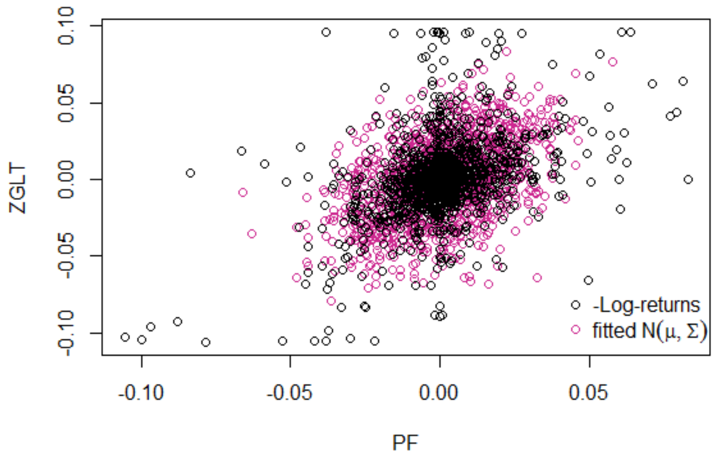

图 6：拟合二维正态分布样本散点图

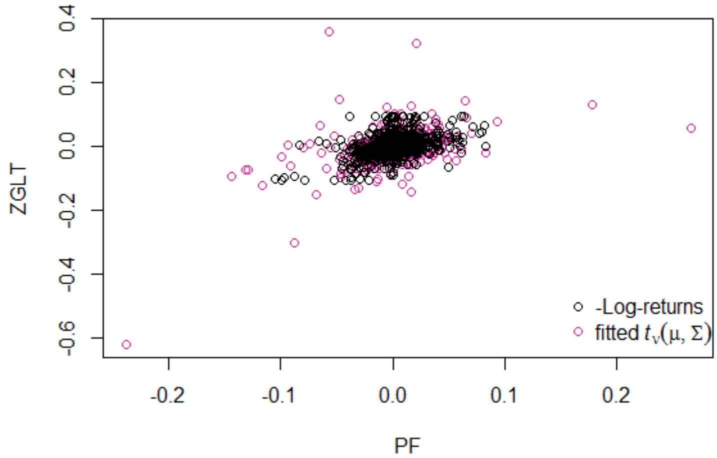

图 7：拟合二维 t 分布样本散点图

感谢中国人民大学统计学院杨昊宇同学提供的代码、数据与报告

由上述散点图可以看出，由于二维正态分布不是厚尾的，因此对尾部数据拟合的不是很好，尾部相依性不高；相反，图 7 中二维 t 分布拟合的更好。

3.3 组合风险指标计算：假设在两只股票的投资比例均为 0.5，基于拟合的二维正态分布模型计算后一天的组合收益率 $0.5X_1(t+1)+0.5X_2(t+1)$的 VaR 和 ES 风险指标（α =0.95）

假设市场上有 n 种不同风险的证券，其收益率ζ_i服从高斯分布，即 $\zeta_i \sim N(\mu_i,\sigma_i^2), i=1,2,\ldots,n$，记投资于证券 i 的比重为$\omega_i (i=1,2,\ldots,n)$，那么投资组合$\zeta=\omega_1\zeta_1+\omega_2\zeta_2+\cdots+\omega_n\zeta_n$，并且

$$\zeta \sim N\left(\sum_{i=1}^{n}\omega_i\mu_i, \sum_{i=1}^{n}\omega_i^2\sigma_i^2+\sum_{i=1}^{n}\sum_{j\neq i, j=1}^{n}\omega_i\omega_j\sigma_{ij}\right)$$

其中σ_{ij}为ζ_i和ζ_j的协方差。

由此可以得到基于高斯分布下投资组合的 VaR 计算公式：

$$VaR_c=-\left(\alpha_c^N\sigma+\mu\right)=-\left(-\alpha_c^N\sqrt{\sum_{i=1}^{n}\omega_i^2\sigma_i^2+\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n}\omega_i\omega_j\sigma_{ij}}+\sum_{i=1}^{n}\omega_i\mu_i\right)$$

ES 计算公式为：

$$ES_c=\frac{\sqrt{\sum_{i=1}^{n}\omega_i^2\sigma_i^2+\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n}\omega_i\omega_j\sigma_{ij}}}{(1-c)\sqrt{2\pi}}\exp\left[-\frac{\left(\alpha_c^N\right)^2}{2}\right]+\sum_{i=1}^{n}\omega_i\mu_i$$

按照上述公式，基于拟合的二维正态分布模型计算后一天的组合收益率的 VaR 和 ES 风险指标如表 5：

表 5：VaR 和 ES 风险指标

	VaR	ES
数值	0.0307	0.0391

四、 二元 Garch 模型（日收益率）

4.1 实现 PPT(p.677)中互相关图，观察是否有 lead lag relation。

实现 PPT(p.677)中互相关图，结果如图 8：

感谢中国人民大学统计学院杨昊宇同学提供的代码、数据与报告

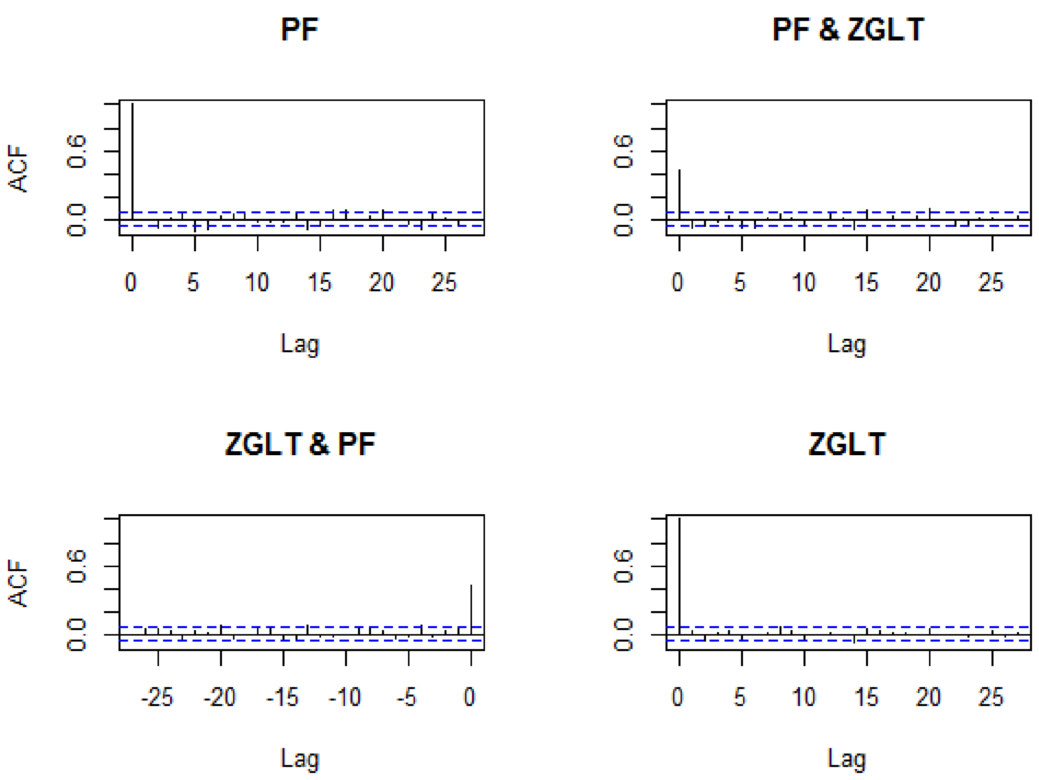

图 8：序列互相关图

由上图可知，ZGLT 和 PF 有延迟相关性，存在滞后关系，PF 领先于 ZGLT 市场。

4.2 拟合正态分布新息的 CCC-GARCH(1,1)和 DCC-GARCH(1,1)过程。

首先拟合正态分布新息的 CCC-GARCH(1,1)过程，拟合结果如下：

Method: Multivariate ML
Message (nlminb): relative convergence (4)
No. of observations: 1220
Sample: 2014-04-22 to 2019-04-19
Coefficients:

	intercept1	intercept2	arch11.1	arch21.1	arch12.1
Estimate:	0.09989441	-0.01422134	0.0988543	0.007298862	-0.02236995
Std. Error:	NA	NA	0.0218479	0.015341808	0.02303538

	arch22.1	garch11.1	garch21.1	garch12.1	garch22.1
Estimate:	0.06195347	0.85463313	0.00854089	0.07192875	0.90354409
Std. Error:	0.01418059	0.04158917	0.02465547	0.05127636	0.02591091

	Elnz2no1	Elnz2no2
Estimate:	-1.57807180	-1.51146000
Std. Error:	0.07474883	0.05967871

Log-likelihood (log-mgarch): 5026.107 Log-likelihood (varma): -4150.461
No. of obs. without zeros: 991 No. of obs. with zeros: 229

感谢中国人民大学统计学院杨昊宇同学提供的代码、数据与报告

拟合正态分布新息的 DCC-GARCH(1,1)过程，拟合结果如下：

----
*   DCC GARCH Fit   *

Distribution : mvt
Model : DCC(1,1)
No. Parameters : 12
[VAR GARCH DCC UncQ] : [0+8+3+1]
No. Series : 2
No. Obs. : 1220
Log-Likelihood : 6655.07
Av.Log-Likelihood : 5.45

Optimal Parameters

	Estimate	Std. Error	t value	Pr(>|t|)
[PF].mu	-0.000048	0.000370	-0.13053	0.896149
[PF].omega	0.000002	0.000006	0.33771	0.735580
[PF].alpha1	0.060069	0.054138	1.10955	0.267193
[PF].beta1	0.936484	0.052360	17.88545	0.000000
[ZGLT].mu	0.000635	0.000495	1.28287	0.199538
[ZGLT].omega	0.000021	0.000018	1.14053	0.254068
[ZGLT].alpha1	0.156818	0.055895	2.80560	0.005022
[ZGLT].beta1	0.827951	0.060331	13.72354	0.000000
[Joint]dcca1	0.059592	0.022758	2.61855	0.008831
[Joint]dccb1	0.897257	0.044058	20.36533	0.000000
[Joint]mshape	4.000000	0.382918	10.44611	0.000000

Information Criteria

Akaike	-10.890
Bayes	-10.840
Shibata	-10.890
Hannan-Quinn	-10.871

感谢中国人民大学统计学院杨昊宇同学提供的代码、数据与报告

基于 DCC-GARCH 模型绘制条件协方差、条件相关系数如图 9：

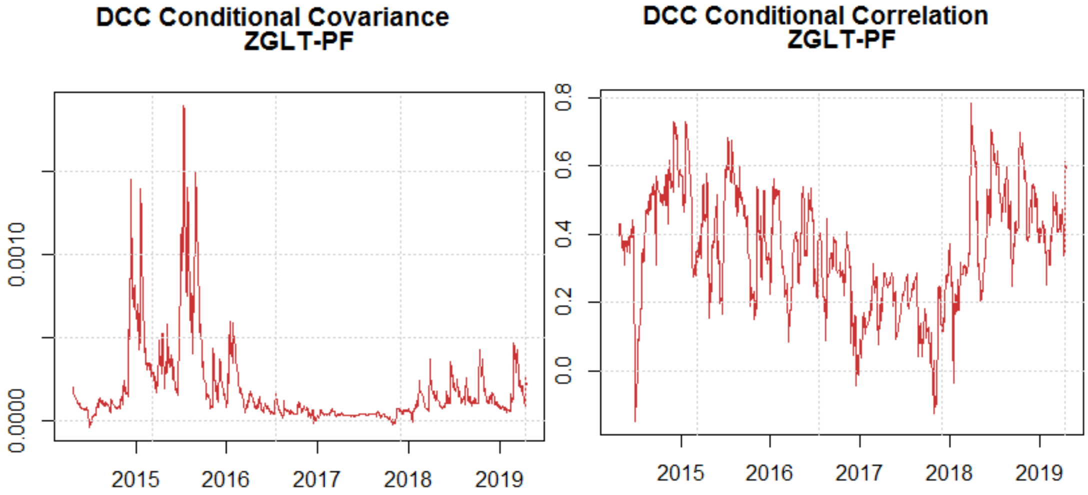

图 9：序列条件协方差、条件相关系数图

4.3 组合风险指标计算：假定在两只股票的投资比例均为 0.5，计算后一天的组合收益率 $0.5X_1(t+1)+0.5X_2(t+1)$的 VaR 和 ES 风险指标（α =0.95）。

计算步骤：

1. 根据拟合的边际 GARCH 模型进行预测，得到 $\mu_{1,t+1}, \sigma_{1,t+1}$ 和 $\mu_{2,t+1}, \sigma_{2,t+1}$；

2. 根据拟合的 DCC-GARCH(1,1)模型，预测相关系数矩阵，得到 ρ_{12}；

3. 计算两只股票的协方差：$\sigma_{12} = \rho_{12}\sigma_1\sigma_2$；

4. 带入公式进行计算。

$$VaR_c = -(\alpha_c^N \sigma + \mu) = -\left(-\alpha_c^N \sqrt{\sum_{i=1}^n \omega_i^2 \sigma_i^2 + \sum_{i=1}^n \sum_{j=1, j\neq i}^n \omega_i \omega_j \sigma_{ij}} + \sum_{i=1}^n \omega_i \mu_i\right)$$

$$ES_c = \frac{\sqrt{\sum_{i=1}^n \omega_i^2 \sigma_i^2 + \sum_{i=1}^n \sum_{j=1, j\neq i}^n \omega_i \omega_j \sigma_{ij}}}{(1-c)\sqrt{2\pi}} \exp\left[-\frac{(\alpha_c^N)^2}{2}\right] + \sum_{i=1}^n \omega_i \mu_i$$

计算结果如表 6：

表 6：VaR 和 ES 风险指标

	VaR	ES
数值	0.0308	0.0390

上述结果与 3.3 中基于拟合的二维正态分布模型计算的 VaR 和 ES 值十分相近，说明两种方法均存在合理性。

感谢中国人民大学统计学院杨昊宇同学提供的代码、数据与报告