

细说回测：模拟路径拓宽回测样本

——机器学习系列报告之四

金融工程深度

该篇报告中，我们引入一个新的能够对收益率进行刻画的 SGED 模型，在此基础上衍生出了 SGED-ARMA 模型和 SGED-GARCH 模型，采用滚动拟合的方式模拟价格序列，并以回测 RSRS 择时模型为例，展示模拟路径通过增大数据样本，降低路径依赖，减少参数过拟程度等方式来提升回测有效性的作用。

- ◆ **广义误差模型的偏度拓展(SGED)**。我们首先从价格收益率存在的尖峰厚尾以及偏度的角度出发，引入了一个新的能够对收益率进行刻画的 SGED 分布，相比于普通的正态分布，SGED 分布能够更好地考虑到价格收益率存在的一些非正态的统计特征。在模拟收益率数据时，本篇报告采用了滚动拟合的方式，并且提取每天的开盘价格进行了额外的拟合工作，以确保价格数据构造的相似性。
- ◆ **收益率自相关性与波动聚集效应**。出于对收益率自相关性和波动聚集效应的考虑，本篇报告分别构建了 SGED-ARMA 模型和 SGED-GARCH 模型，这两个模型能够分别在上述两个方面对 SGED 模型的效果进行提升。由于我们采用了滚动拟合的模拟方式，模拟价格序列的波动聚集效应在各个模型上都有所保证，所以最后选取 SGED-ARMA 模型来模拟价格序列，进行后续的策略有效性研究。
- ◆ **阻力支撑相对强度 (RSRS) 择时指标的有效性及其过拟程度**。将 RSRS 择时策略应用于 SGED-ARMA 模型生成的 300 条模拟价格序列，以及沪深 300 指数价格序列。各个指标的统计结果表明，RSRS 择时模型整体而言确实是较为有效的择时模型，但其参数选择有一定程度的过拟合，以沪深 300 指数历史走势测试的参数构建的策略，其回测结果一定程度上夸大了指标的实际择时能力。
- ◆ **风险提示**：测试结果均基于模型和历史数据，模型存在失效的风险。

分析师

胡骥聪 (执业证书编号：S0930519060002)
021-52523683
hujicong@ebscn.com

刘均伟 (执业证书编号：S0930517040001)
021-52523679
liujunwei@ebscn.com

相关研报

《机器学习：开拓金融量化新前沿——机器学习系列报告之一》2019.01
《数据纵横：探秘 K 线结构新维度——机器学习系列报告之二》2019.02
《数据纵横：拓展差分的操作域——机器学习系列报告之三》2019.04
《基于阻力支撑相对强度 (RSRS) 的市场择时——技术择时系列报告之一》2017.05

目 录

1、 模拟路径对抗路径依赖.....	5
2、 刻画实际收益率特征.....	5
2.1、 描述收益率尖峰厚尾有偏的特征	6
2.2、 处理收益率自相关性与波动聚集效应.....	10
3、 提升回测参考价值：以 RSRS 为例.....	16
3.1、 阻力支撑相对强度原理及 RSRS 指标构建.....	17
3.2、 RSRS 择时模型有效性与参数过拟程度检验	18
3.3、 测试更合适的参数对	19
4、 总结.....	21
5、 风险提示.....	21

图目录

图 1：沪深 300 指数价格走势	7
图 2：每日开盘价收益率的参数估计结果	8
图 3：每日最后一分钟收益率的参数估计结果	9
图 4：SGED 模拟价格路径，正态分布模拟价格路径与市场基准对比图	9
图 5：真实样本收益率的自相关系数和偏自相关系数图	10
图 6：SGED 模拟路径自相关系数和偏自相关系数图	11
图 7：SGED-ARMA 模拟价格路径，SGED 模拟价格路径与市场基准对比图	12
图 8：SGED-ARMA 每日开盘价收益率的参数估计结果	12
图 9：SGED-ARMA 每日最后一分钟收益率的参数估计结果	13
图 10：SGED-ARMA 模拟路径自相关系数和偏自相关系数图	13
图 11：SGED-GARCH 模拟价格路径，SGED 模拟价格路径与市场基准对比图	14
图 12：SGED-GARCH 每日开盘价收益率的参数估计结果	15
图 13：SGED-GARCH 每日最后一分钟收益率的参数估计结果	15
图 14：SGED-GARCH 模拟路径自相关系数和偏自相关系数图	16
图 15：RSRS 择时策略在沪深 300 指数上的净值表现	18
图 16：RSRS 择时策略在 SGED-ARMA 模拟价格序列上的净值表现	19
图 17：改变参数后 RSRS 择时策略在 SGED-ARMA 模拟价格序列上的净值表现	20

表目录

表 1：不同市场状态下阻力支撑相对强弱的含义	17
表 2：RSRS 策略在沪深 300 指数和 300 条模拟价格序列上的统计表现	18
表 3：改变参数后 RSRS 择时策略在沪深 300 指数和 300 条模拟价格序列上的统计表现	20

1、模拟路径对抗路径依赖

回测是量化研究者最频繁接触的事项之一。无论是什么样的投资思路、什么类型的策略，都会通过回测来验证其有效性：即假定在过去的市场中按策略信号交易，会获得怎样的收益表现。然而不像自然科学可以控制变量重复实验，金融市场的样本序列仅此一个，利用历史数据回测不可避免的会有一定路径依赖的成分。如果不能在一定程度上了解策略路径依赖的程度，那即使策略回测有不错的收益表现，在运用策略的时候依然会怀疑这是否只是一个运气产物？策略的有效性与外推性便缺乏一定的说服力。

为了克服上述问题、提升回测的有效性，研究者们探索了不少方法，其中一种做法是基于股票或指数的历史收益率数据，拟合收益率序列的分布，并通过 Monte Carlo 方式生成各种新的价格路径，最后将策略应用于这些路径以说明策略收益的稳定性。在这个过程中，如何找到合适的收益率序列的分布便成为关键，如果拟合的分布和价格的实际分布有着很大的差异，这一模拟过程也就失去了实用意义。

绝大部分的金融标的收益率序列都存在明显的尖峰厚尾现象，有些收益率序列还存在明显的偏度，如果只是使用简单的正态分布去拟合收益率序列，便无法对其进行考量，拟合出来的分布也就失去了这些样本特征。再考虑到样本存在的波动聚集效应和收益率自相关性这些明显的特征，我们同样需要寻找合适的方式对其进行刻画。

本篇报告将就以上几点问题进行探索，希望能够找到合适的分布以及拟合方式，模拟出和特定股票或指数历史收益率数据具有相同统计特征、波动聚集效应、收益率自相关性的样本路径。并在最后利用 RSRS 择时模型，展示模拟路径通过增大数据样本，降低路径依赖，减少参数过拟程度等方式来提升回测有效性的作用。

2、刻画实际收益率特征

使用几何布朗运动刻画对数收益率是一个经典方法，但没有表现出对数收益率分布的偏度和峰度。Panayiotis Theodossiou¹在 2015 年的文章中讨论过，收益率之所以存在偏度和峰度，是由于变量的高阶矩存在相关性，例如条件异方差，非对称波动性，和一些其他的非线性相关性，只有将这些相关性考虑进模型才能完整得对分布进行刻画。之后，他提出了一个广义误差模型的偏度拓展(SGED, Skewed Generalized Error Distribution)，并通过理论和实证说明了这一模型应用于收益率分布的有效性。

本篇报告将采用 SGED，并且通过一定的方式对股票和指数的样本路径进行拟合，希望能够找到更加贴近于真实样本路径的模拟样本。

¹ Panayiotis Theodossiou, 2015, "Skewed Generalized Error Distribution of Financial Assets and Option Pricing", Multinational Finance Journal: vol. 19, no. 4, pp. 223-266

2.1、描述收益率尖峰厚尾有偏的特征

2.1.1、SGED 的分布形态

SGED 有着如下的密度函数形式：

$$f_u = \frac{k^{1-\frac{1}{k}}}{2\varphi} \Gamma\left(\frac{1}{k}\right)^{-1} \exp\left(-\frac{1}{k} \frac{|u|^k}{(1 + \operatorname{sgn}(u)\lambda)^k \varphi^k}\right) \quad \#(1)$$

其中 $u = y - m$, m 是随机变量 y 的众数, φ 是一个与 y 的标准差相关的比例常数, λ 是一个偏度系数, k 是一个峰度系数, $\operatorname{sgn}(\cdot)$ 表示符号函数：

$$\operatorname{sgn}(u) = \begin{cases} 1, & u \geq 0 \\ -1, & u < 0 \end{cases} \quad \#(2)$$

$\Gamma(\cdot)$ 表示 gamma 函数：

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx \quad \#(3)$$

峰度系数 k 控制了分布的尾部和峰值, 偏度系数 λ ($-1 \leq \lambda \leq 1$) 控制了密度函数在众数 m 附近的下降速率, $\lambda > 0$ 表示分布右偏, $\lambda < 0$ 表示分布左偏。

以下将介绍如何用变量 u 的高阶矩来表示密度函数中的一些系数, 使得分布可拟合。我们可以得到变量 u 的 r 阶矩具有如下的形式：

$$M_{u,r} = Eu^r = A_r \varphi^r \quad \#(4)$$

其中：

$$A_r = \frac{1}{2} [(-1)^r (1 - \lambda)^{r+1} + (1 + \lambda)^{r+1}] G_r, \quad \#(5)$$

$$G_r = k^{\frac{r}{k}} \Gamma\left(\frac{r+1}{k}\right) \Gamma\left(\frac{1}{k}\right)^{-1} \quad \#(6)$$

变量 y 的期望和方差也可以用众数 m 和 u 的高阶矩的形式来表示：

$$\mu = E(y) = m + E(u) = m + A_1 \varphi \quad \#(7)$$

$$\sigma^2 = Eu^2 - (Eu)^2 = (A_2 - A_1^2) \varphi^2 \quad \#(8)$$

其中：

$$A_1 = 2\lambda k^{\frac{1}{k}} \Gamma\left(\frac{2}{k}\right) \Gamma\left(\frac{1}{k}\right)^{-1} \quad \#(9)$$

$$A_2 = (1 + 3\lambda^2) k^{\frac{2}{k}} \Gamma\left(\frac{3}{k}\right) \Gamma\left(\frac{1}{k}\right)^{-1} \quad \#(10)$$

根据 u 的高阶矩, Pearson 偏度系数有如下的表达式：

$$\delta = \frac{(\mu - m)}{\sigma} = A_1 / \sqrt{A_2 - A_1^2} \quad \#(11)$$

变量 y 的标准化偏度 SK 和标准化峰度 KU 可以表示成：

$$SK = \frac{A_3 - 3A_2 A_1 + 2A_1^3}{(A_2 - A_1^2)^{\frac{3}{2}}} \quad \#(12)$$

$$KU = \frac{A_4 - 4A_3 A_1 + 6A_2 A_1^2 + 3A_1^4}{(A_2 - A_1^2)^2} \quad \#(13)$$

其中：

$$A_3 = 2\lambda(1 + \lambda^2)k^{\frac{3}{k}}\Gamma\left(\frac{4}{k}\right)\Gamma\left(\frac{1}{k}\right)^{-1} \quad \#(14)$$

$$A_4 = (1 + 10\lambda^2 + 5\lambda^4)k^{\frac{4}{k}}\Gamma\left(\frac{5}{k}\right)\Gamma\left(\frac{1}{k}\right)^{-1} \quad \#(15)$$

有了上面的定义之后，我们可以把常数 φ 和 m 表示成 σ 和 μ 的表达式：

$$\varphi = \gamma\sigma, \quad m = \mu - \delta\sigma, \quad \gamma = 1/\sqrt{A_2 - A_1^2} \quad \#(16)$$

从而之前 SGED 密度函数的表达式就可以完全写成 σ ， μ ， k 和 λ 的表达式：

$$f_y(\sigma, \mu, k, \lambda) = \frac{k^{1-\frac{1}{k}}}{2\gamma\sigma} \Gamma\left(\frac{1}{k}\right)^{-1} \exp\left(-\frac{1}{k} \frac{|y - \mu + \delta\sigma|^k}{(1 + \operatorname{sgn}(y - \mu + \delta\sigma)\lambda)^k \gamma^k \sigma^k}\right) \quad \#(17)$$

这样我们就能用极大似然估计(MLE)来对参数 k 和 λ 进行估计，而 σ ， μ 可由样本的标准差和均值直接代替。

2.1.2、用 SGED 对真实价格路径进行模拟

在这篇报告中我们进行模拟的对象是沪深 300 指数 2009 年 1 月 5 日到 2017 年 12 月 29 日共 9 年时间的分钟价格。

图 1：沪深 300 指数价格走势图



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

在进行模拟的时候，为了使模拟的数据更加贴近真实数据的走势，我们做了如下处理：

- 对模型参数进行滚动拟合：由于同一条价格曲线在不同的时间段所表现出的样本特征并不统一，如均值、波动率等，我们在每一个时间点都对 SGED 模型的参数通过滚动的方式进行了拟合，滚动周期为 240，即用接近过去一个交易日的的数据对当前分钟的样本参数进行估计。
- 提取每天的开盘价格进行拟合：由于每天的第一分钟的数据和前一天收盘价相比，可能会因为集合竞价的原因产生高开或跳空的情况，这使得

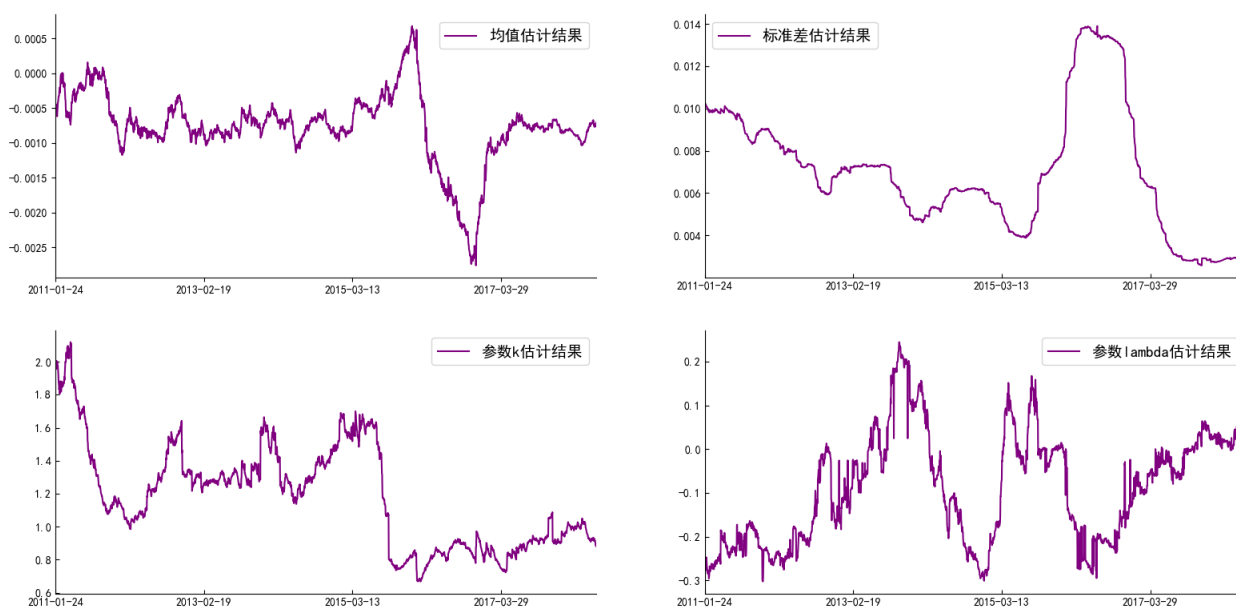
每天第一个分钟的收益率与连续竞价时的分钟收益率有巨大差异。我们在进行参数拟合时将每天的开盘价做了额外的处理，用过去 240 天的跳开收益率对当前开盘第一个分钟的收益率的统计参数进行拟合，最后再重新对每分钟的数据进行时间点上的拼接，得到最终完整的分钟价格收益率。

- 使用拒绝法来产生模拟的价格收益率：由于 SGED 的表达式十分复杂，无法写出其逆函数的显式表达式，这篇报告采用的是拒绝法来产生最后的模拟价格收益率。拒绝法有如下的操作步骤：

1. 确定 SGED 的密度函数 f 的最大值 p ，选定一个常数 c ；
2. 根据价格收益率的取值范围产生一个服从均匀分布的随机数 x ；
3. 产生一个服从取值范围为 $(0, cp)$ 的均匀分布的随机数 y ；
4. 如果 $f(x) > y$ ，则返回 x 作为生成的随机数，否则返回步骤 2。

用上述方法对每一分钟的参数进行 MLE 拟合，每日开盘价收益率的参数估计结果，有如下的形式：

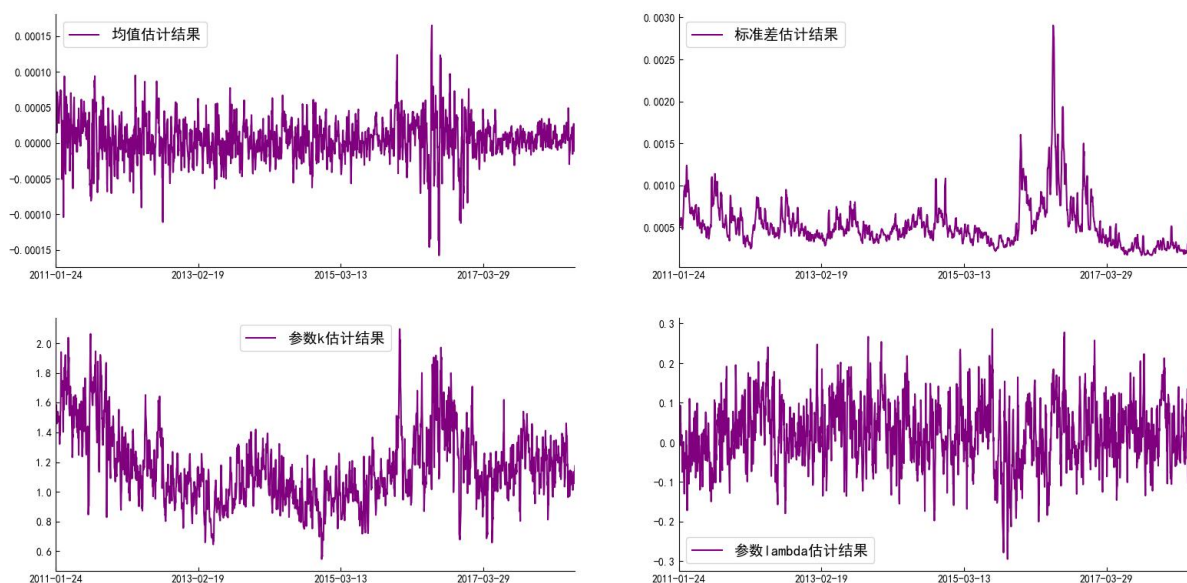
图 2：每日开盘价收益率的参数估计结果



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

以每天最后一分钟收益率参数估计结果作为日内每分钟收益率参数估计结果的代表，有如下的形式：

图 3：每日最后一分钟收益率的参数估计结果

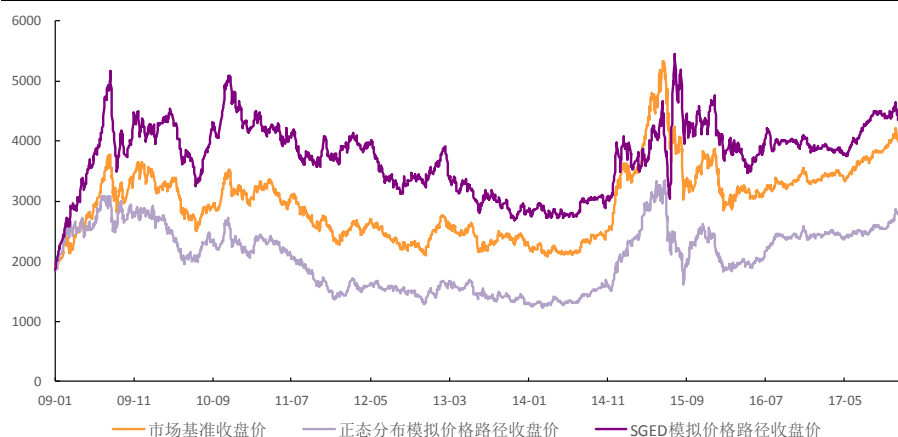


资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

从上面两张图可以看出，每日第一分钟与前一日最后时刻价格的收益率和每日最后一分钟收益率之间存在很大的差异，在四个参数估计上后者都更加平稳，而前者的趋势性更加明显，更能够体现当前市场价格收益率服从的模型是不同的。这也说明了我们在模拟的时候将两者分开估计的意义所在。

利用上面得到的参数，通过 Monte Carlo 方式生成一条新的价格路径，以沪深 300 日收盘价作为市场基准，取以 SGED 为收益率分布函数模拟收盘价结果与取以正态分布为收益率分布函数模拟收盘价结果的对比图如下：

图 4：SGED 模拟价格路径，正态分布模拟价格路径与市场基准对比图



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

利用 SGED 分布刻画单个收益率一定程度上解决了正态分布无法刻画出的尖峰厚尾有偏的性质，然而在模拟数据时还有另一个常见的问题，那就是收益率序列并非独立的，它呈现明显的自相关性与波动聚集效应，在下一章节我们将尝试对该序列性质进行刻画。

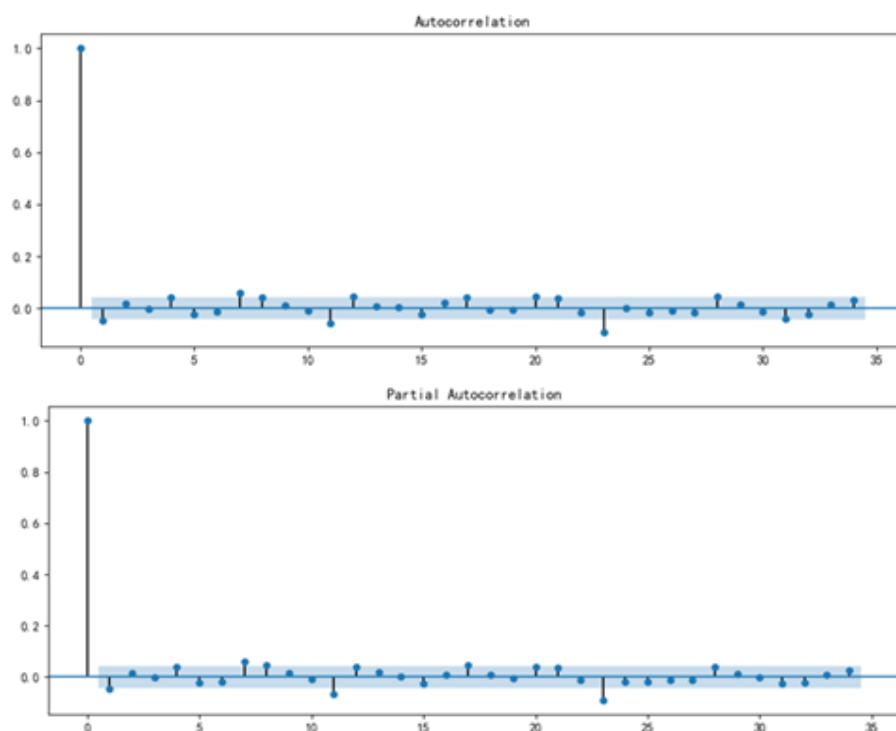
2.2、处理收益率自相关性与波动聚集效应

在之前的讨论中，我们用 SGED 模型拟合价格收益率，从模拟价格路径来看似乎取得了不错的结果。在这一部分的讨论中，我们将加入对于价格收益率的自相关性与波动聚集效应的讨论，希望能够得到更加贴近真实样本路径的模拟价格路径。

2.2.1、SGED-ARMA 模型的构建

我们先画出真实样本收益率的自相关系数和偏自相关系数图：

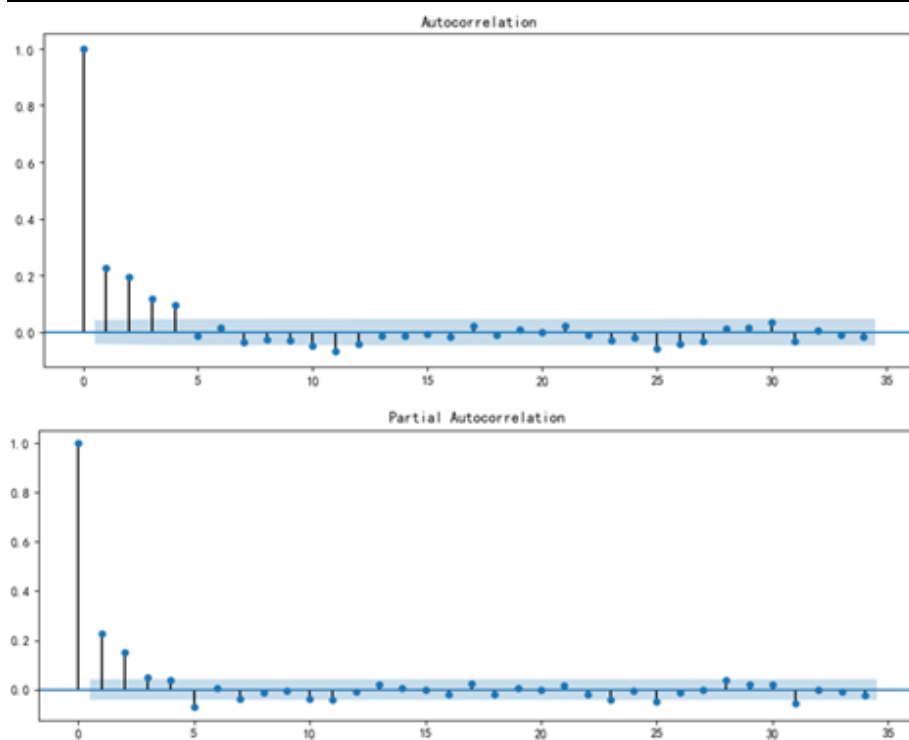
图 5：真实样本收益率的自相关系数和偏自相关系数图



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

从图中我们可以看出，真实样本收益率存在明显的一阶自相关和偏自相关性，下面我们再对 SGED 的模拟结果做出自相关系数和偏自相关系数图：

图 6: SGED 模拟路径自相关系数和偏自相关系数图



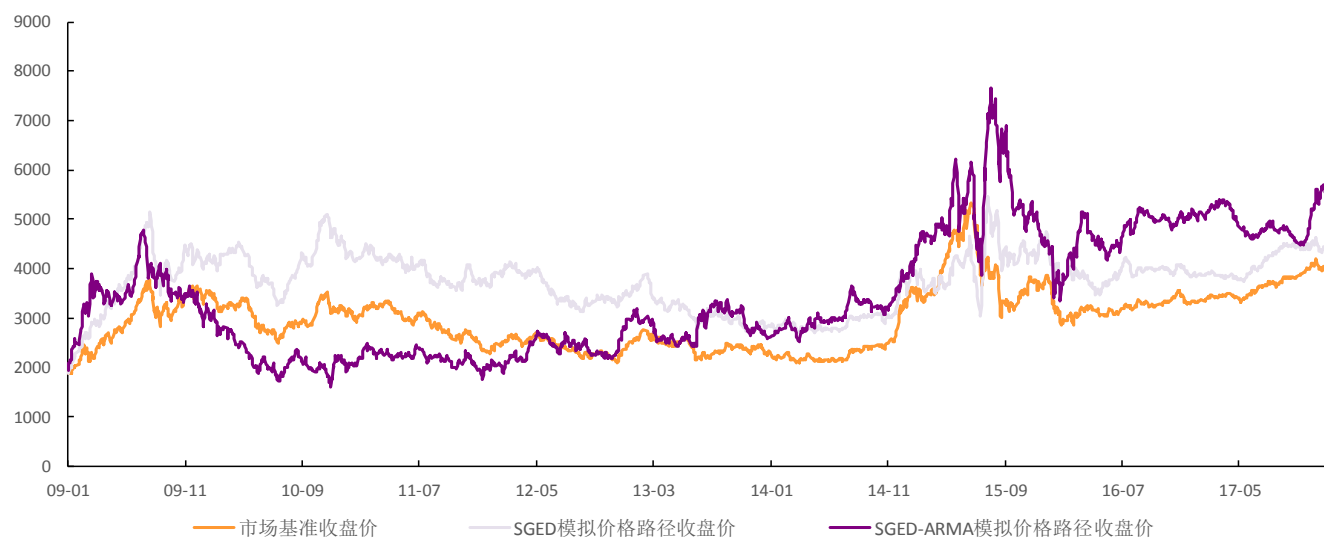
资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

从上图中我们可以看出模拟出来的序列整体上的自相关性存在 4 阶拖尾，偏自相关性存在 3 阶拖尾，这与真实的样本路径之间存在很大的差异，简单用正态分布去模拟路径也有类似的现象。为了能够让我们模拟的价格路径与真实样本路径具有接近的自相关和偏自相关性质，我们通过在拟合 SGED 参数时引入 ARMA 模型的方式来改进 SGED 模型。

在 SGED 模型中我们用样本的均值作为总体期望的估计，但从上图中我们可以看到，样本整体收益率之间其实存在着一阶自相关和偏自相关性，如果只是简单的用均值作为总体期望的估计就会导致最后的结果与真实路径在自相关性上存在明显的差异。

为了避免这一现象的发生，我们用 ARMA 模型对原模型进行改进。在模拟参数时，先对价格收益率建立 ARMA(1,1)模型，用得到的模型来产生新的模拟收益率数据，用模拟收益率数据求得的均值作为总体期望的估计带入 SGED 模型中来估计其他参数。这样做，是希望能够使最后产生的模拟收益率，与样本收益率之间有相同的自相关性和偏自相关性。

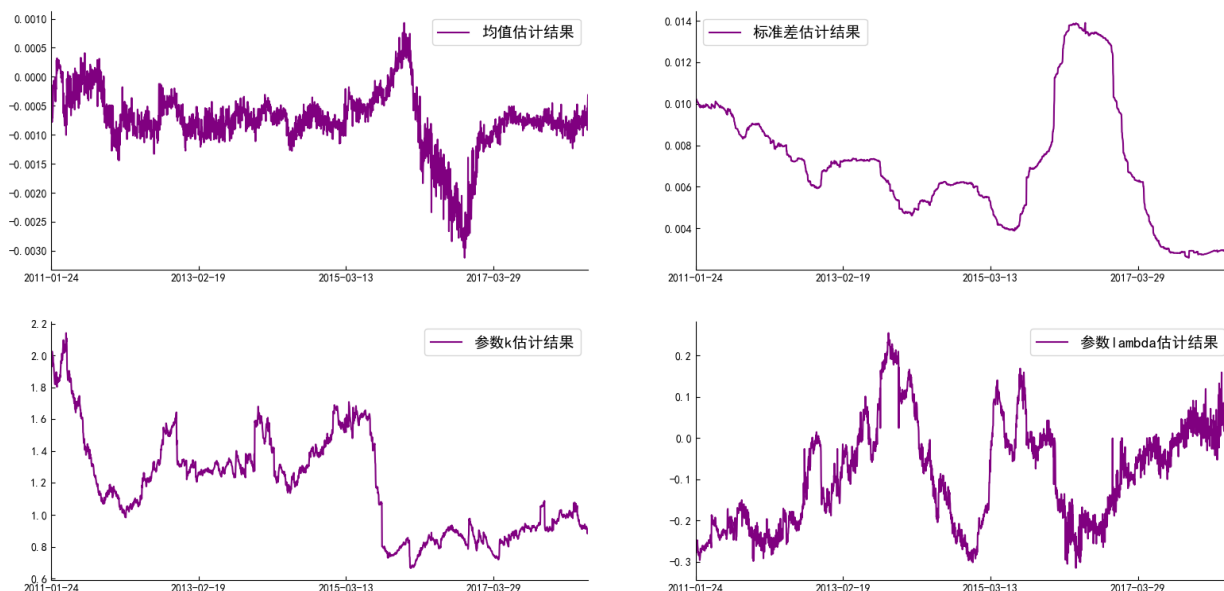
图 7: SGED-ARMA 模拟价格路径, SGED 模拟价格路径与市场基准对比图



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

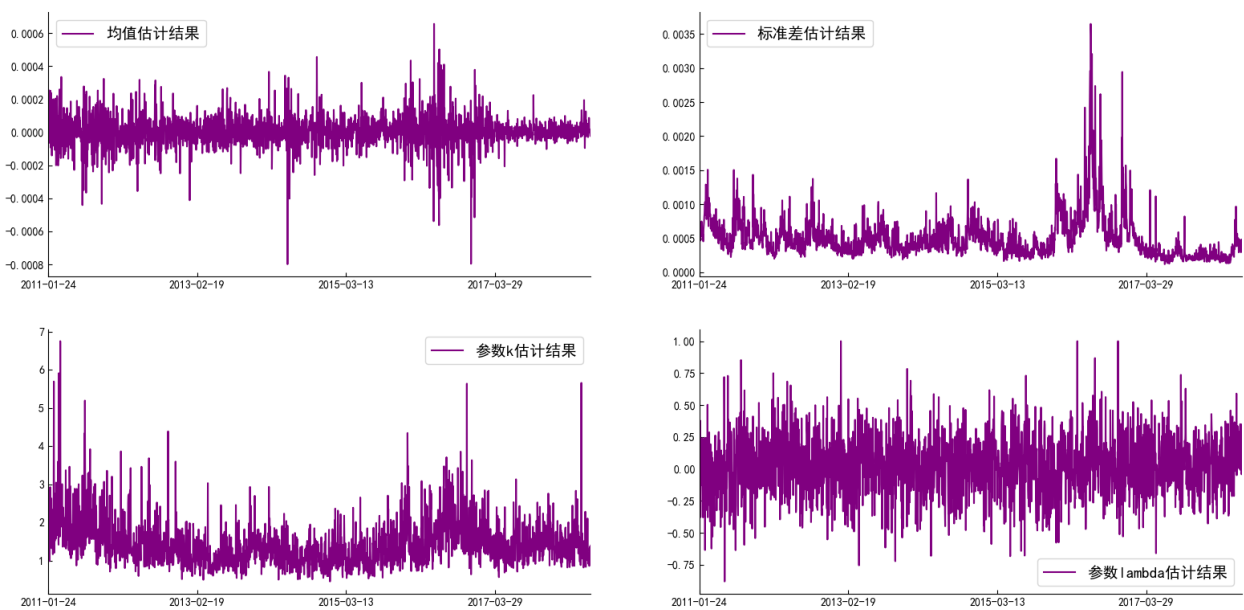
我们同样做出 SGED-ARMA 模拟的每日开盘价收益率的参数估计结果，以及每日最后一分钟收益率的参数估计结果。与 SGED 模型相比，SGED-ARMA 的参数估计结果更加具有波动性，各参数的估计结果取值跨度也更大。

图 8: SGED-ARMA 每日开盘价收益率的参数估计结果



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

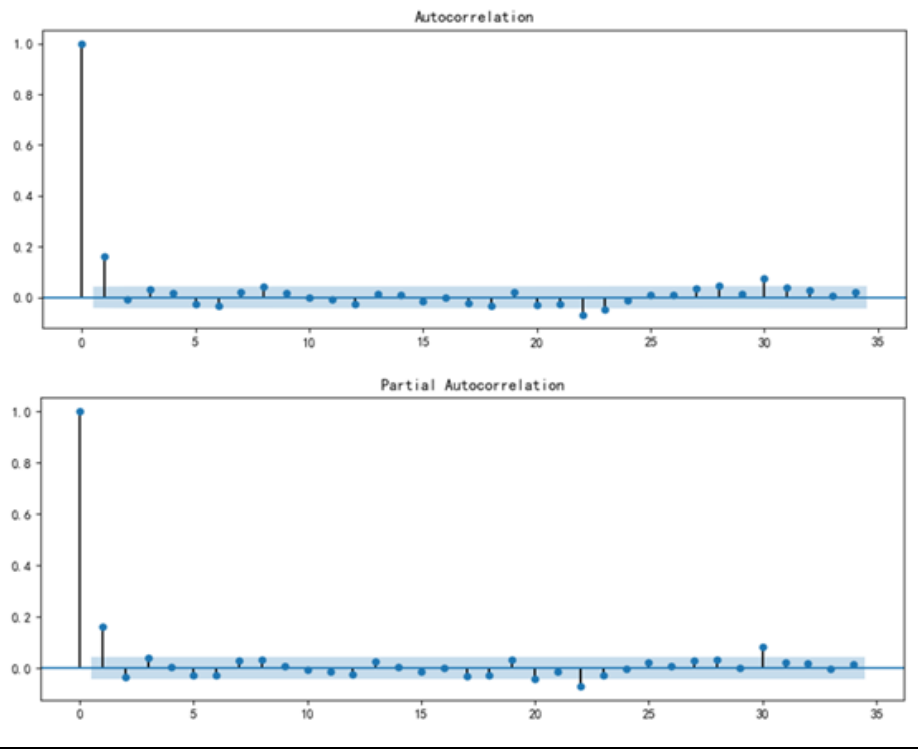
图 9：SGED-ARMA 每日最后一分钟收益率的参数估计结果



资料来源：光大证券研究所， 注：样本区间为 2010-01-01 至 2017-12-31

观察 SGED-ARMA 模拟收益率的自相关系数和偏自相关系数图，可以看出模拟出来的序列整体上已经不再具有高阶的自相关性和偏自相关性，而只存在一阶自相关性和偏自相关性，这与真实样本收益率的情况更为接近。

图 10：SGED-ARMA 模拟路径自相关系数和偏自相关系数图



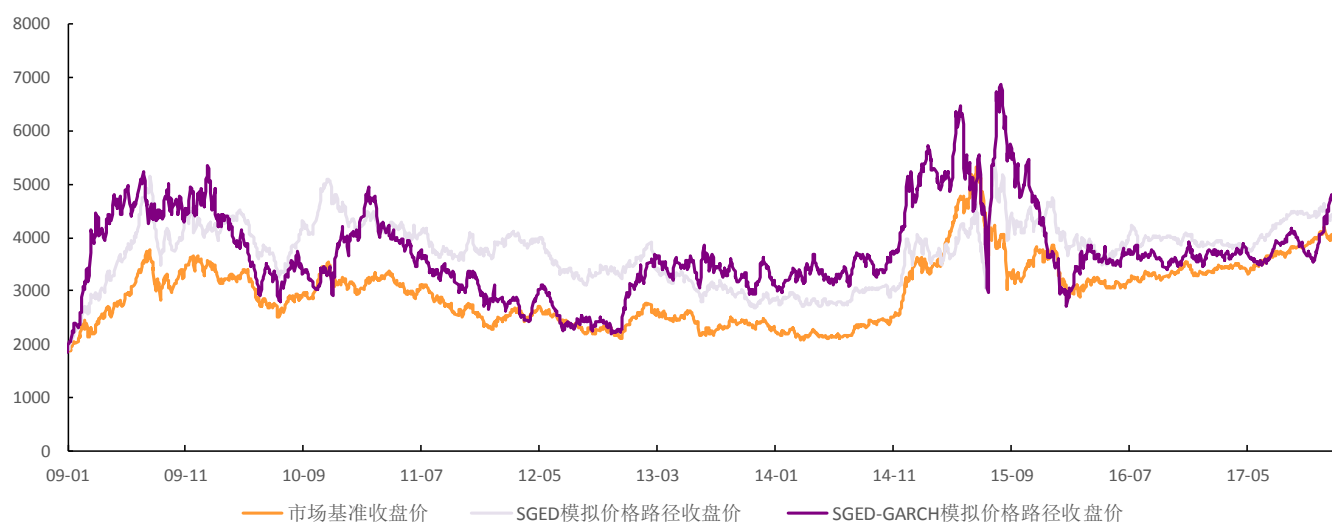
资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

2.2.2、SGED-GARCH 模型的构建

与 SGED-ARMA 模型一样，出于对波动率聚集效应的考虑，我们希望能够构建一个模型，使得模拟的收益率在波动率聚集效应上与真实样本波动率的波动率更为贴近。

在这里我们考虑使用的是 GARCH 模型。在模拟参数时，先对价格收益率建立 GARCH(1,1)模型，用得到的模型来产生新的模拟收益率数据，用模拟收益率数据求得的标准差作为总体标准差的估计带入 SGED 模型中来估计其他参数。这样做，是希望能够使最后产生的模拟收益率，无论是在日内还是整体上都和真实样本路径具有更加贴近的波动率聚集效应。

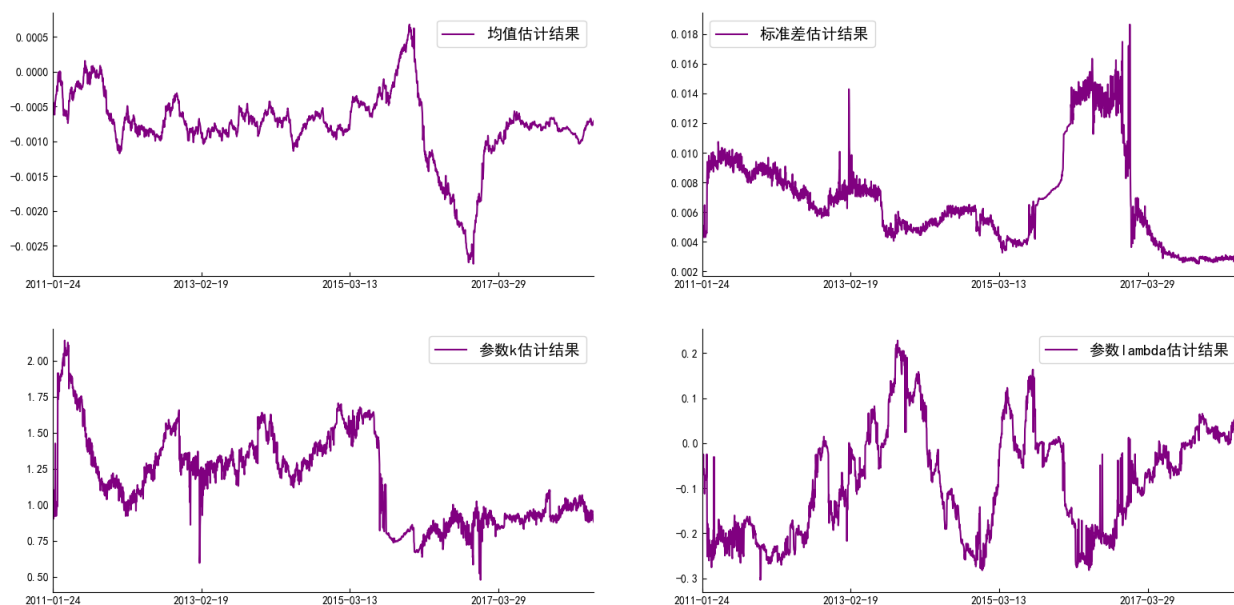
图 11：SGED-GARCH 模拟价格路径，SGED 模拟价格路径与市场基准对比图



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

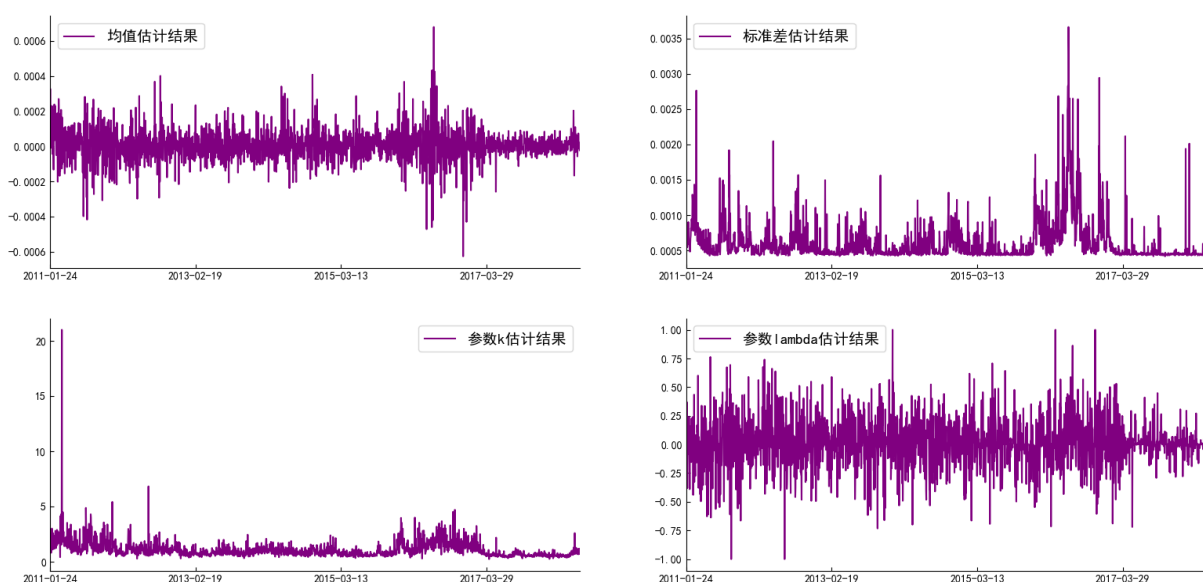
类似于 SGED-ARMA 模型的研究，我们同样做出 SGED-GARCH 模拟的每日开盘价收益率的参数估计，以及每日最后一分钟收益率的参数估计结果。其各参数的估计范围和波动与 SGED 模型的结果差异，远不及 SGED-ARMA 模型与 SGED 模型的结果差异那么大，但可能存在部分分钟

图 12: SGED-GARCH 每日开盘价收益率的参数估计结果



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

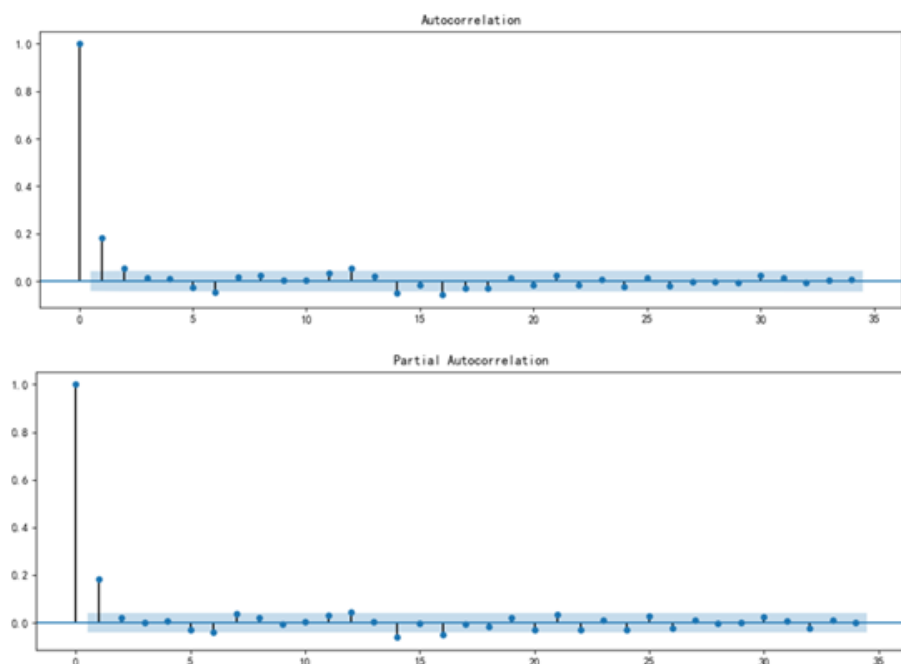
图 13: SGED-GARCH 每日最后一分钟收益率的参数估计结果



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

观察 SGED-GARCH 模拟收益率的自相关系数和偏自相关系数图，可以看出 SGED-GARCH 的模拟收益率存在 2 阶自相关性和 1 阶偏自相关性，这和真实样本之间还是存在较明显差异。

图 14: SGED-GARCH 模拟路径自相关系数和偏自相关系数图



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

2.2.3、模型构建小结

- 在这一部分，出于对收益率自相关性和波动聚集效应的考虑，我们分别构建了 SGED-ARMA 模型和 SGED-GARCH 模型，这两个模型能够分别在上述两个方面对 SGED 模型的效果进行提升。
- 从模拟效果来看，SGED-ARMA 模型使得模拟的收益率与真实样本收益率具有相同的一阶自相关性和一阶偏自相关性，但却缺乏了对波动聚集效应的考虑，当拟合周期拉长、数据频率增加时，模拟数据的波动聚集效应不再明显。SGED-GARCH 模型使得模拟的收益率与真实样本收益率具有更为接近的波动聚集效应，但在自相关性上与真实样本存在差异。
- 由于我们采用滚动拟合参数的方式来生成模拟数据，一定程度上变相地使模拟数据天然带有一定波动聚集的效果，不同模型的波动率聚集效应差异已经不是很大，所以在本篇报告后面的部分，我们主要应用 SGED-ARMA 模型来模拟数据，并以此展示模拟数据在策略回测方面的实用性。

3、提升回测参考价值：以 RSRs 为例

我们曾经在技术择时系列报告中提出了阻力支撑相对强度 (RSRS) 择时指标，并将其应用于各大指数择时，均得到了不错的策略收益表现。但在策略参数的选择上，我们是基于指数历史数据回测，进行优化处理的结果。所以我们的每一个 RSRs 指标策略其实只在一条样本路径上进行了测试，这

样的回测结果是否真实反映了该策略的择时能力？我们使用的参数过拟的程度有多少，样本外是否仍有稳定的效果？这是每个策略使用者关心的问题。

在这一部分，我们将利用 SGED-ARMA 模型生成 300 条模拟沪深 300 价格序列，重新回测 RSRs 模型，来一一审视以上提出的问题，同时展示 SGED 模型对于策略回测带来的助益。

3.1、阻力支撑相对强度原理及 RSRs 指标构建

我们先简单回顾一下 RSRs 指标的原理与构建方式。

在《基于阻力支撑相对强度（RSRs）的市场择时——技术择时系列报告之一》中，我们提出了利用最高价、最低价序列的相对强弱来动态刻画市场的阻力位与支撑位，我们认为阻力位与支撑位实质上反应了交易者对目前市场状态顶底的一种预期判断。如果这种预期判断极易改变，则表明支撑位或阻力位的强度小，有效性弱；如果众多交易者预期较为一致、变动不大，则表明支撑位或阻力位强度高，有效性强。

按照上述逻辑，当支撑位的强度弱于阻力位，则表明市场参与者对于支撑位的分歧大于对于阻力位的分歧，市场倾向于下跌；当支撑位的强度强于阻力位，则表示市场参与者对于支撑位的认可度高于对于阻力位的认可度，市场更倾向于上涨。

表 1：不同市场状态下阻力支撑相对强弱的含义

市场状态	支撑强度显著大于阻力强度	阻力强度显著大于支撑强度
上涨	价格加速上涨	上涨可能结束，转向下跌
震荡	向上突破，转向上涨	向下突破，转向下跌
下跌	下跌结束，市场见底	价格加速下跌

资料来源：光大证券研究所

从最高价与最低价的形成机制出发，每日的最高价与最低价就是一种阻力位与支撑位，它是当日全体市场参与者的交易行为所认可的阻力与支撑。当然我们并非用支撑位与阻力位作突破或反转交易的阈值，而是更关注市场参与者们对于阻力位与支撑位的定位一致性。当日最高价与最低价能迅速反应近期市场对于阻力位与支撑位态度的性质，是我们使用最高价与最低价的最重要原因。

根据上述逻辑，我们建立 RSRs 指标：

1. 取前 N 日的最高价序列与最低价序列。
2. 将两列数据按下式的模型进行滚动 OLS 线性回归。

$$High = \alpha + \beta * Low + \varepsilon, \varepsilon \sim N(0, \sigma) \quad \#(18)$$

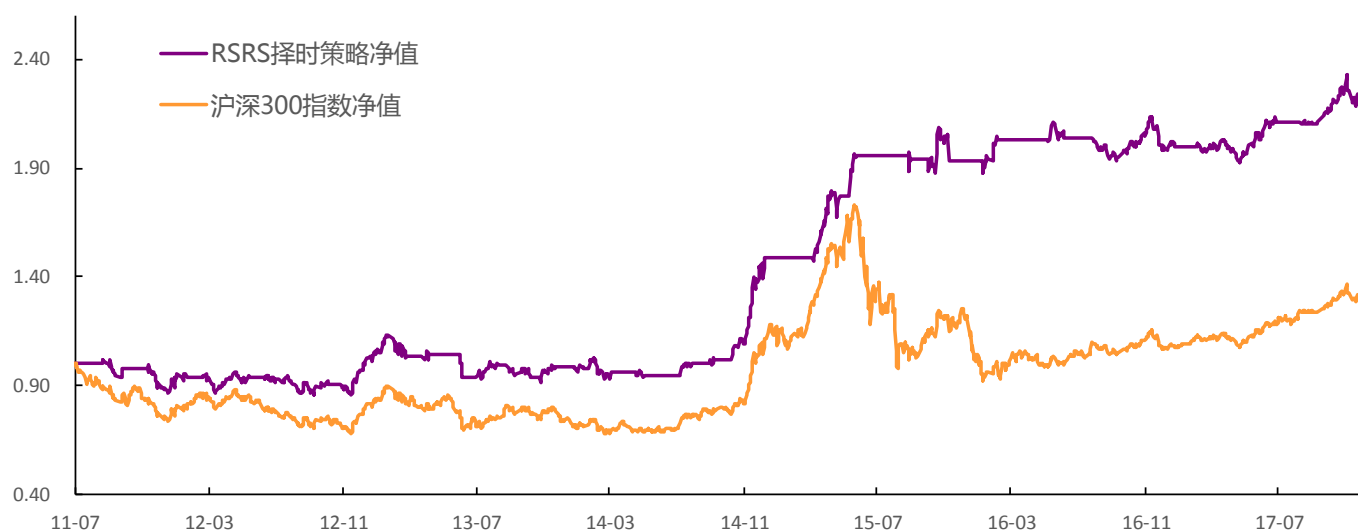
3. 取前 M 日的 β 时间序列，计算当日斜率的 z_score 标准分 z。
4. 将 z 与拟合方程的决定系数相乘，作为当日 RSRs 指标值

RSRS 择时策略则根据 RSRS 指标值的大小变化相应开平仓。当指标值上穿高点阈值 S 时开仓，下穿低点阈值 $-S$ 时平仓。

3.2、RSRS 择时模型有效性与参数过拟程度检验

在之前的报告中,沪深 300 RSRS 择时策略的参数设置为($N=18, M=600, S=0.7$), 这一策略在沪深 300 上的表现如下图

图 15: RSRS 择时策略在沪深 300 指数上的净值表现



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

沿用这一参数，将 RSRS 择时策略应用于 300 条 SGED-ARMA 模拟价格序列。从统计表现中我们可以看到，将 RSRS 择时策略运用于 SGED-ARMA300 条模拟价格序列之后，平均年化收益率 9.42%，略低于在实际沪深 300 指数上得到的 12.96%；年化超额收益平均只有 1.23%，平均夏普比率和最大回撤的表现也都有所下降，平均胜率有所上升、盈亏比下滑。而在交易次数、平均持仓天数等数据上，各序列上测试的 RSRS 择时策略都有接近的表现。这说明将 RSRS 择时策略应用于和沪深 300 指数价格有类似统计性质的价格序列上，仍有一定的择时效果，但基于沪深 300 历史价格数据回测的参数与结果夸大了其实际择时能力。

表 2: RSRS 策略在沪深 300 指数和 300 条模拟价格序列上的统计表现

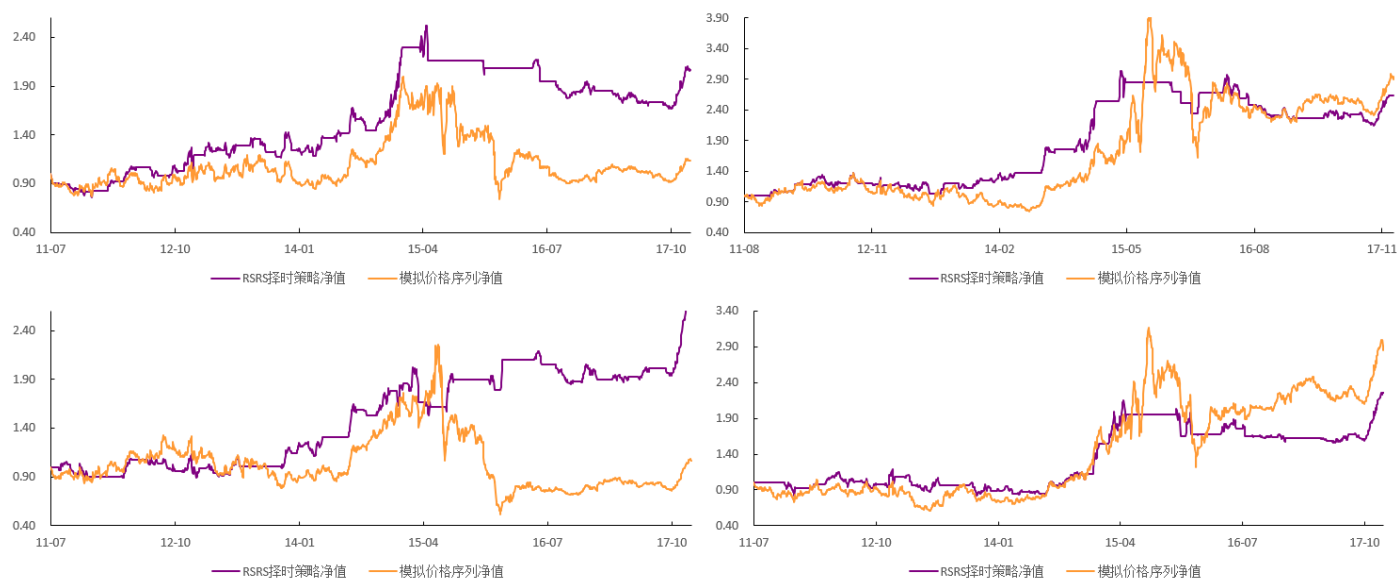
统计项	历史价格序列上 RSRS 策略效果	模拟路径上 RSRS 策略效果平均值
年化收益率	12.96%	9.42%
年化超额收益	8.84%	1.23%
夏普比率	0.89	0.51
最大回撤	-19.02%	-41.10%
持仓总天数	877	790

交易次数	26.00	27.83
平均持仓天数	33.73	28.38
盈利次数	13.00	15.35
亏损次数	13.00	12.48
单次最大盈利	45.47%	38.32%
单次最大亏损	-9.80%	-20.83%
胜率(按次)	50.00%	55.31%
平均盈亏比(按次)	3.88	1.74

资料来源：光大证券研究所

我们随机选取 4 条样本路径,展示了 RSRS 择时策略在这几条样本路径上的净值表现。整体而言,在 300 条模拟价格序列上,RSRS 择时策略在大多数情况下都有着不错的净值表现,但当模拟价格序列拥有比较大的价格波动时,我们的策略变得不再那么有效,策略收益净值不能跑赢标的价格序列。模拟价格序列存在大幅度价格波动可能也是导致超额年化收益降低的原因之一。

图 16: RSRS 择时策略在 SGED-ARMA 模拟价格序列上的净值表现



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

3.3、测试更合适的参数对

在上一部分,我们得到了将 RSRS 择时策略应用于和沪深 300 指数价格有类似统计性质的价格序列上,仍有一定择时效果的结论,但是在模拟序列上,其择时效果差强人意。一个可能的原因是,我们测试的参数是基于实际历史数据上优化得来的,这样的参数对很有可能有严重路径依赖从而在样本外无法得到样本内类似的表现。

接下来,我们将通过运用模拟序列一定程度上解决路径依赖的问题,寻找全局最优参数。参数的选取范围为:

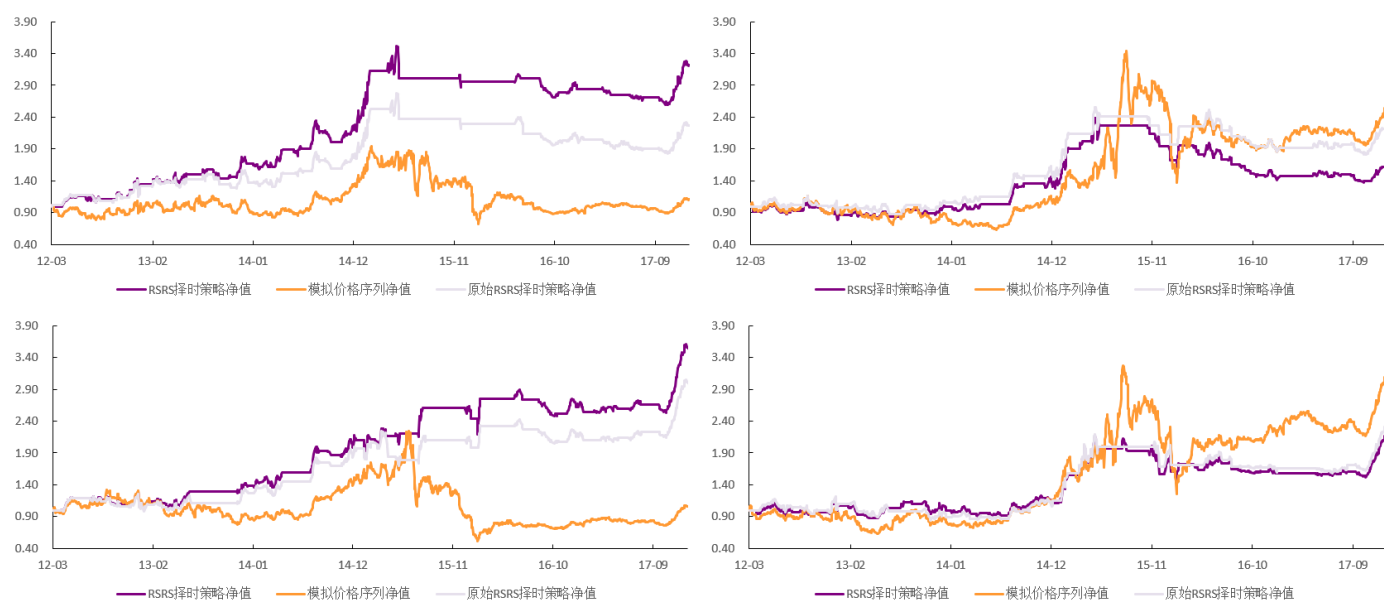
$$N = [15, 16, 17, 18, 19, 20, 21, 22, 23, 24]$$

$$M = [450, 500, 550, 600, 650, 700, 750, 800, 850, 900]$$

$$S = [0.4, 0.5, 0.6, 0.7, 0.8, 0.9]$$

在以所有模拟价格序列的平均年化超额收益作为选取最优参数的标准时，得到的最优参数为(N=17, M=750, S=0.6)。依旧选取之前的 4 条样本路径，下图展示了改变参数后 RSRs 择时策略在这几条样本路径上的净值表现：

图 17：改变参数后 RSRs 择时策略在 SGED-ARMA 模拟价格序列上的净值表现



资料来源：光大证券研究所，注：样本区间为 2010-01-01 至 2017-12-31

将这一修改参数后的 RSRs 择时模型应用于沪深 300 指数以及 300 条 SGED-ARMA 模拟价格序列，从统计数据来看，策略应用于模拟价格序列后有了更高的平均年化收益率和平均年化超额收益率，平均年化超额收益提高了 3.09%。而在其他统计指标上，改变参数对模拟路径的策略表现结果影响不大。同时可以看出在该参数下，策略在沪深 300 实际走势上的效果与在 300 条模拟路径上的择时效果相当，但明显低于之前参数下的择时效果。这中间的差距一定程度上可以理解为由路径依赖造成的过拟对策略实际择时能力的夸大程度。

表 3：改变参数后 RSRs 择时策略在沪深 300 指数和 300 条模拟价格序列上的统计表现

统计项	历史价格序列上 RSRs 策略效果	模拟路径上 RSRs 策略效果平均值
年化收益率	11.78%	10.46%
年化超额收益	4.30%	4.32%
夏普比率	0.83	0.55

最大回撤	-24.40%	-41.03%
持仓总天数	801	727
交易次数	32.00	29.35
平均持仓天数	25.03	24.77
盈利次数	15.00	15.84
亏损次数	17.00	13.52
单次最大盈利	40.80%	37.88%
单次最大亏损	-10.47%	-19.62%
胜率(按次)	46.88%	54.04%
平均盈亏比(按次)	3.11	1.80

资料来源：光大证券研究所

4、总结

- 在本篇报告中，我们首先从价格收益率存在的尖峰厚尾以及偏度的角度出发，引入了一个新的能够对收益率进行刻画的 SGED 模型，这一模型从收益率数据的高阶矩出发进行构造，相比于普通的正态分布，能够更好地考虑到价格收益率存在的一些统计特征。在模拟收益率数据时，采用滚动拟合的方式，并且提取每天的开盘价格进行了额外的拟合工作，以确保了数据构造的准确性。
- SGED 模型的模拟结果虽然接近真实路径，但在自相关性与偏自相关性，以及波动聚集效应上却和真实数据有所差异。出于对收益率自相关性和波动聚集效应的考虑，本篇报告分别构建了 SGED-ARMA 模型和 SGED-GARCH 模型，这两个模型能够分别在上述两个方面对 SGED 模型的效果进行提升。
- 从模拟效果来看，SGED-ARMA 模型使得模拟的收益率与真实样本收益率具有相同的一阶自相关性和一阶偏自相关性，但却缺乏了对波动聚集效应的考虑，当拟合周期拉长时，模拟数据的波动聚集效应不再明显。SGED-GARCH 模型使得模拟的收益率与真实样本收益率具有为接近的波动聚集效应，但在自相关性上与真实样本存在差异。
- 在应用层面上，运用模拟路径可以弥补回测样本少及路径依赖等回测问题，一定程度上减少策略的过拟风险，更有效地评估策略的实际价值。

5、风险提示

本报告中的测试结果均基于模型和历史数据，历史数据存在不被重复验证的可能，模型存在失效的风险。

行业及公司评级体系

评级	说明
买入	未来 6-12 个月的投资收益率领先市场基准指数 15% 以上；
增持	未来 6-12 个月的投资收益率领先市场基准指数 5% 至 15%；
中性	未来 6-12 个月的投资收益率与市场基准指数的变动幅度相差 -5% 至 5%；
减持	未来 6-12 个月的投资收益率落后市场基准指数 5% 至 15%；
卖出	未来 6-12 个月的投资收益率落后市场基准指数 15% 以上；
无评级	因无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使无法给出明确的投资评级。

基准指数说明：A 股主板基准为沪深 300 指数；中小盘基准为中小板指；创业板基准为创业板指；新三板基准为新三板指数；港股基准指数为恒生指数。

分析、估值方法的局限性说明

本报告所包含的分析基于各种假设，不同假设可能导致分析结果出现重大不同。本报告采用的各种估值方法及模型均有其局限性，估值结果不保证所涉及证券能够在该价格交易。

分析师声明

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，并对本报告的内容和观点负责。负责准备以及撰写本报告的所有研究人员在此保证，本研究报告中任何关于发行商或证券所发表的观点均如实反映研究人员的个人观点。研究人员获取报酬的评判因素包括研究的质量和准确性、客户反馈、竞争性因素以及光大证券股份有限公司的整体收益。所有研究人员保证他们报酬的任何一部分不曾与、不与、也将不会与本报告中的具体的推荐意见或观点有直接或间接的联系。

特别声明

光大证券股份有限公司（以下简称“本公司”）创建于 1996 年，系由中国光大（集团）总公司投资控股的全国性综合类股份制证券公司，是中国证监会批准的首批三家创新试点公司之一。根据中国证监会核发的经营证券期货业务许可，本公司的经营范围包括证券投资咨询业务。

本公司经营范围：证券经纪；证券投资咨询；与证券交易、证券投资活动有关的财务顾问；证券承销与保荐；证券自营；为期货公司提供中间介绍业务；证券投资基金代销；融资融券业务；中国证监会批准的其他业务。此外，本公司还通过全资或控股子公司开展资产管理、直接投资、期货、基金管理以及香港证券业务。

本报告由光大证券股份有限公司研究所（以下简称“光大证券研究所”）编写，以合法获得的我们相信为可靠、准确、完整的信息为基础，但不保证我们所获得的原始信息以及报告所载信息之准确性和完整性。光大证券研究所可能将不时补充、修订或更新有关信息，但不保证及时发布该等更新。

本报告中的资料、意见、预测均反映报告初次发布时光大证券研究所的判断，可能需随时进行调整且不予通知。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。客户应自主作出投资决策并自行承担投资风险。本报告中的信息或所表述的意见并未考虑到个别投资者的具体投资目的、财务状况以及特定需求。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。对依据或者使用本报告所造成的一切后果，本公司及作者均不承担任何法律责任。

不同时期，本公司可能会撰写并发布与本报告所载信息、建议及预测不一致的报告。本公司的销售人员、交易人员和其他专业人员可能会向客户提供与本报告中所载观点不同的口头或书面评论或交易策略。本公司的资产管理子公司、自营部门以及其他投资业务板块可能会独立做出与本报告的意见或建议不相一致的投资决策。本公司提醒投资者注意并理解投资证券及投资产品存在的风险，在做出投资决策前，建议投资者务必向专业人士咨询并谨慎抉择。

在法律允许的情况下，本公司及其附属机构可能持有报告中提及的公司所发行证券的头寸并进行交易，也可能为这些公司提供或正在争取提供投资银行、财务顾问或金融产品等相关服务。投资者应当充分考虑本公司及本公司附属机构就报告内容可能存在的利益冲突，勿将本报告作为投资决策的唯一信赖依据。

本报告根据中华人民共和国法律在中华人民共和国境内分发，仅向特定客户传送。本报告的版权仅归本公司所有，未经书面许可，任何机构和个人不得以任何形式、任何目的进行翻版、复制、转载、刊登、发表、篡改或引用。如因侵权行为给本公司造成任何直接或间接的损失，本公司保留追究一切法律责任的权利。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

光大证券股份有限公司 2019 版权所有。

联系我们

上海	北京	深圳
静安区南京西路 1266 号恒隆广场 1 号 写字楼 48 层	西城区月坛北街 2 号月坛大厦东配楼 2 层 复兴门外大街 6 号光大大厦 17 层	福田区深南大道 6011 号 NEO 绿景纪元大厦 A 座 17 楼