第一章　稳健的投资组合优化

1.1　综述

马克维茨投资组合理论并且其在R中的应用已经在第五章中进行了相关的介绍。运用投资组合直接优化所可能遇到的问题，得到了解决。尤其当采用样本估计均值和协方差阵时会由于估计误差从而导致次最优的结果。此外，极端的组合权重以及资产组合不稳定的波动常常可以在事后的模拟中被观察到。另外一般来说，由于交易费用的存在会使得经验事实不可取。从统计学的观点来看，出现这种假象主要归于普通样本估计关于异常点的灵敏度。在其他条件不变的情况下，这些离群数据点对离差估计的影响程度要小于对均值的影响。因此，但不仅仅由于这些原因，最小方差组合也被提倡与均值方差组合比较(具体见第五章的参考文献)。因此需要得到一个受离群值影响较小的可用估计量并且由此产生的估计可以代表大量的样本数据，并且优化方法能直接包含估计的误差。前者通过采用稳健统计学的方法完成，而后者的实现可以通过使用稳健优化(鲁棒优化)方法。

在接下来的两节中将会提出两种从理论观点出发的方法，并讨论其在R中的实现。这一章还给出了基于蒙特卡洛模拟与回测比较的实证应用，包括了稳健投资组合优化的解与普通样本估计解的对比。

1.2　稳健统计

1.2.1　动机

在第三章中已经指出正态假设很多时候很难适用于金融市场中的数据(Pfaff 2010, 第一章有类似叙述)。并且根据一些典型的关于单变量和多变量收益情况的经验看违反这一假设是合乎情理的。但是在第六章中也可以看到数据频率越低其对正态假设的违反程度就越小。

在实证工作中算术平均和样本方差被广泛用于估计对应总体的理论均值和方差。这些估计值可以通过极大似然原理推导得到并且具有渐进一致性和有效性。而且，可以认为它们是渐进正态分布的。然而，当没有对分布的假设时，这些估计量就不再具有这些良好的性质。事实上，算术平均作为对总体均值的估计其对极端观察值是非常敏感的，以至于有时候估计无法很好的反应出大部分数据的特点。类似的，两个随机变量的相关性分析当遇到离群的数据对时就会严重偏离真实值。鉴于此，非常需要探索一种方法可以对离群值免疫并且不需要依赖于模型假设，但是其仍然能传递主要样本观察值的合理分析结果。稳健统计学领域就是用于处理该类问题并且给出以稳健估计值及基于此的推断等形式的解决方案。可以从相关的教材中找到关于稳健统计学的阐述，例如Huber(1981), Hampel $etal.$(1986), Rousseeuw and Leroy (1987), Staudte and Sheather (1990) and Maronna $etal.$(2006)。

以前，对于之前描述的离群值问题一般都通过整理(剔除离群值)或者尾处理(采用一个固定的分位数值来均衡异常观察值)两种方法解决。事实上，这两种方法都可以认为是稳健的方法。不幸运的是这两种处理方法的结果都很大程度上依赖于研究者的主观选择。中位数可能会被用于作为

均值和平均绝对偏差范围的估计。相比算术平均和标准差等普通样本估计而言它们都可以衍生出更多的稳健估计，这些稳健估计量的优点在于其避免了提前指定那些数据点为离群值。因此这样就避免了数据分析中的主观因素。在下一小节介绍最常碰到的稳健统计学方法前，需要注意目前我们还没有对“离群值”和“极端观察值”作出精确的定义。原因很简单，现在没有一种准确清晰的方法去评价某个数据点是否为离群值。这个问题总是相对而言的并且关键依赖于模型或者分布的假设。例如，给出一个标准正态分布以及一个样本观察值5，我们可以非常确定的将其划分为离群值。但是这样的划分当分布的假设是自由为4的t分布或者柯西分布时还成立么?

1.2.2 选择稳健估计量

接下来将会介绍选择稳健估计量以及它们的性质基于Maronna $et al.$(2006), Todorov and Filzmoser(2009)。最常用的稳健性评价估计量是崩溃点。这种方法被定义为样本中离群点的相对份额使得估计量不会出现任意大的值。根据定义，BP的取值范围为0到0.5。算数平均的BP 是0，因为如果单个观察值被某个值替代，均值估计会有明显的波动。与此想法，中位数的BP是0.5。BP的上界可以这么解释，如果超过一半的观察值都是离群值，那么这个样本已经被一定的篡改以至于无法根据总体作出推断。进一步评价稳健估计量合适与否的标准是相对效率(relative efficiency)。这里稳健估计量的方差通过相对于基于严格模型/分布假设推导所得最优估计量的方差表示。同样地，它可以被解释成一个百分数，即表示必须要增加多少样本量才可以使两者的方差相等。

1.2.3 M-与MM类估计量

早在1964年Huber就已经提出了M-估计量。类名表达了这类估计量与极大似然法则的相似之处(见Huber 1964, 1981)，这对于接下来所叙述的单变量样本估计量是非常方便约束的。根据极大似然法则，未知的参数θ被决定后会使得其最可能产生取自分布函数为$F(x,\theta)$(对应密度函数为$f(\cdot)$)的简单随机样本(i.i.d.)$\{x_1,\cdots,x_n\}$。由于i.i.d. 的假设，故联合分布等于边缘分布的乘积，它的极大值点为:

$$\hat{\theta} = \arg\max(\prod_{i=1}^{n} f(x_i,\theta))$$

由于对数变换是严格单调的，因此优化问题一般可以转化为求负对数似然函数的极小值点,

$$\hat{\theta} = \arg\max(-\sum_{i=1}^{n} \log(f(x_i,\theta)))$$

类似的，M-估计量可以定义为函数$\rho(x,\theta)$和的极小值点,

$$\hat{\theta} = \arg\max(-\sum_{i=1}^{n} \rho(x_i,\theta))$$

从广义的角度来看这类估计量包含了极大似然估计即$\rho(\cdot) = -\log f(x,\theta)$以及最小二乘法(LS)，后者的$\rho(\cdot)$为误差平方。函数$\rho(\cdot)$必须满足对称性、正定性、全局最小值为零(大于等于零)，当然当存在模型/分布假设且不受反例影响(不违反假设)时该函数可以给出较好的估计。M估计量和极大似然与最小二乘法的区别在于$\rho(\cdot)$的形式。对于前者(M估计量)极端数据点占据了很小的权重因此对于参数估计的影响更小。在实证工作中有两种形式的$\rho(\cdot)$被经常使用。首先介绍Huber函数，其被定义为

$$\rho_k(x) = \begin{cases} x^2 & \text{if}|x| \le k \\ 2k|x| - k^2 & \text{if}|x| > k. \end{cases}$$

该函数在中心区域($|x| \le k$)是二次的，而在外部区域$|x| > k$则是线性的。当$k \to \infty$和$k \to 0$时M估计量将分别等价于算术平均和中位数。此外，由Tukey提出的$bi-square$是另一种在实证工作中常常遇到的函数。给出的形式为

$$\rho_k(x) = \begin{cases} 1-[1-(x/k)^2]^3 & \text{if} |x| \le k \\ 1 & \text{if} |x| > k. \end{cases}$$

该函数对任意大x的绝对值有界。对于具有高狭峰对称分布来说相比Huber 函数而言$bi-square$函数更合适，因为此时可以完全克服离群值的影响。

直接代入方程(10.3)，通过使梯度$\psi(x) = \delta\rho(x,\theta)/\delta\theta$为零，一般容易找到方程的解。

MM类估计量最早由Yohai(1987,1991)在关于稳健回归分析的文章中提出，Lopuhaä将该估计量推广至多元数据分析。第一步利用M估计量对数据的离差进行估计，之后基于此就实现了对位置参数的又一种类型的M估计，其紧密地结合了第一步中的方差——协方差矩阵。

1.2.4 基于robust scaling的估计量

有一类可由robust scaling的推导得到的稳健估计量也非常常见。分别被称为MVE估计量($minimumvolumeellipsoid$)，MCD估计量($minimumcovariancedeterminant$)和S估计量。考虑一个p维随机矩阵$\mathbf{x} = (x_1, \cdots, x_p)' \in \mathbb{R}^p$，服从多元正态分布，即$\mathbf{x} \sim \mathfrak{N}(\mu, \Sigma)$。给出均值向量$\mu = \mathrm{E}(\mathbf{x}) = (\mathrm{E}(x_1), \cdots, \mathrm{E}(x_p))'$，协方差矩阵$\Sigma = \mathrm{Var}(x) = \mathrm{E}((\mathrm{x}-\mu)(\mathrm{x}-\mu)')$。距离通过样本间马氏距离度量，即：$d(\mathbf{x}, \mu, \Sigma) = (\mathrm{x}-\mu)'\Sigma^{-1}(\mathrm{x}-\mu)$。当估计的协方差矩阵的最小特征值趋于零时可以得到使距离为零的平凡解。为了排除这种情况，要求$|\hat{\Sigma}| = 1$，这里$|\cdot|$ 表示矩阵的行列式。之前所提到的三种针对$\hat{\mu}$和$\hat{\Sigma}$的稳健估计量都是基于下面的优化问题：

$$\arg\min \hat{\sigma}(\mathbf{d}(\mathbf{X}, \hat{mu}, \hat{\Sigma}))$$

这里$\mathbf{X}$是全样本，$\mathbf{d}$是距离向量即$d(\mathbf{x}_i, \hat{mu}, \hat{\Sigma})$，$i = p+1, \cdots, N$，$\hat{\sigma}$则是一个稳健度量。

MVE和MCD估计量由Rousseeuw (1985)以及Rousseeuw和Leroy (1987)。在使用MVE估计量时最小的数据点集需满足条件即至少有一半的样本观察值被包含用以计算。其通过利用$\hat{\sigma}$的中位数实现，该估计量的收敛速度仅仅是$N^{\frac{1}{3}}$。在使用MCD稳健估计量时需选取长度$h > N/2$的子样本使方差-协方差矩阵的行列式最小。均值向量取决于所选数据点的p算数平均，而方差-协方差矩阵则根据修正因子调整，使最终的估计$\hat{\Sigma}$服从正态分布。BP是$h = \int[(N+p+1)/2]$的最大值，但在区间$[(N+p+1)/2, N]$上可以取任何值。Davies(1987)介绍了S类估计量。字母"S"表示$scale$。关于上述的优化，按照$\frac{1}{N}\sum_{i=1}^{N}\rho(d_i)$ 插入对规模的M 估计量，这里$\rho(\cdot)$必须是一个有界函数且$\delta \in (0,1)$。

Stahel-Donoho估计量

Stahel-Donoho估计量由Stahel和Donoho分别在1981年和1982年提出。根据其过程，离群程度取决于数据矩阵X 与向量$a \in \mathbb{R}^p$的一个数学规划。该向量由给定的观察值$\mathbf{x}$，关于完整的数据集X的稳健的均值和协方差阵估计量$\hat{\mu}$和$\hat{\Sigma}$ 为

$$t(\mathbf{x}, \mathbf{a}) = \frac{\mathbf{x}'\mathbf{a} - \hat{\mu}(X\mathbf{a})}{\hat{\sigma}(X\mathbf{a})}$$

离群程度由上述方程取最大值的解(对于所有归一化后向量$\mathbf{a}$)。对应的解即作为X的权重使采用的函数$t(\mathbf{x}_i)$非增。

OGK估计量

具有仿射不变性的鲁棒估计量的一个问题为结果的非凸优化。Maronna和Zamar提出了可以在估计两随机变量协方差时避免该问题的估计量

$$s_{jk} = \frac{1}{4}(\sigma[\frac{Y_j}{\sigma(Y_j)} + \frac{Y_k}{\sigma(Y_k)}]^2 - \sigma[\frac{Y_j}{\sigma(Y_j)} - \frac{Y_k}{\sigma(Y_k)}]^2)$$

这种形式的分散估计由Gnanadesikan和Kettenring(1972)提出。如果协方差由上述方程决定，那么仿射不变性可以抵消掉简单的计算。然而，如果一旦针对X的数据对直接应用分散估计量，那么导出的方差-协方差矩阵将不再正定。因为这一原因，Maronna和Zamar(2002)提出了X的正交化并因此把该估计量称之为*orthogonalized Gnanadesikan - Kettenring* (OGK)。如果σ 的稳健估计量被用于成对协方差s_{jk},$j = 1, \cdots, p$ 且$k = 1, \cdots, p$，那么导出的方差-协方差矩阵也是稳健的。

1.3 稳健/鲁棒优化

1.3.1 动机

在这一节中介绍用于优化投资组合的鲁棒方法和技术。之前的章节中采用了稳健估计量而不是经典估计量来解决未知的位置参数与离差参数估计。不管哪种情况下最后的估计 — 无论是通过极大似然获得还是通过鲁棒估计方法 — 都是作为固定的且是被直接利用的数学程序输入。然而在这些情况中参数值常常是不确定的。与此相反我们将会介绍和研究数学程序的优化方法。现在"鲁棒(稳健)"的定义将与之前章节有所不同，之前用于减少离群值影响的点估计并与经典的极大似然估计量对比。现在它被定义为一种优化方法，其可以给出不会因替代参数产生负面影响的解，例如当期望收益的结果并不那么令人满意时。此外严格意义上鲁棒优化方法与随机优化不同因为后者基于对参数特定的分布假设，而在鲁棒优化却一般不要求。简单地说，鲁棒优化的目标是在可行的参数集中导出最优的解。关于方法上的一些问题和困难的进一步介绍在Ben-Tal与Nemirovski(1998), Cornuejols与T*ütün*ü(2007), Fabozzi*et al.*(2007), Meucci(2005,Chapter 9), Scherer(2010,Chapter 5)与Tünt*ücü*与König (2004)等资料中给出。

1.3.2

鲁棒优化的概念可以通过均值-方差投资组合阐明，虽然该方法还可以在其他种类的优化中应用。经典的投资组合优化方法通过下式给出

$$R_\lambda = \arg\min(1-\omega)\sqrt{\omega'\Sigma\omega} - \lambda\omega'\mu$$

这里ω为长度为$N \times 1$的资产组合权重向量，Ω是可行解的集合。N种资产的期望收益包含在了向量μ中，并认为方差- 协方差矩阵正定。参数λ允许在[0,1]任意取值并且决定了投资组合的收益与风险间的权重。当然上面的问题表示可以用进一步的约束条件改进，例如权重的约束，非负的约束以及对单个资产权重及部分资产组合的约束等等。上述公式中的数学程序包含了资产组合最小方差与最大收益率两种特定情形，即当λ的取值分别为$\lambda = 0$与$\lambda = 1$时。

目前对于未知参数(μ, Σ)的点估计已经被广泛使用，既可以由经典估计量导出也可以由稳健估计量导出，正如之前章节的那样。此外

1.4 R包概要

在这一节中，只介绍在多元数据分析中可以部分或全部用于稳健估计方法的R包。因此，R包可以覆盖稳健回归，稳健时间序列分析以及其他为包括在这份概要中的稳健方法的用途，读者可以参考CRAN 上的关于概述用于鲁棒方法R 包的Task View。

1.4.1 covRobust包

covRobust包由Wang在2003年编写。该包收录在了CRAN上“鲁棒”和“多元”两个Task Views中。包中只包含了cov.nnve()一个函数，用于实现Wang和Raftery(2002)所介绍的协方差和均值的估计量。离群值通过最近邻方法检测因此函数名的后缀事实上是“nearest neighbour variance estimator”(最近邻方差估计量)的首字母缩写。该函数的特点在于有两个主要的参数datamat和k分别对应于观测矩阵以及考虑所选择的邻近点的数目，缺省值为12。剩下的参数可以用来控制和调整优化的过程。函数的返回形式是一个带有一下内容的列表：稳健协方差估计(cov)，稳健均值向量估计(mu)以及后验概率(postprob)也就是数据点被判断为离群点的概率(list)。列表中的元素classification包含了(0,1)离散化的后验概率。最后，最近邻清除的初始结果包含在了列表的innc中，当然innc本身为一包含了上述四个方面的列表。

1.4.2 fPortfolio包

fPortfolio包是一个可以用于执行很多不同种类的资产组合优化任务的强大的R包。它是Rmetrics捆绑包中的一部分，并且该包被认为是CRAN 上“金融”Task View的核心包。Würtz等对该包的功能有一个简洁的描述。在这一节中，我们更关心稳健估计量是如何在投资组合优化中被使用的，在第12章中对该扩展包有一个更深入的讨论。

对象类型`fPFOLIOSPEC`的关键位置是`model`。这一位置是它自己的一个列表对象并且包含了名为`estimator`的元素。它的值是取决于用于估计资产回报均值与协方差阵的函数，并且它的缺省值是引用于经典积矩估计量的`covEstimator`。对一个给定资产组合规范的估计量的类型可以通过函数`getEstimator()`查询，同时也可以通过函数`setEstimator()`调整改变。The latter function expects the name of a function which returns estimates for the locations and dispersion.它会返回一个带有元素`mu`和`sigma`即分别表示均值和方差的列表对象。在fPortfolio包中用于鲁棒控制投资组合的函数是`kendallEstimator()`, `spearmanEstimator()`, `mcdEstimator()`, `mveEstimator()`, `covMcdEstimator()`以及`covOGKEstimator()`。前两个函数采用了Kendall和Spearman的秩相关作为资产回报率间相关性的度量的方式。而`mveEstimator()`和`covMcdEstimator()`的不同点在于前者调用了MASS包中的函数而后者使用了robustbase包的函数。类似地，MVE估计量的函数`mveEstimator()`源于MASS包而OGK估计量的函数`covOGKEstimator()`则是源于robustbase包。然而，这里需要再次强调的是使用者可以提供自己的估计量并且在Würtz(2010, 20章)给出的例子中即是源自他自己阐述的方法。

1.4.3 MASS包

MASS包是伴随《Modern Applied Statistics with S》一书(Venables和Ripley, 2002)而产生的包。由于它的重要性以及涉及到广泛的统计学方法，该包是R中很基础的一部分。MASS包包含在了CRAN上的"Distributions","Econometrics", "Environmetrics","Multivariate","Pharmaconetics", "Psychometrics","Robust" 以及"SocialSciences"等Task View中。包中使用S3类和方法并自带了`NAMESPACE`文件。下面只叙述跟鲁棒统计相关的函数。

Huber M-估计量通过函数huber()和hubers()实现，前者估计数据向量的位置参数。平均绝对偏差(MAD)被用作规模估计并且对样本尾处理的缺省值设为$k = 1.5$。后者能够被用于特定的位置或范围参数或使用者没有提供的情形。两者都返回一带有位置估计和规模估计元素的列表对象。

对于多元数据，均值向量和协方差阵的稳健估计可以通过函数`cov.rob()`计算得到。可以使用两种方法，分别是最小椭圆体积和最小协方差行列式，前者是缺省方法。一只利用积矩实现的经典协方差估计可以通过设置参数`method=’ classical’`实现。函数`cov.mve()`与`cob.mcd()`是对前两种方法的封装。该函数返回一个包含均值(`center`)和离差阵(`cov`)的列表对象。如果相关阵也

需要计算的话(缺省情况是`cor=FALSE`)，那返回的列表还会包含元素`cor`。剩下的一些列表项则包含了数值优化的结果信息。

最后，多元t分布的离差阵可以通过函数cov.trob()估计。t分布提供了相比标准正态分布的稳健性即考虑到由厚尾产生的离群点。该函数允许对样本观察值加入权重向量。自由度可以通过参数`nu`指定；其缺省值为5。参数`cov`是一个逻辑变量(缺省值为`FALSE`)，其提供了与函数`cov.rob`一样的用途。其优化过程可以通过参数`maxit` 和`tol`控制，其分别限制最大迭代次数和收敛条件；缺省值分别为25和0.01。该函数的返回一个列表对象，且参数`cov`,`center`,`cor`分别对应表示了方差，位置参数以及协方差的估计。类似于函数`cov.rob()`，剩余的列表元素还包含了数值优化的信息。

1.4.4 robustbase包

robustbase包包含了基本的和主要的用于稳健估计的函数(Rousseeuw等，2012)。该包在鲁棒统计中可以称之为构建其他包的基础材料。同时该包包含在CRAN的‘Multivariate’、‘Robust’以及‘SocialSciences’三个Task Views中。它使用S4方法和类，但也支持S3方法。对于一些需要繁重计算的鲁棒方法，尤其是线性模型，通过与C语言的交互实现。

关于鲁棒统计的函数`covMcd()`和`covOGK()`可以用于估计离差。前者是最小协方差行列式估计(MCD)的实现，而后者则实现OGK的估计量。两者都返回一列表对象，且从中可提取位置参数(`center`)和离差(`cov`)估计。

广义Huber M估计量可以通过函数`huberM()`执行。它返回一个包含鲁棒位置估计(`mu`)以及范围估计(`s`)的列表，这里MAD作为缺省设置，且迭代次数返回在列表元素`it`中。

1.4.5 robust包

robust包is an R port of the Insightful package of the same name (see Wang et al. 2010)。该包是针对鲁棒方法和鲁棒统计的一般使用者。它是CRAN上‘Robust’和‘SocialSciences’两个Task View的一部分，并且被视为核心部分，类似于robustbase包对于‘Robust’。该包依赖于包MASS,lattice,robustbase,rrcov,stats。线性模型和广义线性模型的估计通过与C语言交互实现。

针对鲁棒多元统计可以利用函数`covRob()`。这里用于鲁棒统计的数据要求被转化为矩阵或者数据框。使用者可以通过设置逻辑参数`corr`选择返回协方差还是相关系数，缺省情况是计算方差-协方差矩阵。此外，参数`distance`可以被用作决定是否采用马氏距离计算的开关(逻辑参数)，缺省值是`TRUE`。

缺失数据的值可以通过针对参数`na.action`指定一个函数处理；缺省方式是`na.fail`，`na.omit`可以被替代使用。被使用的鲁棒估计方法通过参数`estim`决定。并通过特定的字符串在如下几种多元估计量中选择：`'mcd'`对应快速MCD算法;’`weighted`’对应重加权MCD估计;’`donostah`’针对Stahel - Donoho估计量;’`M`’针对M-估计量;’`pairwiseQC`’针对orthogonalized quadrant correlation 估计量;’`pairwiseGK`’针对orthogonalized Gnanadesikan - Kettenring估计量。该参数的缺省值是’`auto`’，在这种情况下函数在Stahel - Donoho, MCD或者pairwise quadrant correlation估计量中通过权衡计算时间与对位置和离差估计的合理性自动选择。使用Stahel - Donoho估计量的前提是数据集包含少于1000个观察值少于10个变量或少于5000个观察值且少于五个变量。如果变量总数大于10小于20且最大不超过5000个数据向量，那么可以使用MCD估计量。在所有的高维情形下可以采用orthogonalized pairwise correlation方法。MCD估计量则是基于robustbase包和接下来介绍的rrcov包(见Section 10.4.6)。函数`covRob()`接受一个对象用于控制鲁棒估计c。该函数型参数可以通过函数`covRob.control()`构造，其返回一个控制参数的列表。这些同样也可以通过省略号参数当`estim`没有设置为它的缺省值’auto’。

函数返回一个S3类的列表对象`covRob`。列表元素`call`说明使用者如何调用该函数，也就是说使用参数的种类。因此，泛型函数`update()`可以被用于该对象。离差的估计(协方差或相关系

数)则是作为元素`cov`返回，位置向量则是作为元素`center`返回。如果马氏距离可以被计算，那么它会出现列表元素`dist`。如果鲁棒估计算法可以被初始化(初值)，那么它们的值以`raw.cov`、`raw.center`、`raw.dist`返回，分别对应离差、位置向量以及距离。所选估计量的名字则是作为列表元素`estim`返回，控制参数则是作为列表对象`control`。常用类的对象都会支持verb"print()"、`plot()`、`summary()`方法，并且后者还有它们自己的`print()`方法。对于`plot()`方法，使用者可以交互式地在展现特征值、马氏距离平方根以及椭圆图中选择。离差阵的特征值估计包含在常用S3类对象`summary.covRob`中。

最后，函数`ccov()`是经典估计量的实现工具并返回一个常用S3类`cov`的列表对象。包开发者的主要目的是提供一个函数可以返回一个对象结构上类似于S3类对象`covRob`，并由此促成一种简单的估计对比。这些对象同样都支持verb"print()"、`plot()`、`summary()`方法。

1.4.6 rrcov包

rrcov包专门单独用于鲁棒多元统计的包。同样它依赖于robustbase包，但是提供了一种自包含的S4对象和方法实现鲁棒估计量。进一步依赖的包还有methods、pcaPP及mvtnorm。该包包含在CRAN上的‘Multivariate’和‘Robust’两项Task View中。并且在后者中该包被认为是一个核心的包。

除鲁棒多元统计外，在包中基于鲁棒方法的主成分分析和线性判别分析同样也可以实现，但是仅会提供鲁棒估计量。

类似于robust包，均值和方差的经典估计量可以通过函数`CovClassic()`实现，其返回一个S4对象`CovClassic`。使用者可以通过设置逻辑参数`unbiased`选择是抽样还是总体估计，缺省值是`TRUE`。

使用者同样可以在五种不同的绘图

1.4.7 Rsocp

Psocp包是Rmetrics suite of packages的一部分，在写这本书的时候，该包还在R-forge上。包中封装的函数用于`socp`(Lobo)软件的实现。求解SOCP 的算法通过C 语言开发。相关C 代码一般可以通过R语言交互。除这些高层语言函数外，作者还提供了matlab脚本。原始软件仍然可以在`http://stanford.edu/?boyd/old_software/socp/` 获取。

包中包含了两个函数：`socp()`与`socpControl()`。前者通过C语言封装。

1.5 实证应用

1.5.1 投资组合模拟:鲁棒统计VS经典统计

在第一个实证应用中进行了对比经典估计量与鲁棒估计量的模拟。

首先，把所需要的包加载至工作空间。通过包含在copular包(见9.4.2)中的函数生成随机样本。之后，

对于第二个DGP，

1.5.2 投资组合回测：鲁棒统计VS经典统计

在这一节中将考虑最小方差投资组合策略与主要股指充分投资的回测。对于将经典估计量和鲁棒估计量用于主要的估计中。涉及的六种股指为S&P 500, Nikkei 225, FTSE 100, CAC 40, DAX and Hang Seng。月末指数的样本从1991年7月至2011年6月30日，因此包含了240个观测值。

在表10.5中首先把必要的包加载至内存中。伴随这本书的数据集包含在了FRAPO包中。一般来说，投资组合回测可以

包载入内存后，将数据集导入并将其转化为zoo对象。

盒状图见图10.1

计算回测的代码见表10.6。首先，一个字符向量

1.5.3 投资组合回测：鲁棒优化