基于系统性风险和夏普指数的社保基金投资绩效分析

杜子平(博士生导师)　方兴兴

(天津科技大学经济与管理学院　天津　300222)

【摘要】本文利用资本资产定价模型(CAPM)中的贝塔系数来刻画全国社保基金投资组合的系统性风险,通过与市场收益系统性风险的大小做对比来说明投资组合的整体风险状况。并且以CVaR代替VaR计算夏普指数的变形形式(RAROC)来反映投资组合的风险收益状况。而CVaR则是利用Pair - Copula拟合投资组合的分布之后,通过蒙特卡罗模拟求出的。

【关键词】Pair - Copula　CVaR　RAROC　贝塔系数　投资绩效　系统性风险

一、引言

1964年夏普(sharp)提出的资本资产定价模型(CAPM)刻画了资产收益与系统性风险之间的关系,并采用贝塔系数对资产的系统性风险进行测度。该模型认为资产的回报率只因它们系统性风险的不同而不同。其模型公式:$R_i=\alpha_i+\beta_i R_M+\varepsilon_i$,其中:$R_i$表示资产收益;$R_M$表示市场收益;$\varepsilon_i$表示误差项。等式两边取方差得到:$Var(R_i)=\beta_i^2\sigma_M^2+\sigma^2_{\varepsilon_i}$,其中:$\beta_i^2\sigma_M^2$表示资产i的系统性风险,$\sigma^2_{\varepsilon_i}$表示资产i的非系统性风险。又因为$\beta_i=cov(R_i,R_M)_i/\sigma^2_M=\rho_{iM}\cdot\sigma_i\cdot\sigma_m/\sigma^2_m$,由此得到$\beta^2_i\sigma^2_M/\sigma^2_i=\rho^2_{iM}$。不难发现可以用$\rho^2_{iM}$表示资产i的系统性风险占总风险的大小,则非系统性风险占总风险的大小为:$1-\rho^2_{iM}$。

在经典投资基金投资绩效评价模型中,夏普指数是一个具有代表性的评价模型,其公式为:$S_p=(R_p-R_f)/s_p$,其中:S_p表示夏普业绩指数;R_p表示基金的平均收益率;R_f表示基金的平均无风险利率;s_p表示投资收益的标准差。不过,以标准差作为风险考虑了收益向上变动的正向收益(上行风险),这显然不符合实际投资者的心理。

1994年,JP摩根公司率先提出了VaR的概念,它是投资者关心的下行风险。基于此提出了经风险调整的收益(RAROC),它表示单位损失资本的回报大小,有效地度量了获取收益的风险效率。但以VaR作为下行风险存在以下缺陷:①VaR在非正太分布下不满足次可加性,不能作为一致性风险度量工具;②VaR只是一个分位数,并不一定具有连续性;③VaR只是一个分位点(即最大预计损失),分位点下面的情况则无法得知。针对VaR的这些不足,条件风险价值即CVaR由Rockafellar率先提出,CVaR是组合损失超过VaR的条件均值,反映了超额损失的平均水平。本文拟将CVaR代替VaR来计算RAROC的值作为基金的投资绩效指标,所以CVaR的准确计算成为本文的关键。

二、全国社保基金投资组合系统性风险的度量

系统性风险指全局性的共同因素引起的投资收益的可能变动,这种共同因素以同样的方式对所有有价证券产生影响。系统性风险主要包括战争风险、政策风险、经济周期性波动风险、利率风险、购买力风险、汇率风险等。这种风险即使投资分散化足够充分也不能分散,所以又称为不可分散风险。

据《2012全国社保基金理事会基金年度报告》统计,截至2012年年末,全国社会保障基金数额达到11 082.75亿元,累计投资收益额3 492.45亿元,年均投资收益率达8.29%。全国社保基金的投资工具非常广泛,总的来说分为境内投资和境外投资。考虑到在计算贝塔系数时要用到市场收益类指标,而从全球范围来说难以找到这样的指标,因此本文只考虑国内情况。境内投资可以分为四大类投资工具,即固定收益资产投资、股票投资、企业债投资和基金投资。本文选择上证综合指数和深圳成分股指数的加权值作为市场收益。利用前面介绍的CAPM公式就可以推算出系统性风险占总风险的比例大小。据统计,我国市场收益系统性风险占总风险的比例是66.67%。

三、全国社保基金投资组合CVaR值的准确计算

在CVaR的众多计算方法中,蒙特卡罗方法是最为有效的方法,此方法对于资产组合的不同分布状况及各种非线性的情形都可以得到令人满意的结果。进行蒙特卡罗模拟之前的关键之一是要知道资产组合的概率密度分布状况及各资产之间的相关结构。高维数据建模可以采

用Pair - Copula(也称藤Copula或Vine Copula)的方法。该方法将N元密度函数分解成N(N - 1)/2个Pair - Copula模块和边缘密度函数的乘积,并且允许每一个Pair - Copula模块选择合适的Copula类型进行拟合。

常用藤结构有两种情形,即C藤和D藤。选择C藤或者D藤的一个常用原则是:当数据中有一个变量是引导其他变量的关键变量时,选用C藤分解,而当各变量相对独立时,选用D藤分解。对于C藤和D藤经Pair - Copula分解后的联合密度函数如下所示:

$$\text{C藤:} f(x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} f(x_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j,j+i|1,\ldots,j-1}$$
$$\{F(x_j|x_1, \ldots, x_{j-1}), F(x_{j+i}|x_1, \ldots x_{j-1})\}$$

$$\text{D藤:} f(x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} f(x_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{i,i+j|i+1,\ldots,i+j-1}$$
$$\{F(x_i|x_{i+1}, \ldots, x_{i+j-1}), F(x_{j+i}|x_{i+1}, \ldots x_{i+j-1})\}$$

选择好适当的藤结构后,就是各Pair - Copula函数族的选择和参数估计问题了。Brechmann(2010)指出,针对二元Copula,赤池信息准则(AIC)是一个可靠的选择标准,故本文将以AIC作为各Pair - Copula函数的选择标准。对于各Pair - Copula函数的参数估计,Aasetal(2009)发展了标准的极大似然估计方法:首先单独地对每个Pair - Copula采用极大似然法(ML)估计出其参数作为初值,将估计出的参数初值代入对数极大似然函数得到参数估计的终值,一般参数初值与终值差别不大。n维C藤和D藤的对数极大似然函数如下所示:

$$\text{C藤:} L(\theta) = \sum_{j=1}^{n-1} \sum_{i=1}^{n-j} \sum_{t=1}^{T} \log[c_{j,i+j|1,\ldots,j-1}\{F(x_{j,t}|x_{1,t}, \ldots, x_{j-1,t}),$$
$$F(x_{i+j}|x_{1,t}, \ldots, x_{j-1,t}), \theta\}]$$

$$\text{D藤:} L(\theta) = \sum_{j=1}^{n-1} \sum_{i=1}^{n-j} \sum_{t=1}^{T} \log[c_{i,i+j|i+1,\ldots,i+j-1}\{F(x_i|x_{i+1}, \ldots, x_{i+j-1}),$$
$$F(x_{i+j}|x_{i+1}, \ldots, x_{i+j-1}), \theta\}]$$

参数估计完毕后,通过Monte Carlo模拟仿真技术生成服从多元联合分布的仿真序列,具体过程如下:步骤一:生成n个服从[0,1]均匀分布且相互独立的随机数$\{\omega_i, i=1,2,\ldots,n\}$;步骤二:逐步算出服从Pair - Copula分解模块的仿真序列($v_1, v_2, \ldots, v_n$),$v_1 = \omega_1$,$v_2 = F^{-1}_{2|1}(\omega_2|v_1)$,$v_3 = F^{-1}_{3|1,2}(\omega_3|v_1, v_2)$,$\cdots$,$v_n = F^{-1}_{n|1,2,\ldots,n-1}(\omega_n|v_1, v_2, \ldots, v_{n-1})$,其中每个条件分布函数可由下面的公式得到:$F(x|v) = \partial C_{x,v_j|v_{-j}}(F(x|v_{-j}), F(v_j|v_{-j}))/\partial F(v_j|v_{-j})$;步骤三:重复步骤二2 000次,结合边缘分布得到$n \times 2\ 000$的收益率矩阵Y。在一定的置信度下就可以求出投资组合的CVaR值,以CVaR代替VaR最终算出资产组合的投资绩效。

四、实证分析

《全国社会保障基金投资管理暂行办法》等法律法规对全国社会保障基金的投资品种和投资比例均做出了明确规定:国内投资方面,要求存款、国债及金融债投资比例不低于40%(一般是50%),企业债投资比例不高于10%,股票及基金投资比例不超过40%。

本文选择上证180指数代表股票投资、上证基金指数代表基金投资、上证国债指数代表固定收益投资、上证企债指数代表企业债投资、上证综合指数的60%和深圳成分指数的40%作为我国市场收益。样本数据时间区间为2005年1月4日至2013年12月25日共2 178个交易数据。数据来源于新浪财经,数据处理采用Eviews、R及matlab软件。

将指数的每日收盘价作为价格,按公式$x_t = \ln(p_t/p_{t-1})$得到每日收益率序列,收益率的基本统计特征见表1。

表1　　样本收益率的基本统计特性

	上证国债	上证基金	上证180	上证企债
Mean	0.000 173	0.000 690	0.000 350	0.000 254
Median	0.000 121	0.000 156	0.000 972	0.000 210
Maximun	0.007 456	0.094 098	0.089 486	0.016 722
Minimun	-0.005 416	-0.093 435	-0.097 525	-0.009 900
Std.Dev	0.000 856	0.016 728	0.018 620	0.001 365
Skewness	0.611 292	0.100 592	-0.306 436	1.432 712
Kurtosis	15.215 13	7.176 685	5.868 086	24.783 45
J-B	13 670.12	1 586.051	780.230 2	43 787.57
P	0.000 000	0.000 000	0.000 000	0.000 000
ADF(1%显著水平临界值)	-36.294 22 (-3.433 158)	-45.556 62 (-3.433 158)	-46.298 61 (-3.433 158)	-10.079 33 (-3.433 158)

由表1可知,四个收益率序列都不同程度存在高峰后尾现象,且J~B统计值及其概率拒绝收益率服从正太分布的假设。ADF值表明四个收益率序列不存在单位根,是平稳序列,可进行统计学建模。

通过对各序列进行ARCH - LM检验,发现均具有明显的ARCH效应,分别对四个收益率序列进行GARCH(1,1)模型建模,其模型如下:

$$x_t = \mu_t + \varepsilon_t$$
$$\varepsilon t = \sqrt{h_t v_t}$$
$$h_t = w_0 + \alpha_1 \varepsilon^2_{t-1} + \beta_1 h_{t-1}$$

其中:v_t服从标准正态分布或者t分布或者广义误差分布等。由于上面收益序列呈现出明显的尖峰后尾现象,所以拟假设残差服从t分布。表2为各收益率序列GARCH(1,1)-t模型的参数估计。

K - S值及其相伴概率表明,经概率积分转换后的残差序列服从[0,1]均匀分布,且残差序列与t分布的Q - Q图表明假设残差序列服从标准化t分布是恰当的,说明GARCH(1,1)-t能很好地拟合上述数据。

基于Kendall'S tau及C藤和D藤的选择标准,选择合适的Pair - Copula分解类型。经GARCH - t(1,1)过滤后的两两残差间的Kendall'S tau相关系数见表3。

表2　GARCH-t(1,1)模型参数估计结果

GARCH-t	上证国债	上证企债	上证180	上证基金
u	0.000 113	0.000 163	0.000 516	0.000 124
AR(1)	0.078 721	0.188 638	-	-
w_0	3.64E-09	3.69E-09	1.79E-06	2.03E-06
α	0.302 225	0.129 704	0.040 462	0.055 232
β	0.874 531	0.882 650	0.955 750	0.941 211
自由度	2.339 577	5.004 873	5.406 751	4.360 351
K-S统计量(P)	0.021 8 (0.247 0)	0.015 5 (0.663 8)	0.016 9 (0.560 4)	0.016 7 (0.574 4)

表3　各残差序列间的Kendall'S tau

	上证国债	上证基金	上证180	上证企债
上证国债	1.000 0	-0.012 4	-0.023 3	0.189 2
上证基金	-0.012 4	1.000 0	0.777 1	-0.033 1
上证180	-0.023 3	0.777 1	1.000 0	-0.033 9
上证企债	0.189 2	-0.033 1	-0.033 9	1.000 0

由表3可知,除了上证基金与上证180以及上证企债与上证国债的相关性较强外,其他的相关性都较弱。四个指数间没有先导变量的存在,所以选择D藤分解。根据相关性强弱,第一棵树的排列为:上证基金→上证180→上证企债→上证国债。为表述方便,从左到右编号分别为1、2、3、4。利用R软件CDVine包里的相关函数,首先选择适当的Copula函数拟合每棵树上的数据,接着求出每个Pair-Copula函数的参数初值,最后最大化对数似然函数求出各Pair-Copula函数的终值。结果见表4。

表4　各Pair-Copula的类型、参数估计、对数似然值、AIC及BIC的值

树(j)	参数	Copula	初　值	终　值
1	ρ_{12}	t	0.945 1	0.945 5
	ν_{12}		2.000 1	2.000 1
	ρ_{23}	t	-0.052 96	-0.048 50
	ν_{23}		5.912 8	5.766 1
	ρ_{34}	t	0.305 9	0.314 5
	ν_{34}		3.699 6	3.861 7
2	$\rho_{13\|2}$	t	0.017 49	0.020 19
	$\nu_{13\|2}$		5.189 8	5.151 8
	$\theta_{24\|3}$	rotated Clayton	-0.037 23	-0.038 43
3	$\rho_{14\|23}$	t	0.001 175 6	0.001 071
	$\nu_{14\|23}$		11.100 3	11.099 9
Log.likelihood			2 689.334	2 689.511
AIC			-5 356.668	-5 357.021
BIC			-5 294.131	-5 294.483

参数估计完后,利用蒙特卡罗模拟法模拟出一系列伪随机数代入边缘分布模型得到2 000组t+1时刻收益率的模拟值。

根据给定的投资权重算出不同置信度下的CVaR、RAROC、贝塔系数及系统性风险的值,结果见表5。

表5　不同置信水平的CVaR、RAROC、Beta系数、系统性风险比例

	上证基金	上证180	上证企债	上证国债	CVaR	RAROC	Beta	系统性风险
99%	30%	10%	10%	50%	0.019 5	0.018 1	0.356 3	0.888 6
95%					0.012 6	0.028 1		
90%					0.009 9	0.035 7		

由表5可以看出,单从风险角度来说,全国社保基金投资组合的系统性风险占总风险的比例高达88.86%,比市场收益系统性风险比例(66.67%)高20%多,说明社保基金投资组合经过调整还有分散非系统性风险的余地。且从CVaR的值可以看出处于风险资本的平均值有99%的把握不会超过总资本的1.95%。从收益和风险的综合角度来说,由于全国社保基金是作为未来养老支付压力的战略储备基金,追求收益势在必行,因此每损失一单位风险资本有99%的把握可以得到1.81%的风险收益还是可取的。

五、结论

本文通过计算贝塔系数得出全国社保基金投资组合系统性风险的大小,并且与市场收益系统性风险做对比,说明社保基金投资组合经过调整还有分散非系统性风险的潜力。

针对VaR的一些不足,在传统求解RAROC的基础上以CVaR代替VaR,这样能更好地说明风险与收益之间的关系。在求解CVaR时,利用Pair-Copula分解拟合投资组合的联合分布,并且通过蒙特卡罗模拟法模拟出服从多元联合分布的仿真序列能更好地反映未来收益情况,算出的CVaR值更准确。

【注】本文系国家自然科学基金资助项目"时序非线性关联Copula理论建模及在金融领域的应用研究"(编号:71071111)的阶段性研究成果。

主要参考文献

1. Sharp William. Capital asset prices: a theory of market equilibrium under condition of risk. Journal of Finance, 1964;12

2. Sharp W. F.. Mutual Fund Performance. Journal of Business, 1966;39

3. 菲利普·乔瑞著.郑伏虎,万峰,杨瑞琪译.风险价值VaR.北京:中信出版社,2010

4. 杜子平等.基于"藤"结构的高维动态Copula的构建.教学的实践与认识,2009;10

5. Artzner P., Delbaen F., Eber J. M. et al.. Coherent measures of risk. Mathematical finance, 1999;9