第53卷第5期
2013年9月

大连理工大学学报
Journal of Dalian University of Technology

Vol. 53, No. 5
Sept. 2013

应用数学

文章编号：1000-8608(2013)05-0772-05

基于 f-散度的一致风险度量

任凤英[*1,2]， 李兴斯[1]

（1.大连理工大学 运筹学与控制论研究所，辽宁 大连 116024；
2.北京师范大学珠海分校 应用数学学院，广东 珠海 519087）

摘要： 进一步研究了带有熵约束的一致风险度量，鉴于它的许多良好性质，提出了带有 f-散度约束的一致风险度量，获得了一大组可以灵活选取的风险度量，并且给出了一致风险度量理论中可接受集的控制方法.另外，讨论了将这类风险度量直接用于优化问题的优点以及它们的稳健化处理方法，得到了一致风险度量的新的表现形式，此形式是 $CVaR$ 的上界，在应用中是更易处理的.

关键词： 一致风险度量；凸风险度量；一致熵风险度量；一致 f-散度风险度量

中图分类号： O221.5 **文献标志码：** A

0 引 言

根据一致风险度量理论，对于一组风险资产构成的组合 X，决策者能够知道需要多少资本储备，使得组合 X 的风险可以接受.其基本思想是首先定义一个可接受集合 A，需要 A 满足一些合理的性质，这组性质形成风险度量的一致性公理，任何一个满足一致性公理的风险度量都可以表示为

$$\rho(X)=\sup_{Q} E_Q(-X)$$

这个表达式也称为一致风险度量的对偶表达式，其中 $Q\in\Sigma$，Σ 代表概率测度的集合，它的大小本质上是由可接受集合 A 决定的.因为要使风险 X 变得可以接受，那么所需的最小资本储备是 $\rho(X)=\inf_{m}\{m+X\in A\}$.集合 A 的大小决定了概率测度 $Q\in\Sigma$ 的变化范围.如果知道如何说明可接受集 A，就能构造相应的风险度量 $\rho_A(X)$；反之，如果有一个一致风险度量 $\rho(X)$，就能表示出可接受集合 $A_\rho=\{X\mid\rho(X)\leqslant 0\}$.在文献[1]中，通过锥优化的对偶关系，说明了 A 与 Σ 的相互制约关系.由于定量地表达 A 是非常困难的，对于 $CVaR$ 来说，可接受集合对应的概率测度集是 $\Sigma=\left\{Q\in M_1\mid Q\ll P,\frac{\mathrm{d}Q}{\mathrm{d}P}\leqslant\frac{1}{\alpha}\right\}$，其中 M_1 是所有概率测度的集合.因此，有 $\log\frac{\mathrm{d}Q}{\mathrm{d}P}\leqslant c$ 和 $H(Q\mid P)\leqslant c$，$c=-\log\alpha$.受这一表达形式的启发，文献[1]提出在上面的风险度量的对偶表达式中，用相对熵（Kullback-Leibler 散度）来控制 Σ 的大小，产生一个优化模型如下：

$$\begin{aligned}&\sup_{Q} E_Q(-X)\\ &\text{s.t.}\quad E_Q\left(\log\frac{\mathrm{d}Q}{\mathrm{d}P}\right)\leqslant c\\ &\qquad\int q(x)\,\mathrm{d}x=1\\ &\qquad q(x)\geqslant 0\end{aligned}$$

其中 c 由可接受的风险水平确定，通过控制参数 c 的变化来约束概率测度的集合，并对应不同的可接受集 A_c，表示不同的代理人有不同的容忍度，同时也表示一组不同的风险度量 $\rho_c(X)$.$H(Q\mid P)$ 是 Q 关于 P 的相对熵，定义如下：

$$H(Q\mid P)=\begin{cases}E_Q\left(\log\frac{\mathrm{d}Q}{\mathrm{d}P}\right); & Q\ll P\\ +\infty; & \text{其他}\end{cases}$$

最近，Follmer 等[2] 从分布不变性、模型稳健

收稿日期： 2012-03-16； **修回日期：** 2013-07-01.
作者简介： 任凤英*(1961-)，女，博士生，副教授，E-mail：rui5926@sina.com.

性、独立风险聚合下的行为以及大偏差界等几个方面研究了一致风险度量的这种特殊形式，并称之为一致熵风险度量，记为

$$\rho_c(X) := \sup_{Q:H(Q|P)\leqslant c} E_Q(-X)$$

文献[2]是从凸熵风险度量的一致调整的角度来研究 $\rho_c(X)$ 的性质的. 凸熵风险度量的定义如下：

$$e_\gamma(X) := \frac{1}{\gamma}\log E_P(\mathrm{e}^{-\gamma X}) = \sup_Q\left\{E_Q(-X) - \frac{1}{\gamma}H(Q \mid P)\right\}$$

其中参数 $\gamma \in [0, +\infty)$，是风险厌恶系数. 文献[2]分析了 $\rho_c(X)$ 的一些性质，更多地侧重于这两个风险度量之间的差别. 在精算问题中应用这两个风险度量时，$\rho_c(X)$ 比 $e_\gamma(X)$ 的表现更好. 本文要在一致风险度量的统一表达式中，通过描述概率测度集合 Σ，间接地表示各种不同的可接受集合 A. 除了相对熵外，也扩展到一般的 f- 散度，称之为一致 f- 散度风险度量. 由于有一大组 f- 散度，产生了一组丰富的一致 f- 散度风险度量，并讨论它们的性质和应用问题.

1　记号与定义

设 (Ω, F) 为可测空间，X 是定义在其上的有界可测随机变量的线性空间. 让 A 表示可接受凸锥，$X^+ \subset A$ 且 $A \cap (-A) = \{0\}$. 那么泛函 $\rho(\cdot)$ 定义如下：

$$\rho(X) := \inf\{m \in \mathbf{R} \mid X + m \in A\}$$

可接受集合 A 的性质导致 $\rho(X)$ 有如下 4 条性质，称为一致性公理，即对于 $X, Y \in L^\infty$ 和 $m \in \mathbf{R}$，$\lambda \in \mathbf{R}^+$，有

(1) 单调性：$X \geqslant Y \Rightarrow \rho(X) \leqslant \rho(Y)$

(2) 平移不变性：$\rho(X+m) = \rho(X) - m$

(3) 正齐性：$\rho(\lambda X) = \lambda\rho(X)$

(4) 次可加性：$\rho(X+Y) \leqslant \rho(X) + \rho(Y)$

满足这些性质的风险度量称为一致风险度量，它们有统一的对偶表达式：

$$\rho(X) = \sup_{Q\in\Sigma} E_Q(-X)$$

当不要求有正齐性时，一致风险度量扩展为凸风险度量，它们的统一的对偶表达式为

$$\rho(X) = \sup_{Q\in\Sigma_1}\{E_Q(-X) - \alpha(Q)\}$$

其中 Σ_1 表示在 X 上的所有概率测度，$\alpha(Q)$ 定义如下：

$$\alpha(Q) := \sup_{X\in A_\rho} E_Q(-X)$$

当 $\alpha(Q) \in \{0, +\infty\}$ 时，凸风险度量的表达式化为一致风险度量的表达式[3-6].

2　带约束的一致风险度量

2.1　一致熵风险度量

从引言的两个表达式可知，一致熵风险度量是带有 Kullback-Leibler 散度约束的风险度量；凸熵风险度量是带有熵罚的凸风险度量.

定义 1　对给定的 $c > 0$，一致熵风险度量 ρ_c 定义如下：

$$\rho_c(X) := \sup_{Q:H(Q|P)\leqslant c} E_Q(-X);\ X \in L^\infty$$

其中 $Q \in \Sigma$，Σ 是一致风险度量定义中所有概率测度的集合. c 的选取是为了控制概率测度集合 Σ 的大小.

$\rho_c(X)$ 的基本性质：

(1) 对任何的 $\alpha \in (0,1)$ 和 $X \in L^\infty$，

$$VaR_\alpha(X) \leqslant CVaR_\alpha \leqslant \rho_{c(\alpha)}(X)$$

其中 $c(\alpha) = -\log\alpha > 0$.

左边的不等式是已知的[2]，右边的不等式证明如下：

因为 $CVaR_\alpha(-X) = \sup_{Q\in\Sigma_1} E(-X)$，$\Sigma_1 = \left\{Q \in M_1 \mid Q \ll P, \frac{\mathrm{d}Q}{\mathrm{d}P} \leqslant \frac{1}{\alpha}\right\}$，由 $\frac{\mathrm{d}Q}{\mathrm{d}P} \leqslant \frac{1}{\alpha}$，两边取对数得 $\log\frac{\mathrm{d}Q}{\mathrm{d}P} \leqslant -\log\alpha = c(\alpha)$，再两边取期望得 $E_Q\left(\log\frac{\mathrm{d}Q}{\mathrm{d}P}\right) \leqslant c(\alpha)$，$\Sigma_2 = \left\{Q \in M_1 \mid Q \ll P, E_Q\left(\log\frac{\mathrm{d}Q}{\mathrm{d}P}\right) \leqslant c(\alpha)\right\}$，反过来一般是不成立的. 因此，$CVaR_\alpha(X) = \sup_{Q\in\Sigma_1}(-X) \leqslant \sup_{Q\in\Sigma_2}(-X) = \rho_{c(\alpha)}(-X) \leqslant \rho_c(-X)$. 其中 c 是比 $c(\alpha)$ 更大的数. 可见 $\rho_{c(\alpha)}(-X)$ 和 $\rho_c(-X)$ 是依次比 $CVaR_\alpha(-X)$ 更保守的风险度量.

(2) ρ_c 是分布不变的，上连续且下连续；ρ_c 关于 c 是递增的且满足下列极限：

$$\lim_{c\downarrow 0}\rho_c(X) = E_P(-X),\ \lim_{c\uparrow +\infty}\rho_c(X) = \operatorname{esssup}(-X)$$

(3) 根据 ρ_c 的定义,其中的最大值是可达的,即

$$\rho_c(X)=\max_{Q\in\Sigma,H(Q|P)\leqslant c}E_Q(-X)=E_{\bar{Q}}(-X)$$

(4) 当考虑 n 个不同的随机变量时,ρ_c 有下面的性质:

$$\pi_{c,n}=\frac{1}{n}\rho_c(X_1+\cdots+X_n),\ \lim_{n\uparrow+\infty}\pi_{c,n}=E(-X_1)$$

这个性质说明:如果这 n 个随机变量表示 n 个保险合约,当合约数量增加时,每个合约的费用将下降到公平费用. 性质(2)~(4)的证明见文献[2].

2.2 一致 f-散度风险度量

定义 2 设概率测度 Q 关于 P 连续,$f:(0,+\infty)\to\mathbf{R}$ 是任意一个凸函数,使得 $f(1)=0$,那么 Q 和 P 之间的 f- 散度是

$$D_f(Q\mid P):=\begin{cases}\int f\left(\dfrac{\mathrm{d}Q}{\mathrm{d}P}\right)\mathrm{d}P; & \text{积分存在}\\ +\infty; & \text{否则}\end{cases}$$

其中 $f(0)=\lim\limits_{x\downarrow 0}f(x)$,$0f(0/0)=0$,和 $0f(a/0)=a\lim\limits_{u\to+\infty}f(u)/u$.

在离散的情况下,$D_f(q\mid p)=\sum\limits_{i=1}^{n}p_if(\frac{q_i}{p_i})$,分布的熵定义为 $H(q)=\sum\limits_{i=1}^{n}(-p_i\ln p_i)$.

由于凸函数 f 可有多种选取方式,f- 散度包含了一大组关于概率分布间距离的度量. 另外,也会用到 f 的共轭函数 f^*,定义如下:

$$f^*:\mathbf{R}\to\mathbf{R}\cup\{+\infty\},\ f^*(s)=\sup_{t\geqslant 0}\{st-f(t)\}$$

下面是常用的凸函数及其代表的散度或称距离.

例 1 $f(x)=x\log x$, 对应 Kullback-Leibler 相对熵散度

$$f(Q\mid P)=\int\frac{q(x)}{p(x)}\ln\frac{q(x)}{p(x)}\mathrm{d}P=\int q(x)\ln\frac{q(x)}{p(x)}\mathrm{d}x$$

例 2 $f(x)=|x-1|$(全变差散度)

例 3 $f(x)=(\sqrt{x}-1)^2$ (Hellinger 散度)

例 4 $f(x)=(x-1)^2$(χ^2 散度)

例 5 $f(x)=\ln(2/(1+u))+u(2u/(1+u))$(Jensen-Shannon 散度).

分布的熵与 f- 散度有密切的联系,如下面的式子:

$$JS(q\mid p)=\ln 2\left(2H\left(\frac{p+q}{2}\right)-H(q)-H(p)\right)$$

$$KL(q\mid u)=\ln 2(\log n-H(q))$$

容易看到,当其中的 u 代表离散的均匀分布时,每种状态的概率是 $1/n$. 当控制 $KL(q\mid u)\leqslant c$ 相当于控制 $H(q)\geqslant\log n-\frac{c}{\ln 2}$,当不知道未来的概率分布时,一般假设每种状态出现的概率为 $1/n$,这时熵最大是 $\log n$,因此,$\log n-\frac{c}{\ln 2}\leqslant H(q)\leqslant\log n$,这一约束可解释为从等权赋概到概率权重的调整,目标函数将导致大的权重移向左尾部.

已知的其他的散度还有很多,f- 散度的一些基本性质如下:

性质 1 $D_f(Q\mid P)=D_{f^*}(P\mid Q)$ 当且仅当 $\exists c\in\mathbf{R}$:$f^*(u)-f(u)=c(u-1)$,f^* 是 f 的伴随共轭函数.

性质 2 $D_{f_1}(Q\mid P)=D_f(Q\mid P)$,当且仅当 $\exists c\in\mathbf{R}$:$f_1(u)-f(u)=c(u-1)$.

性质 3 $f(1)\leqslant D_f(Q\mid P)\leqslant f(0)+f^*(0)$.

性质 4 $D_f(Q\mid P)\geqslant f(1)=0$,其中等式成立当且仅当 $Q=P$.

Q、P 是任意的概率测度.

定义 3 对给定的 $c>0$,一致 f- 散度风险度量 $\rho_{(f,c)}$ 定义如下:

$$\rho_{(f,c)}(X):=\sup_{D_f(Q|P)\leqslant c}E_Q(-X);\ X\in L^\infty$$

其中 $Q\in\Sigma$,Σ 是一致风险度量定义中所有概率测度的集合. 显然一致熵风险度量是它的特例,将从优化的角度分析一致 f- 散度风险度量. 它们的性质概括在下面的定理中.

定理 1 $\rho_{(f,c)}(X)$ 是一致的且它的最优值是可达的,$\rho_{(f,c)}(-X)=E_{\bar{Q}}(-X)$,它有一个自然的上界 $\lambda c+\int(\lambda f)^*(-x)\mathrm{d}p(x)$,它的对偶表达式是 $\bar{\lambda}c+E_P[(-X)\bar{\lambda}f^*]$.

证明 记 $\frac{\mathrm{d}Q}{\mathrm{d}P}=\frac{q(x)}{p(x)}=\phi(x)$,那么一致 f- 散度风险度量可以表示为如下的凸优化问题:

$$
\begin{aligned}
(\mathrm{P})\quad & \max_{\phi}\int(-x)\phi(x)\mathrm{d}p(x)\\
\text{s.t.}\quad & \int f(\phi(x))\mathrm{d}p(x)\leqslant c\\
& \int\phi(x)\mathrm{d}p(x)=1\\
& \phi(x)\geqslant 0
\end{aligned}
$$

由于目标函数的线性和 f 的凸性，这是一个凸优化问题，使用 Lagrange 方法求解.

$$
\begin{aligned}
L(\lambda,\mu)=&\int(-x)\phi(x)\mathrm{d}p(x)-\\
&\lambda\left(\int f[\phi(x)]\mathrm{d}p(x)-c\right)+\\
&\mu\left(\int\phi(x)\mathrm{d}p(x)-1\right)
\end{aligned}
$$

优化问题中的密度函数之比 $\phi(x)\geqslant 0$，联系不等式的 Lagrange 乘子 $\lambda\geqslant 0$. 根据最优解的充分且必要条件，对所有的 $g(x)$，计算 L 关于 $g(x)$ 的方向导数：

$$
\begin{aligned}
D_gL(\lambda,\mu)=&\int(-x)\cdot g(x)\mathrm{d}p(x)-\\
&\lambda\left[\int f'[\phi(x)]\cdot g(x)\mathrm{d}p(x)\right]+\\
&\mu\int g(x)\mathrm{d}p(x)=0
\end{aligned}
$$

因此，$\phi(x)$ 必须使下面的等式

$$(-x)-\lambda f'(\phi(x))+\mu=0$$

几乎处处成立.

$\phi(x)$ 还需要满足下列约束条件和互补松弛条件：

$$\int f[\phi(x)]\mathrm{d}p(x)\leqslant c$$

$$\int\phi(x)\mathrm{d}p(x)=1$$

$$\lambda\left(\int f[\phi(x)]\mathrm{d}p(x)-c\right)=0$$

可以断言 $\lambda>0$，否则，由最优性条件可推出 $x=\mu$. 这是退化的情况. 当 $\lambda>0$，

$$\phi(x)=(f')^{-1}\left(\frac{\mu-x}{\lambda}\right)$$

当 $f(x)=x\log x$ 时，$\phi(x)=\mathrm{e}^{(-x)/\lambda+\mu/\lambda-1}$，最优值是可达的，记为 $E_{\bar{Q}}(-X)$.

如果在优化问题中，默认 $\phi(x)$ 是密度之比，则拉格朗日函数简化为

$$
\begin{aligned}
L(\lambda)=&\int(-x)\phi(x)\mathrm{d}p(x)-\\
&\lambda\left(\int f[\phi(x)]\mathrm{d}p(x)-c\right)
\end{aligned}
$$

$$
\begin{aligned}
\max_{q(x)}L(q(x),\lambda)=&\lambda c+\max_{q(x)}\left\{\int(-x)\phi(x)\mathrm{d}p(x)-\right.\\
&\left.\lambda\int f(\phi(x))\mathrm{d}p(x)\right\}=\\
&\lambda c+\max_{q(x)}\int[(-x)\phi(x)-\\
&\lambda f(\phi(x))]\mathrm{d}p(x)=\\
&\lambda c+\int(\lambda f)^*(-x)\mathrm{d}p(x)
\end{aligned}
$$

因此，得到 $\rho_{(f,c)}$ 的上界和对偶表达式：

$$\rho_{(f,c)}(X)\leqslant\max_{q(x)}L(q(x),\lambda)$$

$$
\begin{aligned}
\rho_{(f,c)}(X)=&E_{\bar{Q}}(-X)=\\
&\min_{\lambda}(\lambda c+E_P[(-X)(\lambda f)^*])=\\
&\bar{\lambda}c+E_P[(-X)(\bar{\lambda}f)^*]
\end{aligned}
$$

3　应　用

3.1　对冲问题中的应用

可以将一致 f- 散度风险度量融于对冲问题中，不必单独增加一个风险度量约束. 一个风险最小化欧式期权对冲问题如下：

$$
\begin{aligned}
&\min_{(\alpha,\beta)}\rho_{(f,c)}(X_T-C_T)\\
\text{s.t.}\quad & \int f(\phi(x))\mathrm{d}p(x)\leqslant c\\
& \int\phi(x)\mathrm{d}p(x)=1\\
& \phi(x)\geqslant 0
\end{aligned}
$$

其中 X_T 代表对冲组合在到期日的值，它由 α 份股票和 β 份无风险债券构成；C_T 是欧式期权在到期日的价值，暂且考虑单期的对冲问题. 那么这个问题就是在可接受的风险容忍水平下，选取风险资产和无风险资产的投资组合，使得一致 f- 散度风险度量最小. 所求得的最优值说明：还需要多少资本储备才能使得出售期权可能发生的负债在可接受的风险容忍水平之下. 由于这个优化问题是一个凸优化，可以将其离散化后求解，它是在多项式时间内可解的[7].

3.2　稳健化一致 f-散度风险度量

上述风险度量假设 P 是已知的，但是在实际

的应用中，P 通常是从历史数据中估计或使用蒙特卡罗方法获得，有时也采用专家的判断，因此所得概率测度 P 不可避免地含有误差. 为了控制模型风险，提出一致 f- 散度风险度量的稳健形式：

$$\rho_{(f,c)}(X)=\sup_{Q\in D_{(f,\lambda^*)}(Q|P)\leqslant c}E_Q(-X)$$

其中 $P_\lambda=\lambda P_1+(1-\lambda)P_2$，$P_1$ 表示来自历史数据的估计；P_2 表示蒙特卡罗模拟的结果；P_λ 是专家通过赋权方式给予的判断，调整 λ 可以表示专家的看法，保守观点是 $D_{(f,\lambda^*)}(Q\mid P_\lambda)=\min\limits_{\lambda}D_f(Q\mid P_\lambda)$. 这需要解两个优化问题，恰好它们都是凸优化，理论上是可解的.

4 结 语

本文探索了如何给出一致风险度量中可接受集的可操作性表示，即把它转化为对 f- 散度的约束，通过上界 c 来表示概率测度间的距离，其中 c 的选取可参考 $CVaR_\alpha$ 中的 α. 通过 c 的变化，所给出的一致 f- 散度风险度量能够表达决策者的风险厌恶，并且有多种表示方式. 这一方法还可以直接用于组合优化和衍生产品的对冲问题，这类优化问题的求解留待以后详细探讨.

参考文献：

[1] 任凤英，李兴斯. 一致风险度量和锥优化分析[J]. 运筹学学报，2010，**14**(2)：95-106.

REN Feng-ying，LI Xing-si. The cone optimization analysis of coherent risk measure [J]. **Operations Research Transactions**，2010，**14**(2)：95-106. (in Chinese)

[2] Follmer H，Knispel T. Entropic risk measures：coherence vs. convexity model ambiguity and robust large deviations [J]. **Stochastics and Dynamics**，2011，**11**(2-3)：333-351.

[3] Artzner P，Delbaen F，Eber J M，*et al*. Coherent measures of risk [J]. **Mathematical Finance**，1999，**9**(3)：203-228.

[4] Föllmer H，Schied A. **Stochastic Finance — An Introduction in Discrete Time**[M]. 2nd ed. Berlin：Walter de Gruyter，2004.

[5] Föllmer H，Schied A. Convex measures of risk and trading constraints [J]. **Finance and Stochastics**，2002，**6**(4)：429-447.

[6] Frittelli M，Rosazza G E. Putting order in risk measures [J]. **Journal of Banking & Finance**，2002，**26**(7)：1473-1486.

[7] Ben-Tal A，Nemirovski A. **Lectures on Modern Convex Optimization：Analysis Algorithms and Engineering Applications** [M]. Philadelphia：SIAM，2001.

Coherent risk measure based on f-divergence

REN Feng-ying*[1,2]， LI Xing-si[1]

(1. Institute of Operations Research and Control，Dalian University of Technology，Dalian 116024，China；
2. College of Applied Mathematics，Beijing Normal University，Zhuhai，Zhuhai 519087，China)

Abstract：The investigation for the coherent risk measures with entropy constraints is conducted，and because it has many good properties，the coherent risk measure with f-divergence constraints is presented. The study obtains a group of flexible risk measures and the methods to control the acceptable set of the coherent risk measure theory. In addition，its advantages used directly in optimization problems and the robustness of method are discussed. The new expression of coherent risk measure is obtained. It is the upper bound of $CVaR$ and more tractable in applications.

Key words：coherent risk measure；convex risk measure；coherent entropy risk measure；coherent f-divergence risk measure