

最大熵分位值估计及其对偶型优化解法

吴福仙, 温卫东

(南京航空航天大学 能源与动力学院, 南京 210016)

摘 要: 针对经典最大熵分位值估计中拉格朗日系数计算目前存在高度非线性、计算结果精度不高或有时难以收敛等问题,提出了一种对偶型-逐次寻优的方法。基于拉格朗日对偶法,推导建立了含有拉格朗日系数优化函数的对偶表达式;在此基础上,基于样本的概率权重矩约束,提出了逐次寻优算法。针对几种常见的概率分布类型和一种较为复杂的概率分布类型,采用对偶型最大熵方法和经典最大熵方法对其概率累积函数和分位值进行计算对比分析表明:对偶型最大熵分位值估计不仅具有非线性程度低、形式简单,而且对偶型-逐次寻优的方法具有比较高的计算精度,优化迭代的收敛性好等特点。

关 键 词: 分位值估计; 对偶型; 最大熵; 逐次优化; 概率权重矩

中图分类号: V232.4

文献标志码: A

Maximum entropy quantile estimation and dual optimization method

WU Fu-xian, WEN Wei-dong

(College of Energy and Power Engineering,

Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: Based on high nonlinearity, low computational accuracy, or hard convergence of Lagrangian calculation in the classic maximum entropy quantile function, a new method called dual-sequential updating method was proposed. Dual expression of optimization function was established on the basis of Lagrangian dual approach. Further sequential updating method was proposed based on the constraint by the probability weighted moments. After analyzing several common distributions and a complex distribution by dual maximum entropy quantile function and classic maximum entropy quantile function, result showed that, dual maximum entropy quantile function had advantage of low nonlinearity and simple form, and dual-sequential updating method had good computational accuracy and optimization convergence.

Key words: quantile estimation; dual; maximum entropy; sequential updating; probability weighted moments

可靠性工程设计中经常要求出给定可靠度对应的分位值,一般事先必须假定概率分布类型,然后进行假设检验,参数估计,再得到分位值函数。而最大熵原理在解决概率性问题时能够不事先假定概率分布而拟合出分位值函数,是处理这

类问题的一种有力工具。

文献[1-6]提出了经典最大熵分位值估计的推导过程,利用非线性最小二乘法-拟牛顿优化法确定出最大熵分位值函数的拉格朗日系数。但由于经典最大熵分位值估计拉格朗日系数计算的高

收稿日期: 2015-05-22

作者简介: 吴福仙(1985—),男,博士生,主要从事航空发动机结构强度振动可靠性研究。E-mail:johnny.2009@163.com

引用格式: 吴福仙,温卫东. 最大熵分位值估计及其对偶型优化解法[J]. 航空动力学报, 2017, 32(2): 416-423. WU Fuxian, WEN Weidong. Maximum entropy quantile estimation and dual optimization method[J]. Journal of Aerospace Power, 2017, 32(2): 416-423.

度非线性及拟牛顿法的缺陷,使得寻优结果对初始点的选取较为敏感.初始点的选取不当,往往造成优化结果不理想或寻优结果不收敛.

针对上述问题,本文提出了对偶型-逐次寻优方法确定拉格朗日系数,并在此基础上提出了精度更高的改进对偶型最大熵分位值估计,算例表明,与经典最大熵分位值估计方法比较,本文提出的方法计算精度高,优化迭代的收敛性好.

文中将由非线性最小二乘法确定拉格朗日系数的最大熵分位值估计称为经典型最大熵分位值估计;由对偶原理确定拉格朗日系数的最大熵分位值估计称为对偶型最大熵分位值估计.

1 概率加权矩

1979 年 Greenwood 提出了概率加权矩,与经典矩的统计相比,概率加权矩克服了小样本经典矩估计出现的大偏差问题.该方法已经在气象、水文、机械工程领域中得到了应用.

设随机变量 X 和概率分布函数 $u(x) = P(X \leq x)$, 则 X 的概率加权矩如下^[1-2, 7-12]:

$$M_{i,j,k} = E[X^i u^j (1-u)^k] = \int_0^1 x(u)^i u^j (1-u)^k du \quad (1)$$

其中 i, j, k 为任意实数.为了取得较好的统计特性,避免观测值的高次方造成的较大抽样误差,一般取 $i=1$;为了计算方便,令 $j=0$ 或 $k=0$, 概率加权矩可表示为

$$M_{1,0,k} = \int_0^1 x(u) (1-u)^k du \quad (2)$$

$$M_{1,j,0} = \int_0^1 x(u) u^j du \quad (3)$$

其中 $M_{1,0,k}$ 为 k 阶超过概率权重矩, $M_{1,j,0}$ 为 j 阶不及概率权重矩.

样本观测系列 $x_1, x_2, \dots, x_n$ 为递增离散系列 ($x_1 \leq x_2 \leq \dots \leq x_n$), 则样本概率加权矩可写成如下公式:

$$M_{1,0,k}^s = \sum_{i=1}^n [x_i (1-p_i)^k \Delta p_i] \quad (4)$$

$$M_{1,j,0}^s = \sum_{i=1}^n [x_i p_i^j \Delta p_i] \quad (5)$$

其中 x_i 为样本从小到大排序后,序号为 i 的观测值; $p_i = P(X \leq x_i)$ 为相应于 x_i 的积累频率; $\Delta p_i = P(X = x_i)$ 为相应于 x_i 的频率; $(1-p_i)^k$ 和 p_i^j 的估计公式为

$$(1-p_i)^k = \frac{C_{n-i}^k}{C_{n-1}^k} = \frac{(n-k-1)(n-k-2)\dots(n-k-i+1)}{(n-1)(n-2)\dots(n-i+1)} \quad (6)$$

$$p_i^j = \frac{C_{i-1}^j}{C_{n-1}^j} = \frac{(i-1)(i-2)\dots(i-j)}{(n-1)(n-2)\dots(n-j)} \quad (7)$$

其中 $(i=1, 2, \dots, n; j=0, 1, \dots, n; k=0, 1, \dots, n)$, $\Delta p_i = \frac{1}{n}$; 则 $M_{1,0,k}^s$ 和 $M_{1,j,0}^s$ 分别是 $M_{1,0,k}$ 和 $M_{1,j,0}$ 的无偏估计值.

对于小样本概率加权矩具有良好的无偏性.因此,利用样本的概率加权矩来估计概率统计量精度较高.

2 对偶型最大熵分位值估计

2.1 对偶型不及概率最大熵分位值估计

1) 经典不及概率最大熵分位值估计

设 $x(u)$ 为随机变量 X 在概率累积值为 u 的分位值 ($u(x) = P(X \leq x)$). 其不及概率分位值最大熵定义为^[1, 13-18]

$$\begin{aligned} \max \quad & S = - \int_0^1 [x(u) \ln x(u)] du \\ \text{s. t.} \quad & \int_0^1 u^j x(u) du = b_j \quad (j = 0, 1, \dots, n) \end{aligned} \quad (8)$$

其中 $b_j = M_{1,j,0}$ ($j = 0, 1, \dots, n$), n 为约束条件个数.

现有文献[1, 13-18]中确定不及概率最大熵分位值解析式及其系数一般采用如下形式:

不及概率最大熵分位值函数的解析形式为

$$x(u) = \exp\left(\sum_{j=0}^n \lambda_j u^j\right) \quad (9)$$

$$\lambda_0 = -\ln\left[\int_0^1 \exp\left(\sum_{j=1}^n \lambda_j u^j\right) du / b_0\right] \quad (10)$$

其中 $\lambda_0, \lambda_1, \dots, \lambda_j$ 为拉格朗日乘子.一般情况,利用最小二乘法原理,使残差余量 e_i 的平方和为最小的条件来确定 λ_j ^[1, 13-18], 即

$$\min \quad e^2 = \sum_{i=1}^n e_i^2 \quad (11)$$

其中

$$e_i = 1 - b_0 \frac{\int_0^1 u^i \exp\left(\sum_{j=1}^n \lambda_j u^j\right) du}{b_i \int_0^1 \left(\sum_{j=1}^n \lambda_j u^j\right) du} \quad i = 1, 2, \dots, n \quad (12)$$

目前普遍采用牛顿迭代寻优求解上式.但是,采用牛顿迭代寻优求解存在如下不足,即初值点对寻优结果比较敏感.其一,初值点选取不合理,寻优结果就不理想.其二,上述残差 e 是关于 λ_j 的高度非线性,往往造成寻优结果难以收敛.

2) 对偶型不及概率最大熵分位值估计

根据式(8),本文重新构造拉格朗日函数 T ,且 $\lambda_0, \lambda_1, \dots, \lambda_j$ 为拉格朗日乘子,将其转化为无约束优化问题,即有

$$\begin{aligned} \max \quad T = & - \int_0^1 [x(u) \ln x(u)] du - (\lambda_0 - 1) \cdot \\ & \left[\int_0^1 x(u) du - b_0 \right] - \sum_{j=1}^n \lambda_j \left[\int_0^1 u^j x(u) du - b_j \right] \end{aligned} \quad (13)$$

由驻值条件 $\frac{\partial T}{\partial x(u)} = 0$, 得到不及概率最大熵分位值函数的解析形式为

$$x(u) = \exp\left(-\sum_{j=0}^n \lambda_j u^j\right) \quad (14)$$

为了确定 λ_0 与 $\lambda_j (j=1, 2, \dots, n)$ 的关系,将式(14)代入约束条件 $\int_0^1 x(u) du = b_0$ 中,得到

$$\lambda_0 = \ln \left[\int_0^1 \exp\left(-\sum_{j=1}^n \lambda_j u^j\right) du / b_0 \right] \quad (15)$$

将式(14)代入式(13),得到拉格朗日对偶形式的寻优函数 P 为

$$\begin{aligned} \min \quad P = & - \int_0^1 \exp\left(-\sum_{i=0}^n \lambda_i u^i\right) \left(-\sum_{j=0}^n \lambda_j u^j\right) du - \\ & (\lambda_0 - 1) \left[\int_0^1 \exp\left(-\sum_{j=0}^n \lambda_j u^j\right) du - b_0 \right] - \\ & \sum_{j=1}^n \lambda_j \left[\int_0^1 \exp\left(-\sum_{i=0}^n \lambda_i u^i\right) u^j du - b_j \right] = \sum_{j=0}^n \lambda_j b_j \end{aligned} \quad (16)$$

对式(16)进行寻优即可得到拉格朗日系数 λ_j . 进一步可确定不及概率最大熵分位值函数的式(14).

2.2 对偶型超越概率最大熵分位值估计

1) 经典超越概率最大熵分位值估计

设 $x(q)$ 为随机变量 X 超越概率值为 q 的分位值 ($q = P(X \geq x(q))$). $x(u)$ 为随机变量 X 在不及概率为 u 的分位值 ($u = P(X \leq x(u))$), 且满足

$q = 1 - u$, 由概率权重矩可得

$$\begin{aligned} M_{1,0,k} &= \int_0^1 x(u) (1-u)^k du = \\ & \int_0^1 x(q) q^k d(1-q) = \int_0^1 x(q) q^k dq \end{aligned} \quad (17)$$

则超越概率最大熵定义为^[14-15, 18]

$$\begin{aligned} \max \quad U = & - \int_0^1 [x(q) \ln x(q)] dq \\ \text{s. t.} \quad & \int_0^1 q^k x(q) dq = a_k \quad (k = 0, 1, \dots, m) \end{aligned} \quad (18)$$

其中 $a_k = M_{1,0,k} (k = 0, 1, \dots, m)$, m 为约束条件个数.

与不及概率分位值函数相似,超越概率最大熵分位值函数的解析形式为

$$x(q) = \exp\left(\sum_{k=0}^m \gamma_k q^k\right) \quad (19)$$

$$\gamma_0 = -\ln \left[\int_0^1 \exp\left(\sum_{k=1}^m \gamma_k q^k\right) dq / a_0 \right] \quad (20)$$

在确定分位值函数 $\gamma_0, \gamma_1, \dots, \gamma_k$ 时,同样遇到优化函数结构形式与式(11)和式(12)类似,也同样存在类似的不足.

2) 对偶型超越概率最大熵分位值估计

根据式(18),本文重新构造拉格朗日函数 R ,且 $\gamma_0, \gamma_1, \dots, \gamma_k$ 为拉格朗日乘子,将其转化为无约束优化问题,即有

$$\begin{aligned} \max \quad R = & - \int_0^1 [x(q) \ln x(q)] dq - \\ & (\gamma_0 - 1) \left[\int_0^1 x(q) dq - a_0 \right] - \\ & \sum_{k=1}^m \gamma_k \left[\int_0^1 q^k x(q) dq - a_k \right] \end{aligned} \quad (21)$$

由驻值条件 $\frac{\partial R}{\partial x(q)} = 0$, 得超越概率最大熵分位值函数解析形式为

$$x(q) = \exp\left(-\sum_{k=0}^m \gamma_k q^k\right) \quad (22)$$

为了确定 γ_0 与 $\gamma_k (k=1, 2, \dots, m)$ 的关系,将

式(22)代入约束条件 $\int_0^1 x(q) dq = a_0$ 中,得到

$$\gamma_0 = \ln \left[\int_0^1 \exp\left(-\sum_{k=1}^m \gamma_k q^k\right) dq / a_0 \right] \quad (23)$$

将式(22)代入式(21),得到拉格朗日对偶形式的寻优函数 W 为

$$\begin{aligned} \min \quad W = & - \int_0^1 \exp\left(-\sum_{i=0}^m \gamma_i q^i\right) \left(-\sum_{k=0}^m \gamma_k q^k\right) dq - \\ & (\gamma_0 - 1) \left[\int_0^1 \exp\left(-\sum_{k=0}^m \gamma_k q^k\right) dq - a_0 \right] - \\ & \sum_{k=1}^m \gamma_k \left[\int_0^1 \exp\left(-\sum_{i=0}^m \gamma_i q^i\right) q^k dq - a_k \right] = \sum_{k=0}^m \gamma_k a_k \end{aligned} \tag{24}$$

对式(24)进行寻优即可得到拉格朗日系数 $\gamma_k (k=1, 2, \cdots, m)$, 进一步可确定超越概率最大熵分位值函数式(22)。

2.3 改进的对偶型最大熵分位值估计

不及概率最大熵分位值估计在拟合百分位值函数时,其低概率累积分位点相对其他分位点误差大,而超越概率最大熵分位值函数估计在拟合百分位值函数时,其高概率累积分位点相对其他分位点误差大。将两者进行线性组合可以提高分位值函数的估计精度。

设 $x(u)$ 为随机变量 X 在概率累积值为 u 的分位值 ($u = P(X \leq x(u))$), 由式(14)和式(22), 则改进的分位值函数最大熵解析式为

$$\begin{aligned} x(u) = & u \exp\left(-\sum_{j=0}^n \lambda_j u^j\right) + \\ & (1-u) \exp\left[-\sum_{k=0}^m \gamma_k (1-u)^k\right] \end{aligned} \tag{25}$$

3 对偶型最大熵分位值估计的优化算法及实现

为了进一步解决经典最大熵、对偶型最大熵以及改进的最大熵等分位值函数(即式(9)、式(14)、式(19)、式(22)和式(25))的求解过程中寻优计算拉格朗日乘子的初始点敏感性问题,即寻优结果精确不高、寻优不收敛和求解效率低等问题,本文提出采用逐次寻优的优化算法求解拉格朗日乘子。

由于在 Bayes 分析及数据处理中,各阶样本矩 b_k 或 a_k 是不相关的^[3],所以本文的逐次寻优算法在寻优计算拉格朗日乘子时,从低阶到高阶逐次加入概率权重矩,并更新最大熵函数表达式。

下面以不及概率最大熵分位值函数的计算实现过程为例,阐述本文提出的逐次寻优算法的步骤

设拉格朗日乘子的总数为 $M, n=1, 2, 3, \cdots, M-1$, 那么

1) 当 $n=1$ 时,优化变量为 $V_1=[\lambda_1]$,其初始值 $V_1=0$, 则 $\lambda_0 = \ln \left[\int_0^1 \exp(-\lambda_1 u) du / b_0 \right]$, 并根据优化函数 $\min P = \sum_{j=0}^1 \lambda_j b_j$, 采用单纯形优化求出最优变量为 $V_1^*=[\lambda_1^{*(1)}]$ 。

2) 当 $n=2$ 时,优化变量为 $V_2=[\lambda_1 \quad \lambda_2]$,其初始值为 $V_2 = [\lambda_1^{*(1)} \quad 0]$, 令 $\lambda_0 = \ln \left[\int_0^1 \exp\left(-\sum_{j=1}^2 \lambda_j u^j\right) du / b_0 \right]$, 根据优化函数 $\min P = \sum_{j=0}^2 \lambda_j b_j$, 采用单纯形优化求出最优变量为 $V_2^*=[\lambda_1^{*(2)} \quad \lambda_2^{*(2)}]$ 。

3) 增加 n 值,如此往复,直至 $n=M-1$ 步,得到最终的最优变量 V_n^* , 并由 $\lambda_0 = \ln \left[\int_0^1 \exp\left(-\sum_{j=1}^n \lambda_j u^j\right) du / b_0 \right]$ 得出 λ_0 。

4) 由最终的拉格朗日系数 $\lambda_0, \lambda_1, \cdots, \lambda_{M-1}$, 代入式(15)即可得到不及概率最大熵分位值函数。超越概率最大熵分位值的计算步骤与上述相同。

4 算例分析与讨论

4.1 常见概率分布类型的最大熵分位值计算分析

应用本文方法对常见的正态分布、对数正态分布、威布尔分布、指数分布等 4 种分布类型的不同概率分布分位值计算,以显示本文方法处理不同概率分布分位值的适应性。

假设随机样本分别服从威布尔分布 $W(2, 2, 0.5)$ (形状参数为 2, 位置参数为 2, 尺度参数为 0.5), 正态分布 $N(1, 0.2)$ (期望为 1, 标准差为 0.2), 对数正态分布 $\ln N$ (对数期望为 0, 对数标准差为 0.3), 指数分布 E (率参数为 4)。

通过 Monte Carlo 方法生成 1000 个样本,采用本文的对偶型不及概率最大熵分位值估计法、超越概率最大熵分位值估计法以及改进最大熵分位值估计法分别对该 4 种分布类型的分位值进行计算,并取所计算样本的前 10 阶概率权重矩进行约束,得到对偶型最大熵分位值函数曲线与理论累积函数曲线的对比见图 1。

从图 1 可以看出,在大样本条件下,对偶型不及概率最大熵分位值估计法、超越概率最大熵分位值估计法以及改进对偶型最大熵分位值估计法的计算结果与理论概率累积函数结果均十分吻合,表明本文提出的对偶型最大熵分位值逐次寻

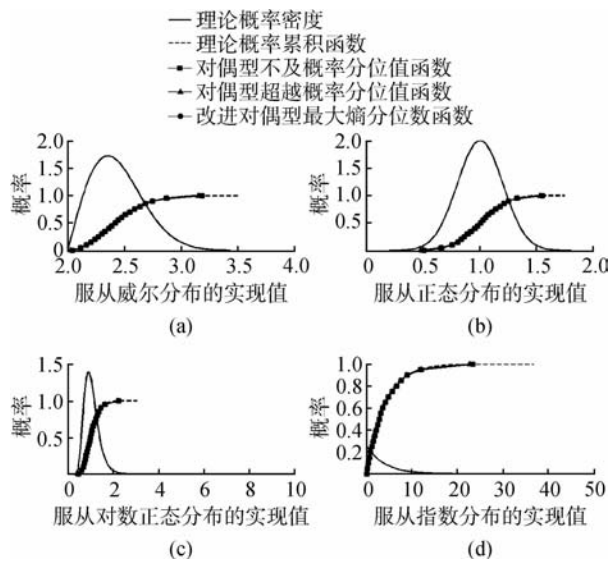


图 1 常见分布类型的对偶型最大熵分位值计算结果对比
Fig. 1 Comparison of dual maximum entropy quantile calculation results of common distribution types

优方法能够很好地模拟常见类型的概率分布;且公式表达形式统一、简单。

4.2 对偶型最大熵与经典最大熵分位值计算精度对比

通过 Monte Carlo 方法生成 1000 个样本,分别服从威布尔分布 $W(2,2,0.5)$ (形状参数为 2,位置参数为 2,尺度参数为 0.5),正态分布 $N(1,0.2)$ (期望为 1,标准差为 0.2),对数正态分布 $\ln N(0,0.3)$ (对数期望为 0,对数标准差为 0.3),指数分布 $E(4)$ (率参数为 4)。

将本文的对偶型超越概率分位值最大熵估计、和经典超越概率分位值最大熵估计分别采用本文的逐次寻优算法和经典的拟牛顿算法计算分位值(不及概率分位值计算类同),取所计算样本的前 10 阶概率权重矩进行约束,得到最大熵分位值函数曲线与理论概率累积函数曲线的对比见图 2,不同分位值函数估计方法产生的方均根误差(RMSE)见表 1。

从图 2 及表 1 的结果可以看到拟合曲线的精确程度从高到低的排序为:对偶型-逐次寻优法、对偶型-拟牛顿法、经典型-逐次寻优法、经典型-拟牛顿法。

分析其原因为:经典型-拟牛顿法得出的拟合曲线与理论曲线相差很大,由于优化函数的高度非线性存在寻优困难,同时拟牛顿法的初始点选取十分重要。本算例中由于初始点的选取不合适,导致优化不收敛。而逐次寻优法避免初始点的选

取不当造成的优化不收敛问题,使经典型-逐次寻优法比经典型-拟牛顿法有更高的计算精度。对偶型分位值函数简化优化函数,使其不再高度非线性,得到的计算曲线较经典型具有更高的精度。

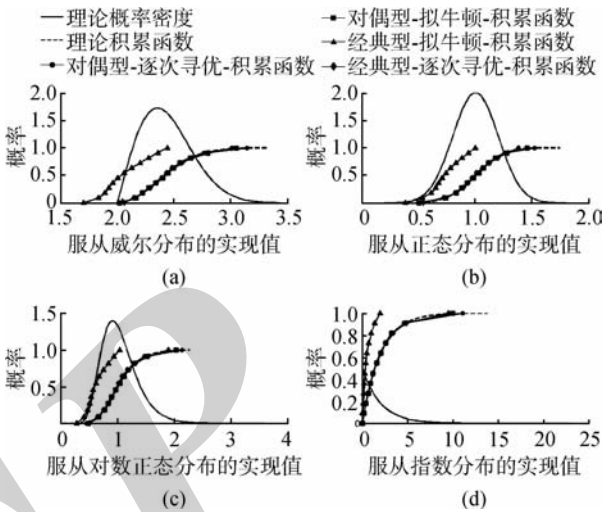


图 2 对偶型分位值函数与经典分位值函数对比
Fig. 2 Comparison between dual quantile function and classic quantile function

表 1 不同分位值估计方法产生的方均根误差
Table 1 RMSE caused by different quantile estimation method

分布类型	估计方法			
	对偶型-逐次寻优	对偶型-拟牛顿法	经典型-逐次寻优	经典型-拟牛顿法
威布尔分布	0.012	0.018	0.029	0.383
正态分布	0.010	0.014	0.024	0.274
对数正态	0.022	0.030	0.044	0.428
指数分布	0.166	0.220	0.223	2.058

4.3 对偶型最大熵逐次寻优方法对分位值计算精度讨论

通过 Monte Carlo 方法生成 1000 个服从正态分布 $N(10,1)$ 的样本,应用对偶型不及概率最大熵分位值估计法、超越概率最大熵分位值估计法以及改进的最大熵分位值估计法等 3 种,对概率累积函数进行计算,给出累积概率在 $[0,0.2]$, $[0.8,1]$, $[0,1]$ 曲线结果进行对比见图 3~图 5,及各概率区间的分位值估计方均根方差 RMSE 见表 2~表 4(其他分布如对数正态分布,威布尔分布,指数分布等概率累积函数计算精度也有类似的结果)。

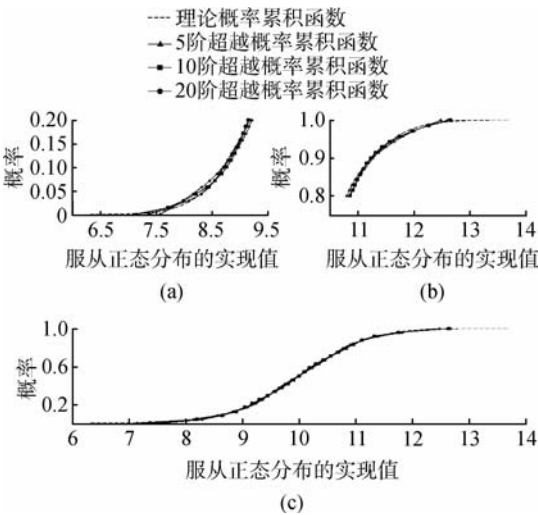


图 3 不同阶数的超越概率累积函数曲线
Fig. 3 Cumulative function curves of exceedance probability with different orders

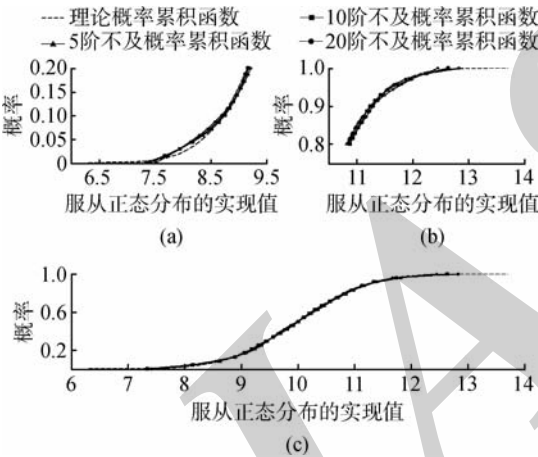


图 4 不同阶数的不及概率累积函数曲线
Fig. 4 Cumulative function curves of insufficient probability with different orders

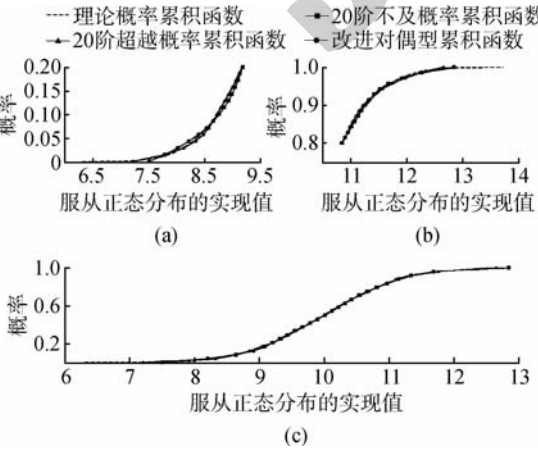


图 5 3 种不同最大熵累积函数曲线对比
Fig. 5 Comparison of three kinds of maximum entropy cumulative function curves

表 2 对偶型超越概率分位值估计方均根误差
Table 2 RMSE of dual exceedance probability quantile estimation

概率 区间	对偶型超越概率累积函数阶数		
	5 阶	10 阶	20 阶
$[0,0.2]$	0.1339	0.0971	0.0747
$[0.8,1]$	0.1233	0.1006	0.1001
$[0,1]$	0.0835	0.0647	0.0591

表 3 对偶型不及概率分位值估计方均根误差
Table 3 RMSE of dual insufficient probability quantile estimation

概率 区间	对偶型不及概率累积函数阶数		
	5 阶	10 阶	20 阶
$[0,0.2]$	0.1328	0.1012	0.1007
$[0.8,1]$	0.1182	0.0956	0.0708
$[0,1]$	0.0824	0.0641	0.0584

表 4 3 种不同最大熵分位值估计方均根误差
Table 4 RMSE of three kinds of maximum entropy quantile estimation

概率 区间	最大熵累积函数		
	20 阶不及概率 累积函数	20 阶超越概率 累积函数	20 阶改进对 偶型累积函数
$[0,0.2]$	0.1048	0.0770	0.0757
$[0.8,1]$	0.0643	0.1002	0.0637
$[0,1]$	0.0572	0.0579	0.0460

从图 3~图 5 及表 2~表 4 可以看出

- 1) 3 种最大熵估计法都能很好地计算常见分布类型的累积函数曲线,随着约束条件阶数的增加(n 阶约束对应的最大熵称为 n 阶最大熵估计),计算的曲线更加逼近理论概率累积函数曲线.
- 2) 从图 3(a)和图 3(b)及表 2 可以看出,超越概率最大熵分位值估计法得到的累积函数曲线在低可靠度的部分随着阶数增大其精度提高变化不大,在高可靠度的部分随着阶数增大其精度提高变化明显.从图 4(a)和图 4(b)及表 3 可以看出,不及概率最大熵分位值估计法得到的累积函数曲线在低可靠度的部分随着阶数增大其精度提高变化明显,在高可靠度的部分随着阶数增大其精度提高变化不大.

3) 从图 5 及表 4 可以看出,由于改进的最大熵分位值函数估计法是不及概率最大熵分位值估计法与超越概率最大熵分位值估计法上的线性组合,综合了两者的优点,并且避开了两者的不足,估计出的概率累积曲线较前两者方法精确。

4.4 改进对偶型最大熵逐次寻优方法在复杂分布函数中的应用

最大熵分位值估计法不仅能方便地应用在常见的单峰随机变量,且对于多峰的复杂随机变量也能很好的适用。利用改进对偶型最大熵分位值估计法及经典最大熵分位值估计法对随机变量 X 进行累积函数估计,随机变量 X 的理论概率密度为 $f(x) = \frac{1}{2\sqrt{2\pi} \times 0.5} \cdot \exp\left[-\frac{(x-20)^2}{2 \times 0.5^2}\right] + \frac{1}{2\sqrt{2\pi} \times 1} \cdot \exp\left[-\frac{(x-23)^2}{2 \times 1^2}\right]$ 。通过 Monte Carlo 方法生成 1000 个服从 X 分布的子样,利用 20 阶改进对偶型最大熵分位值估计法、经典最大熵分位值估计法对子样的累积函数曲线进行拟合,计算结果对比见图 6 及表 5。

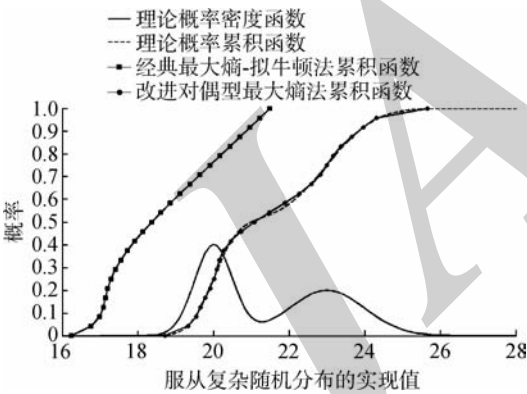


图 6 复杂随机变量的累积函数曲线对比

Fig. 6 Comparison of cumulative probability function curves about complex variable

表 5 改进对偶型及经典型最大熵分位值方均根误差
Table 5 RMSE of advanced dual and classic maximum entropy quantile

概率 区间	经典型最大熵法 方均根误差	改进对偶型最大 熵法方均根误差
[0,1]	2.8576	0.0650

从图 6 及表 5 可以看到,在一定的样本条件下,经典的熵分位值函数曲线与理论曲线相差较大。其原因在于经典最大熵法优化函数高度

非线性且优化结果受初始点影响较大;而改进对偶型最大熵分位值逐次寻优方法得到的累积函数曲线与理论累积函数曲线吻合较好,精度达到工程的使用要求。

5 结 论

1) 基于拉格朗日乘子对偶法,推导建立了含有拉格朗日系数优化函数的对偶表达式,与经典的非线性最小二乘法相比,函数形式简单,非线性程度低,不论采用本文提出的逐次寻优算法和经典的拟牛顿优化算法,计算精度均较高。

2) 本文提出的一种逐次寻优算法求解拉格朗日乘子,不仅能够克服初值敏感性问题,优化收敛好,而且求解效率较高。

3) 本文提出的对偶型-逐次寻优方法与其他方法的相比,计算精度较高。其中改进的对偶型-逐次寻优方法的计算精度最高,并且能够很好地适应于复杂分布类型的累积函数和分位值计算。

参考文献:

[1] Pandey M D. Direct estimation of quantile functions using the maximum entropy principle [J]. Structural Safety, 2000, 22(1): 61-79.

[2] 袁修开,吕震宙,岳珠峰. 小样本下分位数函数的 Bootstrap 置信区间估计 [J]. 航空学报, 2012, 33(10): 1842-1849.

YUAN Kaixiu, LU Zhenzhou, ZHU Yuefeng. Bootstrap confidence interval of quantile function estimation for small samples [J]. Acta Aeronautica et Astronautica Sinica, 2012, 33(10): 1842-1849. (in Chinese)

[3] Wu X. Calculation of maximum entropy densities with application to income distribution [J]. Journal of Econometrics, 2003, 115(2): 347-354.

[4] Shi X, Teixeira A P, Zhang J, et al. Structural reliability analysis based on probabilistic response modelling using the maximum entropy method [J]. Engineering Structures, 2014, 70(9): 106-116.

[5] Hehua Z, Yulong Z, Xiaojun L. Estimation of the fracture diameter distributions using the maximum entropy principle [J]. International Journal of Rock Mechanics and Mining Sciences, 2014, 72: 127-137.

[6] Li C, Wang W, Wang S. Maximum entropy monte carlo method for the evaluation of dam overtopping probability [J]. Disaster Advances, 2012, 5(4): 1143-1147.

[7] 邓建,古德生,李夕兵. 确定可靠性分析 Weibull 分布参数的概率加权矩法 [J]. 计算力学学报, 2004, 21(5): 609-613.

DENG Jian, GU Desheng, LI Xibing. Parameters and quantile estimation for fatigue life distribution using probability weighted moments [J]. Chinese Journal of Computational

Mechanics,2004,21(5):609-613. (in Chinese)

[8] Deng J,Pandey M D. Estimation of minimum cross-entropy quantile function using fractional probability weighted moments[J]. Probabilistic Engineering Mechanics, 2009, 24 (1):43-50.

[9] Caeiro F, Ivette G M, Vandewalle B. Semi-parametric probability-weighted moments estimation revisited [J]. Methodology and Computing in Applied Probability,2014, 16(1):1-29.

[10] Li X,Zuo Y,Zhuang X,et al. Estimation of fracture trace length distributions using probability weighted moments and L-moments[J]. Engineering Geology, 2014, 168(2): 69-85.

[11] ZHANG Ming,BAI Shaoquang. Application of probability-weighted mixed moment method to parameters estimation for three-parameter probability distribution[J]. Water Resources and Power,2013,31(8):16-18.

[12] ZHANG Ming, BAI Shaoguang. Probability-weighted mixed moments method and its application to estimation parameters of pearson type: III distribution[J]. Water Resources and Power,2012,30(9):20-29.

[13] Deng J,Pandey M D,Xie W C. Maximum entropy principle and partial probability weighted moments[R]. Washington D C:American Institute of Physics,2012.

[14] Deng J,Pandey M D. Derivation of sample oriented quantile function using maximum entropy and self-determined probability weighted moments[J]. Environmetrics, 2010, 21(2):113-132.

[15] Deng J,Pandey M D. Using partial probability weighted moments and partial maximum entropy to estimate quantiles from censored samples[J]. Probabilistic Engineering Mechanics,2009,24(3):407-417.

[16] Yu L J,Zhang X N. Maximum entropy models for mode split forecasting[R]. New York:International Conference on Service Operations,Logistics and Informatics,2009.

[17] Yu L J. A new calculation algorithm for quantile function with application to headway distribution[R]. New York: International Conference on Service Operations, Logistics and Informatics,2008.

[18] Deng J,Pandey M D. Estimation of the maximum entropy quantile function using fractional probability weighted moments[J]. Structural Safety,2008,30(4):307-319.

(编辑:王碧珺)