关于单位成本信道容量的注记

田大钢
（上海理工大学 复杂系统科学研究中心，上海 200093）

摘 要：针对单位成本信道容量的计算问题，首先将一个联系单位成本信道容量与容量成本函数的基本定理推广到一般凸函数的情况，并证明了推广的单位成本容量函数具有单峰性，然后给出了一个利用单峰函数的性质计算单位成本信道容量的算法。同时，针对信道容量成本函数的计算问题，证明了这类带约束的信道容量的计算问题可以化为一个目标函数具有自协调性的凸优化问题，并给出了具体的算法和算法的计算量。

关键词：信道容量；单位成本信道容量；自协调函数；计算量
中图分类号：TN911.2 **文献标识码**：A **文章编号**：1000-436X(2011)04-0032-07

Notes on channel capacity per unit cost

TIAN Da-gang
(Research Center for Complex System Science, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: For the computation of the capacity per unit cost, an extension of the basic theorem connect the capacity per unit cost with the capacity-cost function was made, and the single peak of the generalized capacity-cost function per unit cost was proved. Then, an algorithm to locate the optimal point of the generalized capacity-cost function per unit cost based on the single peak of the function was proposed. Meanwhile, it was proved that the computation of the capacity-cost function, which is a constrained channel capacity issue, can be expressed as a convex optimization, to which the objective function is self- concordant. The algorithm and its computational complexity were given.

Key words: channel capacity; channel capacity per unit cost; self-concordant functions; computational complexity

1 引言

信道容量是传输信道的最重要的参数之一，代表了信道传输信息的最大能力。与信道容量计算有关的问题是信息理论的经典问题，一直受到人们的重视[1~3]。文献[1]给出了这方面的一些精细的新成果，并说明，不存在求解一般离散无记忆信道的信道容量的显式公式（二元信道及一些特殊情况除外）。这使得具有多项式时间性的迭代算法变得十分重要。传统的 Arimoto-Blahut 算法具有多项式时间性，也是目前实际中使用的主要方法，但对带约束的信道容量问题的作用十分有限。一般来说，信道的使用不是无代价的，如果考虑信号传输的代价，就出现带约束条件的信道容量问题，如信道容量成本函数问题[4]、单位成本信道容量问题[5~10]和能量受限的信道容量[11]问题。除了成本之类的约束外，通信系统中的约束形式还包括安全性约束[12]，公平性约束和流量控制等[13]，表现在信道容量的计算上，就是带各种约束的信道容量问题。对这些带约束的信道容量问题，目前还没有具有普遍性的有

收稿日期：2010-05-06；**修回日期**：2011-01-22
基金项目：上海市重点学科基金资助项目（S30504）；国家科技支撑计划基金资助项目（2008BADA6B01）
Foundation Items: The Leading Academic Discipline Project of Shanghai Municipal Government (S30504); The National Science and Technology Pillar Program (2008BADA6B01)

效算法。

文献[7]是关于单位成本信道容量的计算问题的经典文献，这是一个典型的带约束信道容量问题，本文给出其中一个基本定理的推广形式，并证明了表示单位成本信道容量的函数具有单峰性，利用单峰函数的性质，得到一个具有多项式时间性的算法。

信道容量的计算本来就是一个凸优化问题。在凸优化领域，以自协调函数理论为代表的内点算法，已经导致了算法理论的一场革命[14]。本文将证明，在信道容量的计算问题中，尤其是带约束条件的信道容量的计算问题中，可以而且应该利用自协调函数理论。

2　基本结果

定理　设 $D\in R^n$ 是凸集，$F(x)$是 D 上有上界的上凸函数，$G(x)$是 D 上的下凸函数。记 $a=\inf\{G(x)|x\in D\}$，$b=\sup\{G(x)|x\in D\}$，并设 $a\geqslant 0$。记 $C(\beta)=\sup\{F(x)|G(x)\leqslant\beta\}$，这是推广形式的容量成本函数。则

1) $C(\beta)$是不减的上凸函数。

2) 如下定义的广义单位成本容量函数

$$\rho(\beta)=\frac{C(\beta)}{\beta}$$

是单调或单峰函数。

3) 下式成立：

$$\sup_D\frac{F(x)}{G(x)}=\sup_{a<\beta<b}\frac{C(\beta)}{\beta}$$

证明

1) 设$\beta_1>\beta_2$，则$\{x|G(x)\leqslant\beta_2\}\subseteq\{x|G(x)\leqslant\beta_1\}$，所以 $C(\beta_2)\leqslant C(\beta_1)$；对任何$\varepsilon>0$，设 x_1 满足 $G(x_1)\leqslant\beta_1$ 和 $C(\beta_1)=F(x_1)+\varepsilon$，设 x_2 满足 $G(x_2)\leqslant\beta_2$ 和 $C(\beta_2)=F(x_2)+\varepsilon$，由 G 的下凸性，有：

$$\begin{aligned}&G(\alpha x_1+(1-\alpha)x_2)\\&\leqslant\alpha G(x_1)+(1-\alpha)G(x_2)\\&\leqslant\alpha\beta_1+(1-\alpha)\beta_2\end{aligned}$$

由 F 的上凸性和 $C(\beta)$的定义，得：

$$\begin{aligned}&\alpha C(\beta_1)+(1-\alpha)C(\beta_2)\\&=\alpha F(x_1)+(1-\alpha)F(x_2)+\varepsilon\\&\leqslant F(\alpha x_1+(1-\alpha)x_2)+\varepsilon\\&\leqslant C(\alpha\beta_1+(1-\alpha)\beta_2)+\varepsilon\end{aligned}$$

而ε是任意的，所以 $C(\beta)$是上凸的。

2) 不妨设 $C(\beta)$是可微函数，否则，用逼近好了[15]。设存在$\beta_0\in(a,b)$，使得

$$\rho'(\beta)|_{\beta=\beta_0}=0$$

即

$$C'(\beta_0)\beta_0-C(\beta_0)=0$$

任取$\beta_1>\beta_0$，由 $C(\beta)$的上凸性，有 $C'(\beta_1)\leqslant C'(\beta_0)$。令

$$y_1(\beta)=C'(\beta_1)(\beta-\beta_0)+C(\beta_0)$$

这是经过点（β_0，$C(\beta_0)$），斜率为 $C'(\beta_1)$的直线。由 $C(\beta)$的单调不减性和上凸性得：$y_1(\beta_1)\leqslant C(\beta_1)$，于是

$$\begin{aligned}C'(\beta_1)-\frac{C(\beta_1)}{\beta_1}&\leqslant C'(\beta_1)-\frac{y_1(\beta_1)}{\beta_1}\\&=C'(\beta_1)\frac{\beta_0}{\beta_1}-\frac{C(\beta_0)}{\beta_1}\\&\leqslant C'(\beta_0)\frac{\beta_0}{\beta_1}-\frac{C(\beta_0)}{\beta_1}=0\end{aligned}$$

即$\rho'(\beta_1)\leqslant 0$。任取$\beta_2<\beta_0$，由 $C(\beta)$的上凸性，有 $C'(\beta_2)\geqslant C'(\beta_0)$。令

$$y_0(\beta)=C'(\beta_0)(\beta-\beta_0)+C(\beta_0)$$

这是经过点（β_0, $C(\beta_0)$），斜率为 $C'(\beta_0)$ 的直线。由 $C(\beta)$的单调不减性和上凸性得：$y_0(\beta_2)\geqslant C(\beta_2)$，于是，

$$\begin{aligned}C'(\beta_2)-\frac{C(\beta_2)}{\beta_2}&\geqslant C'(\beta_2)-\frac{y_0(\beta_2)}{\beta_2}\\&\geqslant C'(\beta_0)-\frac{y_0(\beta_2)}{\beta_2}\\&=C'(\beta_0)\frac{\beta_0}{\beta_2}-\frac{C(\beta_0)}{\beta_2}=0\end{aligned}$$

即$\rho'(\beta_2)\geqslant 0$。从而$\beta_0$是$\rho(\beta)=C(\beta)/\beta$的最大点。

3) 若

$$\sup_D\frac{F(x)}{G(x)}=+\infty$$

则任意的 $N>0$，存在 x_0，使得：

$$\frac{F(x_0)}{G(x_0)}>N$$

设 $G(x_0)=\beta_0$，则

$$C(\beta_0)=\sup\{F(x)|G(x)\leqslant\beta_0\}\geqslant F(x_0)$$

所以

$$\sup_{a<\beta<b}\frac{C(\beta)}{\beta}\geqslant\frac{C(\beta_0)}{\beta_0}\geqslant\frac{F(x_0)}{G(x_0)}>N$$

由 N 的任意性，得：

$$\sup_{a<\beta<b}\frac{C(\beta)}{\beta}=+\infty$$

反过来，若

$$\sup_{a<\beta<b}\frac{C(\beta)}{\beta}=+\infty$$

则对任何 $N>0$，存在 $a<\beta_0<b$，使得

$$\frac{C(\beta_0)}{\beta_0}>N$$

而

$$\sup_D\frac{F(x)}{G(x)}\geqslant\sup_{G(x)\leqslant\beta_0}\frac{F(x)}{G(x)}$$

$$\geqslant\sup_{G(x)\leqslant\beta_0}\frac{F(x)}{\beta_0}=\frac{C(\beta_0)}{\beta_0}>N$$

所以

$$\sup_D\frac{F(x)}{G(x)}=+\infty$$

由以上讨论可知：

$$\sup_D\frac{F(x)}{G(x)}<+\infty \tag{1}$$

与

$$\sup_{a<\beta<b}\frac{C(\beta)}{\beta}<+\infty \tag{2}$$

同时成立。

设式(1)成立，则对任何$\varepsilon>0$，存在 $x_0\in D$，使得：

$$\frac{F(x_0)}{G(x_0)}+\varepsilon=\sup_D\frac{F(x)}{G(x)}$$

设 $G(x_0)=\beta_0$，则

$$\sup_D\frac{F(x)}{G(x)}=\frac{F(x_0)}{G(x_0)}+\varepsilon$$

$$\leqslant\frac{C(\beta_0)}{\beta_0}+\varepsilon\leqslant\sup_{a<\beta<b}\frac{C(\beta)}{\beta}+\varepsilon$$

由ε的任意性，得：

$$\sup_D\frac{F(x)}{G(x)}\leqslant\sup_{a<\beta<b}\frac{C(\beta)}{\beta}$$

另一方面，由于式(2)也成立，故对任何$\varepsilon>0$，存在 $a<\beta_0<b$，使得

$$\frac{C(\beta_0)}{\beta_0}+\varepsilon=\sup_{a<\beta<b}\frac{C(\beta)}{\beta}$$

设 $x_0\in D$，$G(x_0)\leqslant\beta_0$，满足

$$F(x_0)+\varepsilon\beta_0=\sup\{F(x)|\,G(x)\leqslant\beta_0\}=C(\beta_0),$$

则

$$\sup_D\frac{F(x)}{G(x)}\geqslant\sup_{G(x)\leqslant\beta_0}\frac{F(x)}{G(x)}$$

$$\geqslant\frac{F(x_0)}{\beta_0}=\sup_{a<\beta<b}\frac{C(\beta)}{\beta}-2\varepsilon$$

由ε的任意性，得：

$$\sup_D\frac{F(x)}{G(x)}\geqslant\sup_{a<\beta<b}\frac{C(\beta)}{\beta}$$

证毕

取 $F(x)$为互信息，$G(x)$为期望成本，则定理中的结论 3）就是文献[7]的定理 2。取 $F(x)$为熵，则得到单位成本的最大熵。而如果 $G(x)$是二次函数，定理的结论依然有效。

利用定理 1 中的结论 2）可以设计计算

$$\sup(F(x)/G(x))$$

的新算法，例如，可以写出如下的 0.618 法[16]：

令 $a=\inf\{G(x)|\,x\in D\}$，$b=\sup\{G(x)|\,x\in D\}$，并设 $a\geqslant0$。

算法 1

① 确定搜索区间$[a,b]$和精度要求 $\varepsilon>0$。计算 λ，μ

$$\lambda=a+0.382(b-a)$$

$$\mu=a+0.618(b-a)$$

计算 $C(\lambda)/\lambda$和 $C(\mu)/\mu$，令 $k=1$。

② 比较函数值。若 $C(\lambda)/\lambda>C(\mu)/\mu$，则转③；若 $C(\lambda)/\lambda\leqslant C(\mu)/\mu$，则转④。

③ 若 $b-\lambda\leqslant\varepsilon$，则停止，输出$\mu$和 $C(\mu)/\mu$。否则，令

$a=\lambda,\lambda=\mu,C(\lambda)/\lambda=C(\mu)/\mu,\mu=a+0.618(b-a)$，计算 $C(\mu)/\mu$，转⑤。

④ 若$\mu-a\leqslant\varepsilon$，则停止，输出$\lambda$和 $C(\lambda)/\lambda$；否则，令

$b=\mu,\mu=\lambda,C(\mu)/\mu=C(\lambda)/\lambda,\lambda=a+0.382(b-a)$，计算 $C(\lambda)/\lambda$，转⑤。

⑤ $k=k+1$，转②。

这里包括 $a=0$ 的情况，即零成本信号的情况。

3　计算 $C(\mu)$的自协调函数方法

上一节的讨论，依赖于 $C(\beta)=\sup\{F(x)|\ G(x)\leqslant\beta\}$的有效计算。这是一个带约束的凸规划问题，在一般情况下，给出一个计算复杂性很清楚的算法并不容易，尤其当约束是非线性的时候。不过，当 $F(x)$是信道容量或熵函数，$G(x)$是线性或二次函数时，引入适当的障碍函数后，可以利用自协调函数理论，建立有效的算法。

下文中涉及的对数函数都是以 2 为底的对数函数。

3.1　自协调函数

自协调函数理论的提出[17]，对凸优化理论的发展具有里程碑意义。这个理论揭示了线性规划问题内点算法的本质特性，并利用这些特性得到了求解凸优化问题的多项式时间算法，使得非线性优化理论更具备了实际操作的意义。文献[18]从不同的角度阐述了这个理论，文献[19]给出了一个通俗且实用的介绍。为了方便，这里对这个理论的主要概念和结论作一个简介。

定义 1（一元自协调函数）设 $D\subset R$，是开凸集。$f: D\to R$ 是三次可微的凸函数。称 f 是 D 上的自协调函数，如果对任何 $x\in D$，有：

$$|f'''(x)|\leqslant 2f''(x)^{3/2} \tag{3}$$

定义 2（多元自协调函数）设 $D\subset R^n$，是开凸集。$f: D\to R$ 是三次可微的凸函数。如果对任何 $x\in D$，$\boldsymbol{h}\in R^n$，记 $T=\{t\,|\,x+t\boldsymbol{h}\in D\}$，$g(t)=f(x+t\boldsymbol{h})$是 T 上的自协调函数，则称 f 是 D 上的自协调函数。

对于线性函数和正定二次函数来说，由于式(3)的左边都是 0，所以它们都是（平凡的）自协调函数。而应用中最典型的一元非平凡的自协调函数是$-\ln(x)$。

利用以下性质可以得到更多的自协调函数：

1) 若 f_i 是开凸集 $D_i\subset R^n$ 上的自协调函数，$i=1,\cdots,m$，记 $D=\cap_{i=1}^m D_i$，则

$$f=f_1+f_2+\cdots+f_n$$

是 D 上的自协调函数；

2) 设 $\boldsymbol{A}$: $R^m\to R^n$ 是线性映射，$x\in R^n$，$\boldsymbol{b}\in R^n$，若 $f(x)$是 $D\subset R^n$ 上的自协调函数，则 $f(\boldsymbol{A}y+\boldsymbol{b})$关于 y 是 $\bar{D}=\{y\in R^m\,|\,\boldsymbol{A}y+\boldsymbol{b}\in D\}$ 上的自协调函数；

3) 设 f_i 是开凸集 $D_i\subset R^{n_i}$ 上的自协调函数，$i=1,\cdots,m$，记 $D=D_1\times\cdots\times D_m$，则

$$f(x_1,\cdots,x_m)=f_1(x_1)+\cdots+f_m(x_m)$$

是 D 上的自协调函数。

3.2　自协调函数与凸规划

在下面的讨论中，对任何 $x^0\in D$，设水平集 $S=\{x\in D\,|\,f(x)\leqslant f(x^0)\}$是闭的。这意味着，当 x 趋于 D 的边界时，$f(x)$趋于无穷。

考虑优化问题

$$\min_D f(x) \tag{4}$$

这里 D 是 R^n 的非空开凸集，f 是 D 上的三阶连续可微的凸函数。设 $f(x)$下有界且自协调。

问题(4)的牛顿算法表述如下。

牛顿迭代

给定初始点 $x\in D$，允许误差$\varepsilon>0$。

1) 计算

$$\Delta x_{\mathrm{nt}}=-\nabla^2 f(x)^{-1}\nabla f(x);$$

$$\lambda(f,x)=(\nabla f(x)^{\mathrm{T}}\nabla^2 f(x)^{-1}\nabla f(x))^{1/2}$$

2) 停止准则：如果$\lambda(f,x)^2/2\leqslant\varepsilon$，则停止；否则执行 3）。

3) 如果 $\lambda(f,x)>1/4$，则

$$x=x+\frac{1}{1+\lambda(f,x)}\Delta x_{\mathrm{nt}}$$

否则 $x=x+\Delta x_{\mathrm{nt}}$，返回 1）。

特别之处在于，当 $f(x)$是自协调函数时，当$\lambda(f,x)>0.25$ 时，迭代一次目标函数值将至少减少 0.026 856。而当$\lambda(f,x_i)\leqslant 0.25$ 时，对于每个 $j\geqslant i$，有$\lambda(f,x_{j+1})\leqslant 2\lambda^2(f,x_j)\leqslant(1/2)\lambda(f,x_j)$和 $f(x_j)-p^*\leqslant\lambda^2(f,x_j)/2(1-\lambda(f,x_j))$。这里 p^*是式(4)的最优值。

上述牛顿算法的迭代次数不会超过下式：

$$O(1)(f(x^{(0)})-p^*)+\mathrm{lblb}(1/\varepsilon) \tag{5}$$

这里 $O(1)$与问题的规模无关。

对于有等式约束 $\boldsymbol{Ax}=\boldsymbol{b}$ 的优化问题，上述牛顿迭代算法和迭代次数的上界式(5)仍然成立。计算时只需注意初始点 $x^{(0)}$必须满足等式约束，并且将其中的牛顿步长Δx_{nt}和牛顿减量$\lambda(f,x)$作相应的改变就可以了。具体说来，牛顿步长Δx_{nt}由如下方程解出。

$$\begin{bmatrix}\nabla^2 f(x) & \boldsymbol{A}^{\mathrm{T}}\\ \boldsymbol{A} & \boldsymbol{0}\end{bmatrix}\begin{bmatrix}\Delta x_{\mathrm{nt}}\\ w\end{bmatrix}=\begin{bmatrix}-\nabla f(x)\\ \boldsymbol{0}\end{bmatrix}$$

这里假设 $\boldsymbol{A}$ 是 $p\times n$ 行满秩矩阵。牛顿减量由

$$\lambda(f,x)=(\Delta x_{\mathrm{nt}}^{\mathrm{T}}\nabla^2 f(x)\Delta x_{\mathrm{nt}})^{1/2}$$

给出。

更一般的优化问题具有如下的形式：

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t} \quad & f_i(x) \leqslant 0, \quad i = 1, \cdots, m \\ & \boldsymbol{A}x = \boldsymbol{b} \end{aligned} \tag{6}$$

这里 $f_0,\cdots,f_m$: $R^n \to R$ 是凸函数。$\boldsymbol{A}$ 是 $p\times n$ 行满秩矩阵。

利用对数障碍函数的方法，式(6)可以化为如下的形式

$$\begin{aligned} \min \quad & tf_0(x) + \phi(x) \\ \text{s.t} \quad & \boldsymbol{A}x = \boldsymbol{b} \end{aligned} \tag{7}$$

这里

$$\phi(x) = -\sum_{i=1}^{m} \text{lb}(-f_i(x))$$

是问题(6)的障碍函数。与只带等式约束的优化问题不同的是问题(7)的算法中要考虑参数 t 的选取。设 $x^*(t)$是(7)的最优解，p^*是问题(6)的最优值，由于有 $f_0(x^*(t))-p^* \leqslant m/t$，因此一种简单的做法是直接选取 $t = m/\varepsilon$，用牛顿迭代法求解，可以在式(5)界定的迭代步之内，获得需要的解。只是真正实施起来，由于这样选取的 t 往往很大，从而导致（$f(x^{(0)})-p^*$）可能很大，计算量也就会很大。实际中一般不使用这种方法。实际中使用的主要是如下的路径跟踪算法。

路径跟踪算法

给定：严格可行点 x，$t = t^{(0)} > 0$，$\mu > 1$ 和 $\varepsilon > 0$。

1) 对式(7)用牛顿迭代，求 $x^*(t)$。

2) 令 $x = x^*(t)$。

3) 如果 $m/t < \varepsilon$，则停止，输出 x；否则，转 4)。

4) $t = \mu t$，返回到 1)。

如果对任何 $t > 0$，$f(x) = tf_0(x) + \varphi(x)$是自协调和下有界的，并且，水平集 S 是闭的，则总的迭代次数不会超过

$$N = \left[\frac{\text{lb}(m/(t^{(0)}\varepsilon))}{\text{lb}\mu}\right]\left(O(1)m(\mu - 1 - \text{lb}\mu) + \text{lblb}(1/\varepsilon)\right)$$

其中，方括号中的部分是 t 从初始点变到 m/ε时，需要的迭代次数。在 t 变化过程中，每次的牛顿迭代次数由圆括号中的部分界定。

如果将 $f(x)$、$f'(x)$和 $f''(x)$本身的计算量用 M 表示，并且考虑到牛顿步长的计算量（主要是逆阵的计算），最终的计算量可以用

$$O(1)(M + n^3)m^{1/2}\text{lb}(V/\varepsilon) \tag{8}$$

来界定。这里的 V 与初始数据的选择有关。实际的计算量往往会比式(8)小很多。

3.3　基于自协调函数的信道容量的计算

设离散信道的输入有 m 个符号 $\boldsymbol{x} = \{x_1,\cdots,x_m\}$，$P(\boldsymbol{x}) = (p_1,\cdots,p_m)$表示 $\boldsymbol{x}$ 的分布；输出有 n 个符号 $\boldsymbol{y} = \{y_1,\cdots,y_n\}$，$Q(\boldsymbol{y}) = (q_1,\cdots,q_n)$表示 $\boldsymbol{y}$ 的分布；转移矩阵为 $T = (p(y_j|x_i))_{m\times n} = (p_{ij}))_{m\times n}$。于是有 $q_j = p_1p_{1j} + \cdots + p_mp_{mj}$，$j = 1,\cdots,n$。

$\boldsymbol{x}$，$\boldsymbol{y}$ 的互信息可以表示为如下的形式

$$I(x;y) = \sum_{i,j} p_i p(y_j \mid x_i)\text{lb}\frac{p(y_j \mid x_i)}{q_j}$$

记

$$c_i = -\sum_{j=1}^{n} p(y_j \mid x_i)\text{lb}\, p(y_j \mid x_i)$$

带约束的信道容量问题可表示为

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} c_i p_i + \sum_{j=1}^{n} q_j \text{lb} q_j \\ \text{s.t} \quad & q_j = \sum_{i=1}^{m} p_i p(y_j \mid x_i), \quad j = 1, \cdots, n \\ & p_1 + \cdots + p_m = 1 \\ & f_k(p_1, \cdots, p_m) \leqslant E_k, \ k = 1, \cdots, r \\ & p_1 > 0, \cdots, p_m > 0 \end{aligned} \tag{9}$$

这里 f_k 是线性函数或正定二次函数。这个凸优化问题的目标函数并不是自协调的。即使加上对数障碍函数 $-\text{lb}p_i$ 后，仍然不是自协调的。当所有的 $p_i > 0$ 时，必然所有的 $q_j > 0$，所以约束 $q_1 > 0,\cdots,q_n > 0$ 是冗余的。不过，加上 $-\text{lb}q_j$ 这些障碍项之后，目标函数就是自协调的了。为了说明这一点，需要证明，对任何 $t > 0$，$f(x) = tx\text{lb}x - \text{lb}x$ 在 $x > 0$ 时是自协调的。事实上，

$$\begin{aligned} & f'(x) = t\,\text{lb}\,x + t - 1/x, \\ & f''(x) = t/x + 1/x^2, \\ & |f'''(x)| = t/x^2 + 2/x^3 \end{aligned}$$

由此得到：

$$|f'''(x)|/f''(x)^{3/2} = (tx+2)/(tx+1)^{3/2} \leqslant 2$$

所以式(9)的目标函数是自协调的。利用对数障碍函数的方法，可以利用路径跟踪算法来求解信道容量问题式(9)。具体说来，就是要解下列问题

$$\min t\sum_{i=1}^{m} c_i p_i + t\sum_{j=1}^{n}\sum_{i=1}^{m} p_i p(y_j \mid x_i)\text{lb}\sum_{i=1}^{m} p_i p(y_j \mid x_i)$$

$$-\sum_{i=1}^{m}\text{lb}p_i - \sum_{j=1}^{n}\text{lb}\sum_{i=1}^{m} p_i p(y_j \mid x_i)$$

$$-\sum_{k=1}^{r}\mathrm{lb}(E_k - f_k(p_1,\cdots,p_m))$$
$$\text{s.t}\quad p_1+\cdots+p_m=1$$

计算量可以用式(8)形式的界给出。

当 $F(x)$是信道容量或熵函数[20]，$G(x)$是线性或正定二次函数[21]时，$C(\beta)=\sup\{F(x)|\,G(x)\leqslant\beta\}$是问题(9)的特殊形式，可以在自协调函数的框架下，在式(8)计算量内，求出符合要求的解。这样对于算法 1 的计算量，可以写出如下的界：

$$O(1)(M+m^3)(m+n)^{1/2}(\mathrm{lb}(V/\varepsilon))^2$$

在不存在一般求解公式的情况下[1]，具有多项式时间性的迭代算法就很好。

4　例子

设信道的转移矩阵 $\boldsymbol{P}$ 如下（10×10）：

$$\boldsymbol{P}=\begin{pmatrix}
0.90 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.02\\
0 & 1.00 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0.01 & 0.02 & 0.01 & 0.90 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01\\
0.50 & 0 & 0.50 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0.10 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10\\
0 & 0.10 & 0.10 & 0 & 0.10 & 0.10 & 0 & 0.10 & 0.50 & 0\\
0 & 0 & 0 & 0 & 0 & 0.40 & 0.10 & 0.10 & 0 & 0.40\\
0.10 & 0.20 & 0.30 & 0.40 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0.20 & 0 & 0.20 & 0 & 0.20 & 0.20 & 0 & 0.20\\
0 & 0.10 & 0.50 & 0 & 0.20 & 0 & 0.10 & 0.10 & 0 & 0
\end{pmatrix}$$

计算中采用自然对数。$\varepsilon=0.000\,01$。初始为等概率分布，$\mu=2$，$t_0=10$，牛顿迭代的步长因子取为 $1/(1+\lambda(f,\ \boldsymbol{x}))$。用路径跟踪算法迭代 70 次，信道容量的计算结果如下：$C=1.370\,656$，最优分布为

0.089 351 8，0.241 440 0，0.174 353 0，0.154 848 3，0.000 000 3，0.081 175 4，0.194 623 4，0.000 000 4，0.064 204 2，0.000 008 0

用 Arimoto-Blahut 交替算法做了计算，共迭代了 180 次。

取成本向量为 $\boldsymbol{b}=(1,2,5,7,6,8,3,9,3,4)^{\mathrm{T}}$，取 $G(\boldsymbol{x})$为期望成本，则 $C(\beta)/\beta$ 的曲线如图 1 所示。其中，横坐标是 β，纵坐标是 $C(\beta)/\beta$。用 0.618 法求解的单位成本容量为 0.580 1，对应的 $\beta=1.832\,5$，起始搜索区间取为[1, 5.8]。迭代次数为 17（路径跟踪次数）×19（0.618 次数）×5（牛顿法次数）+19×19（初始牛顿次数）=1 976 次。固定的 0.618 迭代次数增加了总的计算量。

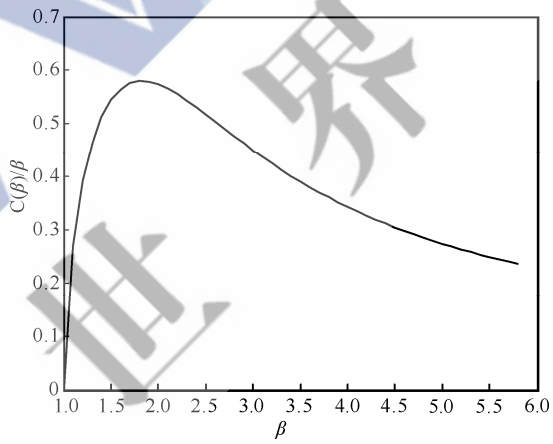

图 1　容量成本函数与成本比的曲线

取 $G(\boldsymbol{x})=\boldsymbol{b}'\boldsymbol{x}+\boldsymbol{x}'\boldsymbol{D}\boldsymbol{x}$ 为正定二次函数，其中，

$$\boldsymbol{D}=\begin{pmatrix}
4.279\,93 & 3.639\,0 & 3.329\,9 & 3.150\,5 & 2.593\,9 & 3.297\,7 & 3.177\,7 & 3.013\,1 & 3.220\,7 & 4.710\,7\\
3.639\,0 & 4.453\,6 & 3.688\,5 & 3.664\,4 & 2.256\,6 & 3.772\,5 & 3.391\,2 & 3.161\,5 & 2.163\,6 & 3.731\,1\\
3.329\,9 & 3.688\,5 & 4.843\,6 & 2.885\,2 & 2.960\,7 & 3.999\,4 & 1.950\,9 & 2.659\,9 & 2.767\,4 & 4.199\,0\\
3.150\,5 & 3.664\,4 & 2.885\,2 & 4.100\,4 & 2.277\,3 & 2.892\,0 & 3.845\,4 & 2.610\,6 & 2.678\,6 & 3.301\,3\\
2.593\,9 & 2.256\,6 & 2.960\,7 & 2.277\,3 & 3.636\,2 & 2.276\,4 & 2.049\,4 & 1.385\,8 & 2.498\,4 & 3.085\,5\\
3.297\,7 & 3.772\,5 & 3.999\,4 & 2.892\,0 & 2.276\,4 & 5.415\,8 & 2.924\,1 & 3.211\,3 & 3.058\,0 & 3.086\,0\\
3.177\,7 & 3.391\,2 & 1.950\,9 & 3.845\,4 & 2.049\,4 & 2.924\,1 & 6.029\,2 & 3.867\,9 & 3.231\,7 & 3.458\,2\\
3.013\,1 & 3.161\,5 & 2.659\,9 & 2.610\,6 & 1.385\,8 & 3.211\,3 & 3.867\,9 & 4.928\,0 & 3.420\,4 & 3.523\,1\\
3.220\,7 & 2.163\,6 & 2.767\,4 & 2.678\,6 & 2.498\,4 & 3.058\,0 & 3.231\,7 & 3.420\,4 & 5.185\,6 & 3.654\,3\\
4.710\,7 & 3.731\,1 & 4.199\,0 & 3.301\,3 & 3.085\,5 & 3.086\,0 & 3.458\,2 & 3.523\,1 & 3.654\,3 & 6.595\,3
\end{pmatrix}$$

$C(\beta)/\beta$ 的曲线如图 2 所示。其中，横坐标是 β，纵坐标是 $C(\beta)/\beta$。用 0.618 法求广义单位成本容量的值等于 0.2156，相应的 β= 5.9093，起始搜索区间为[5.28, 12]。迭代次数为 17（路径跟踪次数）×19（0.618 次数）×4（牛顿法次数）+19×171 =4 541 次。这个例子的初始牛顿迭代次数太多。

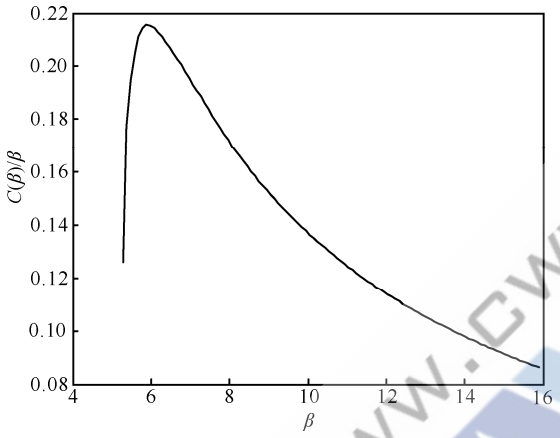

图 2　约束为正定二次函数时，广义容量成本函数与成本比的曲线

5　结束语

单位成本信道容量的表达式对一般凸函数也是成立的；同时 $C(\beta)/\beta$ 具有单峰性，这为设计新的算法提供了条件。利用自协调函数理论，使得计算带约束的信道容量问题变得更加简单，尤其是当约束是非线性的时候，这也包括对 $C(\beta)$的计算问题。

信息论中的一些新问题，例如信息隐藏问题[12]，其特点是优化目标不仅仅与信号的分布有关，而且涉及转移概率的选择，使得目标函数的自协调性问题变得更加复杂，需要进一步研究。另外，删除信道的信道容量的计算问题[9]，也是人们关注的热点之一。

参考文献：

[1] LIANG X B. An algebraic, analytic, and algorithmic investigation on the capacity and capacity-achieving input probability distributions of finite-input-finite-output discrete memoryless channels[J]. IEEE Trans Inf Theory, 2008, 54(3): 1003-1023.

[2] CHIANG M, BOYD S. Geometric programming duals of channel capacity and rate distortion[J]. IEEE Trans Inf Theory, 2004, 50(2): 245-258.

[3] 管宇. 离散信道容量的迭代算法[J]. 应用数学与计算数学学报, 2006, 20(2): 19-27.
GUAN Y. Some alternate algorithms for computing the capacity of arbitrary discrete memoryless channels[J]. Comm on Appl Math and Comput, 2006, 20(2): 19-27.

[4] MCELIECE R J.信息论与编码理论（第二版）[M]. 北京: 电子工业出版社, 2004.
MCELIECE R J. The Theory of Information and Coding (Second Edition)[M]. Beijing: Publishing House of Electronics Industry, 2004.

[5] ALAJAJI F, WHALEN N. The capacity-cost function of discrete additive noise channels with and without feedback[C]. IEEE Trans on Inform Theory, 2000, 46(3): 1131-1140.

[6] KHAYRALLAH A S, NEUHOFF D L. Coding for channels with cost constraints[J]. IEEE Trans on Inform Theory, 1996, 42(3): 854-867.

[7] VERDU S. On channel capacity per unit cost[J]. IEEE Trans on Inform Theory, 1990, 36(5): 1019-1030.

[8] VARSHNEY L. Variations on channel capacity per unit cost[EB/OL]. http://web.mit.edu/lrv/www/writing/cap_cost .pdf , 2005

[9] MITZENMACHER M. A survey of results for deletion channels and related synchronization channels[J]. Probability Surveys, 2009, 6: 1-33.

[10] KLEINER M, RIMOLDI B. On fidelity per unit cost[A]. Proceedings of the 2009 IEEE International Conference on Symposium on Information Theory[C]. Seoul, Korea,2009. 1639-1643.

[11] OBIANUJU N, TOKUNBO O. Achieving maximum possible speed on constrained block transmission systems[J]. EURASIP Journal on Advances in Signal Processing, 2007:1-11.

[12] 叶天语, 钮心忻, 杨义先. 被动攻击下的信息隐藏系统性能分析[J]. 北京邮电大学学报, 2008, 31(3):46-49.
YE T Y, NIU X X, YANG Y X. Analysis on the performances of information hiding system under passive attack[J]. Journal of Beijing University of Posts and Telecommunications, 2008, 31(3): 46-49.

[13] 杨盘隆, 陈贵海. 无线网状网容量分析与优化理论研究[J]. 软件学报, 2008, 19(1): 111-125.
YANG P L, CHEN G H. Research paradigm of capacity analysis and optimizing theory on wireless mesh network[J]. Journal of Software, 2008,19(1):111-125.

[14] WRIGHT M H. The interior-point revolution in optimization: history, recent developments, and lasting consequences[J]. Bull Amer Math Soc, 2005, 42: 39-56.

[15] KOPOTUN K A, LEVIATAN D, SHEVCHUK I A. Convex polynomial approximation in the uniform norm: conclusion[J]. Canad J Math, 2005, 57(6): 1224-1248.

[16] 袁亚湘, 孙文瑜. 最优化理论与方法[M]. 北京: 科学出版社, 1997.
YUAN Y X, SUN W Y. Optimization Theory and Methods[M]. Beijing: Science Press, 1997.

[17] NESTEROV Y, NEMIROVSKII A. Interior-Point Polynomial Methods in Convex Programming[M]. Philadelphia: Society for Industrial and Applied Mathematics, 1994.

（下转第 46 页）