

逼近精确罚函数法求解单阶段随机规划

潘青飞^{1,2}, 王效俐¹

(1. 同济大学 经济与管理学院, 上海 200092; 2. 三明学院 物理与机电工程系, 福建 三明 365004)

摘要: 提出了一种求解单阶段随机规划的算法——逼近精确罚函数法. 首先, 通过离散化随机变量的方法得到逼近原问题的确定非线性规划序列, 然后, 建立精确罚函数并构造无约束最优化问题. 在一定的条件下, 证明了确定非线性规划序列与无约束最优化问题的等价性, 同时也证明了离散序化的解序列收敛到原规划的解.

关键词: 单阶段随机规划; 离散化; 精确罚函数; 收敛

中图分类号: O 221.5

文献标识码: A

An Approximation-exact Penalty Function Method of Solving Single Stage Stochastic Programming

PAN Qingfei^{1,2}, WANG Xiaoli¹

(1. College of Economics and Management, Tongji University, Shanghai 200092, China; 2. Department of Physics & Mechanics-Electronic Engineering, Sanming University, Sanming 365004, China)

Abstract: An approximation-exact penalty function method for solving single stage stochastic programming is presented. Firstly, the determinate nonlinear programming sequences are obtained by means of discretizing random variable. Secondly, an exact penalty function and an unconstrained optimization are constructed correspondingly. Under lenient conditions, some equivalent properties between the determinate nonlinear programming and the unconstrained optimization are proved, and the solution sequence of the determinate nonlinear programming converges to solution to the original problem in some sense.

Key words: single stage stochastic programming; discretization; exact penalty function; convergence

在某种意义上, 多阶段随机规划通常可以转化为单阶段随机规划^[1], 如文献[2]给出了 2 阶段凸随

机规划的转换. 因此, 研究单阶段随机规划有关理论与算法是很有意义的.

考虑以下单阶段随机规划(RP):

$$\begin{aligned} \min & f(x) \\ \text{s. t.} & g_i(x, \xi) \leq 0 (i = 1, \dots, m) \quad x \in \mathbf{R}^n \end{aligned}$$

式中: $f(x) = \int_S g_0(x, \xi) P(d\xi)$, 其中, ξ 为定义在概率空间 (S, \sum, P) 上的连续型随机变量(向量), $S \subset \mathbf{R}^r$, P 为概率; g_i 为函数.

规划(RP)解的理论问题人们已经有所研究^[1,3]. Flam 和 Zowe^[4]首先讨论了单阶段凸随机规划精确罚函数的某些条件. 对于这些条件, 即使 ξ 是连续型随机变量(向量), 用已有的精确罚函数最优化算法, 这类随机规划通常也是不能求解的. 从目标函数和罚函数的构造看, 对应梯度的计算是很复杂的. 因此, 有必要寻找其他解决的方法, 其中就有离散化逼近方法. 近来, 用近似技术解决随机规划已成为重要的方法^[2,5]. 许多学者已就这个问题设计了一些有效的算法, 如最近出现的最大熵方法^[6].

近二三十年, 很多学者研究用精确惩罚法, 尤其是外部惩罚法, 求解非线性规划^[4]. 为此, 笔者结合精确罚函数和逼近的方法, 建立新算法——逼近精确罚函数法, 用它求解规划(RP), 可以克服一般算法的常见困难.

1 逼近精确罚函数的构造

1.1 预备

设 S 是有界连通闭集, ξ 是 S 上的连续型随机变量, 其密度函数为 $\Psi(t)$. 用不同方法划分 S , 其中第 k 种划分 $A_{j,k} (j = 1, \dots, N_k, k = 1, 2, \dots)$ 满足条件:

$$\textcircled{1} P(A_{j,k}) > 0, P(A_{j,k}) = \int_{A_{j,k}} \Psi(t) dt; \textcircled{2} A_{i,k} \cap$$

收稿日期: 2010-01-07

基金项目: 国家自然科学基金资助项目(7084002); 上海市重点学科建设资助项目(B310)

作者简介: 潘青飞(1962—), 男, 副教授, 理学学士, 主要研究方向为优化理论与方法. E-mail: pqf101@yahoo.com.cn

王效俐(1960—), 男, 教授, 博士生导师, 管理学博士, 主要研究方向为管理科学及运营管理. E-mail: xiaoli-wang@tongji.edu.cn

$A_{j,k} = \phi(i \neq j), \bigcup_{j=1}^{N_k} A_{j,k} = S; \textcircled{3} P(\partial A_{j,k}) = 0,$
 $\text{diam} A_{j,k} = \sup_{s,t \in A_{j,k}} \|s - t\| \rightarrow 0 (N_k \rightarrow \infty); \textcircled{4} t_{j,k} \in$
 $A_{j,k}, \max_j \frac{p_{j,k}}{P(A_{j,k})} \rightarrow 1 (N_k \rightarrow \infty).$ 这里 ∂A 指 A 的边
 界, $p_{j,k} = P(\xi = t_{j,k})$, 并假设数列 $N_k \rightarrow \infty (k \rightarrow \infty)$.

命题 1^[2]

$$\sum_{j=1}^{N_k} h(t_{j,k}) p_{j,k} \rightarrow \int_{A_{j,k}} h(\xi) P(d\xi), k \rightarrow \infty \quad (1)$$

成立的充要条件是分划满足条件①~④, 其中 $h \in C(S)$.

令 $\Delta V_{j,k}$ 记 $A_{j,k}$ 的“体积”, 在 $A_{j,k}$ 中选择点 $t_{j,k}$, 并定义 ξ^k

$$p_{j,k} = P(\xi_k = t_{j,k}) = \Psi(t_{j,k}) \Delta V_{j,k} \quad (2)$$

ξ_k 可能不是随机变量, 因为 $\sum_{j=1}^{N_k} \Psi(t_{j,k}) \Delta V_{j,k}$ 不一定等于 1.

命题 2 对于由式(2)定义的 ξ_k , 满足条件①~④, 即也满足式(1).

证明 要保证假设 ① ~ ④ 成立, 只要证明

$$\max_j \frac{p_{j,k}}{P(A_{j,k})} \rightarrow 1 (k \rightarrow \infty), \text{ 而 } |p_{j,k} - P(A_{j,k})| =$$

$$\left| \Psi(t_{j,k}) \Delta V_{j,k} - \int_{A_{j,k}} \Psi(t) dt \right| \leq |\Psi(t_{j,k}) - \Psi(\bar{t}_{j,k})| \cdot$$

$$\Delta V_{j,k} \rightarrow 0 (k \rightarrow \infty), \text{ 其中 } \bar{t}_{j,k} \in A_{j,k}, \text{ 即有}$$

$$\max_j \frac{p_{j,k}}{P(A_{j,k})} \rightarrow 1 (k \rightarrow \infty).$$

对由式(2)定义的 ξ_k , 定义其分布函数

$$G_k(t) = \sum_{t_{j,k} \leq t} \Psi(t_{j,k}) \Delta V_{j,k}$$

由于 S 是有界连通集, 密度函数 Ψ 连续, 假如 $\max_j \Delta V_{j,k} \rightarrow 0 (k \rightarrow \infty)$, 则由定积分的定义, 对任何 $t \in S$ 有

$$\lim_{k \rightarrow \infty} G_k(t) = G(t) \quad (3)$$

其中 $G(t) = \int_{u \leq t} \Psi(u) du$ 是分布函数.

因此, 可以称由式(2)定义的 ξ_k 为广义随机向量, 可用它方便地进行近似计算.

1.2 构造

假设式(1)成立, 对规划(RP)离散化得下列离散近似数学规划(DP):

$$\min \sum_{j=1}^{N_k} g_0(x, t_{j,k}) p_{j,k}$$

$$\text{s. t. } g_i(x, t_{j,k}) \leq 0, j = 1, \dots, N_k,$$

$$i = 1, \dots, m, x \in \mathbf{R}^n$$

显然, 规划(DP)的约束条件数量将随着 k 的增加而迅速增加, 这将为求解序列规划(DP)带来一些不便, 因此有必要使序列规划(DP)的约束条件数量保持不变.

利用精确罚函数^[8]由规划(DP)构造无约束优化问题(UP)

$$\min \rho_k(x, u) \quad x \in \mathbf{R}^n$$

式中:

$$u > 0$$

$$\rho_k(x, u) = u \sum_{j=0}^{N_k} g_0(x, t_{j,k}) p_{j,k} +$$

$$\sum_{i=1}^m \sum_{j=1}^{N_k} [g_i(x, t_{j,k}) + |g_i(x, t_{j,k})|] =$$

$$u f_k(x) + \sum_{i=1}^m \sum_{j=1}^{N_k} [g_i(x, t_{j,k}) + |g_i(x, t_{j,k})|]$$

$$f_k(x) = \sum_{j=1}^{N_k} g_0(x, t_{j,k}) p_{j,k}$$

求解规划(UP)有很多有效的算法^[7].

2 收敛

2.1 假设条件

为了研究逼近规划的收敛, 要引进一些必要的假设, 并分析规划(DP)和(UP)的关系. 假设: ⑤对 S 中几乎处处的 ξ , 可行集 $D = \{x \in \mathbf{R}^n \mid g_i(x, \xi) \leq 0, i = 1, \dots, m\}$ 是非空有界闭凸的, S 是有界连通集; ⑥函数 $g_i(\cdot, \cdot) (i = 0, \dots, m)$ 关于第 1 个变量凸可微, 关于第 2 个变量连续可测; ⑦对任意 $x \in \mathbf{R}^n, i = 1, \dots, m$, 存在有界可测函数 $\delta_i(\xi)$ 满足 $\|\nabla_x g_i(x, \xi)\| \leq \delta_i(\xi)$; ⑧如果 x 是规划(DP)不可行点, 则存在 $\sigma > 1$ 使 $\sigma g_i(x, t_{j,k}) \geq \sum_{t_{j,k} \in I_{i,k}} g_i(x, t_{j,k})$, 其中, $I_{i,k} = \{t_{j,k} \mid g_i(\bar{x}_k, t_{j,k}) = 0, j = 1, \dots, N_k\}$, \bar{x}_k 是规划(DP)的最优解.

假设条件⑤和⑥是确保规划(RP)存在最优解、 ρ_k 为关于变量 x 的连续凸函数的条件; 假设条件⑦是保证离散化逼近规划(UP)收敛的必要条件; 假设条件⑧描述了离散化逼近规划在最优解的渐近性质. 为确保指标集 $I_{i,k}$ 有限, 假设: ⑨梯度 $\nabla_x g_i(\bar{x}_k, t_{j,k}) (i = 0, \dots, m)$ 对 ξ 几乎处处线性无关, 因此, 为不失一般性, 记 $I_{i,k} = \{t_{j,k} \mid g_i(\bar{x}_k, t_{j,k}) = 0,$

$1 \leq \Gamma \leq k_j, j = 1, \dots, N_k$, 显然, 有 $k_j \leq N_k$. 如果 \bar{x}_k 是规划 (DP) 的最优解, 其 Kuhn-Tucker 条件是

$$\sum_{j=1}^{N_k} \nabla_x g_0(\bar{x}_k, t_{j,k}) p_{j,k} + \sum_{i=1}^m \sum_{j=1}^{N_k} \sum_{t=1}^{k_j} \lambda(t_{j_r,k}) \nabla_x \cdot g_i(\bar{x}_k, t_{j_r,k}) = 0, \text{ 其中 } \lambda(t_{j_r,k}) \geq 0 \text{ 是最优解 } \bar{x}_k \text{ 的拉格朗日乘子.}$$

2.2 规划的等价性

定理 1 设假设条件⑤~⑨成立, \bar{x}_k 规划 (DP) 的最优解, 则存在 $u^* > 0$, 使对 $0 \leq u \leq u^*$, \bar{x}_k 也是规划 (UP) 的最优解.

证明 如果 $x \in \mathbf{R}^n$ 是 (DP) 可行点, 则 $\rho_k(x, u) = uf_k(x) \geq uf_k(\bar{x}_k) = \rho_k(\bar{x}_k, u)$

如 $x \in \mathbf{R}^n$ 不是 (DP) 可行点, 则存在 i, j , 使 $g_i(x, t_{j,k}) > 0$.

$$\begin{aligned} f_k(x) - f_k(\bar{x}_k) &\geq \nabla f_k(\bar{x}_k)^T (x - \bar{x}_k) = \\ &= - \sum_{i=1}^m \sum_{j=1}^{N_k} \sum_{t=1}^{k_j} \lambda(t_{j_r,k}) \nabla_x g_i(\bar{x}_k, t_{j_r,k})^T (x - \bar{x}_k) \geq \\ &= \sum_{i=1}^m \sum_{j=1}^{N_k} \sum_{t=1}^{k_j} \lambda(t_{j_r,k}) [-g_i(x, t_{j_r,k}) + g_i(\bar{x}_k, t_{j_r,k})] \geq \\ &= \sum_{i=1}^m \sum_{j=1}^{N_k} \sum_{t=1}^{k_j} \lambda(t_{j_r,k}) [-g_i(x, t_{j_r,k})] \end{aligned}$$

分别考虑如下 2 种情况:

情况 1 $g_i(x, t_{j_r,k}) \leq 0, \rho_k(x, u) - \rho_k(\bar{x}_k, u) =$

$$uf_k(x) + \sum_{i=1}^m \sum_{j=1}^{N_k} [g_i(x, t_{j_r,k}) + |g_i(x, t_{j_r,k})|] -$$

$$uf_k(\bar{x}_k) \geq \sum_{i=1}^m \sum_{j=1}^{N_k} \sum_{t=1}^{k_j} u \lambda(t_{j_r,k}) [-g_i(x, t_{j_r,k})] \geq$$

0. 即对任意 $u \geq 0, \rho_k(x, u) \geq \rho_k(\bar{x}_k, u)$.

情况 2 对某些 $r, g_i(x, t_{j_r,k}) > 0$. 由 $k_j = 0$, 得

$$\sum_{i=1}^m \sum_{j=1}^{N_k} \lambda(t_{j_r,k}) [-g_i(x, t_{j_r,k})], \text{ 对于 } \rho_k(x, u) - \rho_k(\bar{x}_k, u) \text{ 的大小没有贡献.}$$

因此, 为不失一般性, 假设对任意 j 有 $k_j \geq 1$, 则

$$\begin{aligned} \rho_k(x, u) - \rho_k(\bar{x}_k, u) &\geq \sum_{i=1}^m \sum_{j=1}^{N_k} [g_i(x, t_{j_r,k}) + |g_i(x, t_{j_r,k})|] - \sum_{i=1}^m \sum_{j=1}^{N_k} \sum_{t=1}^{k_j} u \lambda(t_{j_r,k}) [g_i(x, t_{j_r,k})] \geq \\ &= \sum_{i=1}^m \sum_{j=1}^{N_k} 2g_i(x, t_{j_r,k}) - \sum_{i=1}^m \sum_{j=1}^{N_k} \sum_{t=1}^{k_j} u \lambda(t_{j_r,k}) [g_i(x, t_{j_r,k})] \geq \\ &= \sum_{i=1}^m \sum_{j=1}^{N_k} \sum_{t=1}^{k_j} [2\sigma^{-1} - u \lambda(t_{j_r,k})] g_i(x, t_{j_r,k}). \end{aligned}$$

令 $2\sigma^{-1} - u \lambda(t_{j_r,k}) \geq 0$, 则取 $0 < u < u^* \leq 2/[\sigma \max_{1 \leq r \leq k_j} \lambda(t_{j_r,k})]$, 即对任意 j , 有 $\rho_k(x, u) \geq$

$\rho_k(\bar{x}_k, u)$. 如果 $\lambda(t_{j_r,k}) = 0$, 则 $2\sigma^{-1} - u \lambda(t_{j_r,k}) \geq 0$ 恒成立.

定理 2 设假设条件①~⑤成立, 且 $x_k(u)$ 是 (UP) 的全局解, 则对充分小的 u 使 $x_k(u) = \bar{x}_k$ 也是 (DP) 的最优解.

证明 由 $\rho_k(x, u) = uf_k(x) + \sum_{i=1}^m \sum_{j=1}^{N_k} [g_i(x, t_{j_r,k}) + |g_i(x, t_{j_r,k})|] = uf_k(x) + \sum_{i=1}^m \sum_{j=1}^{N_k} \max\{0, 2g_i(x, t_{j_r,k})\} \geq uf_k(x) + \sum_{i=1}^m \sum_{j=1}^{N_k} \sum_{t=1}^{k_j} \sigma^{-1} \max\{0, 2g_i(x, t_{j_r,k})\} = \bar{\rho}_k(x, u)$.

显然, $\bar{\rho}_k(x, u)$ 是下列非线性规划的 l_1 精确处罚函数 (P1):

$$\begin{aligned} \min \quad & f_k(x) \\ \text{s. t.} \quad & 2\sigma^{-1} g_i(x, t_{j_r,k}) \leq 0 \quad x \in \mathbf{R}^n, \\ & r = 1, \dots, k_j, j = 1, \dots, N_k, i = 1, \dots, m \end{aligned}$$

对 $x \in \mathbf{R}^n$ 和充分小的 u , 上述规划 (P1) 的全局解 \bar{x}_k 也是 $\bar{\rho}_k(x, u)$ 的全局解^[8], 因而, 对充分小的 u 有 $uf_k(\bar{x}_k) = \bar{\rho}_k(\bar{x}_k, u) \leq \bar{\rho}_k(x, u) \leq \rho_k(x, u)$. 特别是, 当 $uf_k(\bar{x}_k) \leq \rho_k(x_k(u), u)$, 可知 $x_k(u)$ 是 (UP) 的全局解, 则有 $\rho_k(x_k(u), u) \leq \rho_k(\bar{x}_k, u) = uf_k(\bar{x}_k)$. 于是对充分小的 u 有 $\rho_k(\bar{x}_k, u) = \bar{\rho}_k(\bar{x}_k) = \rho_k(x_k(u), u)$, 即 $x_k(u) = \bar{x}_k$. 由指标 I_k 的定义知规划 (P1) 等价规划 (DP).

2.3 收敛性

在研究随机规划逼近中, epi(epigraph) 收敛起着重要的作用. 下面主要探讨规划 (DP) 的解序列如何收敛于规划 (RP) 的解. 用集合 M 和 M_k 分别定义规划 (RP) 和 (DP) 的可行解集 (非空). 同文献 [2] 命题 3.1, 类似可得如下命题:

命题 3 设假设条件⑤~⑨和式 (1) 成立, 则约束集 M 包含序列 $\{x_k\}$ 所有聚点, $x_k \in M_{k_1}$, 其中 $\{M_{k_1}\}$ 是 $\{M_k\}$ 的任意子序列.

命题 4 设假设条件⑤~⑦和式 (1) 成立, 则函数序列 $\{f_k\}$ epi 收敛于 f , 记为 $\lim_{k \rightarrow \infty} f_k \underline{e} f$.

用 \bar{f} 和 \bar{f}_k 分别记问题 (RP) 和 (DP) 的最优值, x_k 是规划 (DP) 的最优解.

定理 3 设假设条件⑤~⑨和式 (1) 成立, 则 $\lim_{k \rightarrow \infty} \bar{f}_k \underline{e} f$, 且 $\{x_k\}$ 存在收敛的子序列 $\{x_{k_1}\}$, 其聚点是规划 (RP) 的最优解.

证明 为不失一般性, 记 \bar{x} 是序列 $\{x_k\}$ 的 1 个

聚点.由命题3即可得 $\bar{f} \leq f(x) \leq \liminf_{k \rightarrow \infty} f_k(x_k)$.

由假设条件⑤得 $\bar{x} \in \mathbf{R}^n$. 下面证明 $\limsup_{k \rightarrow \infty} \bar{f}_k \leq \bar{f}$. 由假设,存在 $\varepsilon > 0$, $\hat{x} \in \mathbf{R}^n$,使得对几乎每处 $\xi \in S$,有 $g_i(\hat{x}, \xi) \leq -\varepsilon < 0 (i = 1, \dots, m)$,则存在 $v \in [0, 1]$,使 $y_v = v\hat{x} + (1-v)\bar{x}$ 是规划(RP)的可行点,并有 $|f(y_v) - f(\bar{x})| \leq |vf(\hat{x}) + (1-v)f(\bar{x}) - f(\bar{x})| = v|f(\hat{x}) - f(\bar{x})|$. 取充分小的 v 使 $|f(y_v) - f(\bar{x})| \leq \varepsilon/2$. 对固定的 $v = v(\varepsilon)$,存在有自然数 K_1 ,使对任意 $k > K_1$ 有 $g_i(y_v, t_{j,k}) \leq v \cdot g_i(\hat{x}, t_{j,k}) + (1-v)g_i(\bar{x}, t_{j,k}) - vg_i(\hat{x}, \xi) + vg_i(\hat{x}, \xi) - (1-v)g_i(\hat{x}, \xi) + (1-v)g_i(\hat{x}, \xi) = v[g_i(\hat{x}, t_{j,k}) - g_i(\hat{x}, \xi)] + (1-v)[g_i(\hat{x}, t_{j,k}) - g_i(\bar{x}, \xi)] + vg_i(\hat{x}, \xi) + (1-v)g_i(\hat{x}, \xi) \leq 0$. 对 $k > K_1$,有 $y_v \in M_k$. 由命题3和假设⑦,使对几乎每处 $\xi \in S$,有 $g_i(\bar{x}, \xi) \leq 0, g_i(\hat{x}, \xi) \leq -\varepsilon < 0 (i = 1, \dots, m)$. 由假设⑥,对 $\forall i \in \{1, \dots, m\}$ 有 $\max_{1 \leq j \leq N_k} |g_i(x, t_{j,k}) - g_i(x, \xi)| \rightarrow 0 (k \rightarrow \infty)$,故 $g_i(x, t_{j,k}) \leq 0$. 再由式(1),存在自然数 K_2 ,使 $k > K_2$ 有 $|f_k(y_v) - f(y_v)| \leq \varepsilon/2$,取 $k > \max\{K_1, K_2\}$,则 $f_k(x_k) - f(\bar{x}) \leq f_k(y_v) - f(\bar{x}) \leq |f_k(y_v) - f(y_v)| + |f(y_v) - f(\bar{x})| \leq \varepsilon$. 由 x_k 是(DP)的最优解得 $f_k(x_k) \leq f_k(y_v)$,从而 $f_k(x_k) \leq f(\bar{x}) + \varepsilon$. 所以 $\limsup_{k \rightarrow \infty} \bar{f}_k \leq \bar{f}$,得 $\lim_{k \rightarrow \infty} \bar{f}_k = \bar{f}$.

由假设⑤和⑥,在序列 $\{x_k\}$ 中必有收敛子序列,其聚点 x 包含在可行集 M 中, \bar{x} 就是规划(RP)的最优解.

3 结论

为了解决随机变量为连续型的单阶段随机规划

问题,借助于离散化的思想,其近似解可以通过求解无约束规划得到,这种算法可克服由约束条件数量大量增加而产生的诸如计算精度等困难,而且这种算法是很灵活的求近似解的方法,可用于求解一般的随机变量为连续型的单阶段随机规划问题.

这里仅仅提供了方法的一般性描述和理论上的证明,可为进一步的研究作准备.并且一些具体问题仍有待解决,例如,在解的逼近过程中如何确定 k 和 u 的关系、什么是终止条件、必要的数值试验,等等.

参考文献:

- [1] 陈志平.一般形式多阶段有补偿问题的理论和算法[D].西安:西安交通大学数学系,1992.
CHEN Zhiping. General form of multi-stage compensation theory and algorithms [D]. Xi'an: Xi'an Jiaotong University. Department of Mathematics, 1992.
- [2] Lepp R. Approximations to stochastic programs with complete recourse[J]. SIAM Journal of Control and Optima, 1990, 28: 382.
- [3] 王金德.随机规划[M].南京:南京大学出版社,1990.
WANG Jinde. Stochastic programming [M]. Nanjing: Nanjing University Press, 1990.
- [4] Flam S D, Zowe J. Exact penalty functions in single-stage stochastic programming[J]. Optimization, 1990, 2: 723.
- [5] Birge R J, Wets R J B. Designing approximation schemes for stochastic optimization problems in particular for stochastic programs with recourse[J]. Mathematical Programming Study, 1986, 27: 54.
- [6] 万仲平,陈开周.求解单阶段随机规划的一种光滑逼近法[J].数学研究与评论,1997,17(4): 565.
WAN Zhongping, CHEN Kaizhou. A smooth approximating method for solving single stage stochastic programming [J]. Journal of Mathematical Research and Exposition, 1997, 17 (4): 565.
- [7] Clark F H. Optimization and nonsmooth analysis[M]. New York: John Wiley, 1983.
- [8] Fletcher R. Practical methods of optimization [M]. New York: John Wiley & Sons, 1987.