一种解带补偿的随机规划的逼近方法*

孙德锋　　王金德

(中科院应用数学研究所)　　(南京大学数学系)

AN APPROXIMATE METHOD FOR STOCHASTIC PROGRAMMING WITH RECOURSE

Sun De-feng　　Wang Jin-de

(*Institute of Applied Mathematics Academia Sinica*)　　(*Dept. of Mathematics, Nanjing University*)

Abstract

In this paper we give a method for solving special nonlinear programming by designing an inexact line search and combining the feasible direction method of Topkis—Veinott case and approximation theory in mathematical programming. In particular, our algorithm is suitable for solving stochastic programming with a complete recourse matrix. Moreover, a practical design is presented and numerical reslts are provided.

1. 引　　言

考虑非线性规划问题:

$$\begin{aligned} &\min \quad f(x), \\ &\text{s.t.} \quad Ax \le b, \end{aligned} \tag{1}$$

其中 $f(x) \in C^1$ 且 $f(x)$ 为凸函数, $A \in I\!R^{m\times n}, x \in I\!R^n, b \in I\!R^m$. (1) 的一般形式可用可行方向法 (Topkis-Veinott 情形) 得到一个 Fritz-John 点. 但当 $f(x)$ 或 $\nabla f(x)$ 太复杂以致难以计算时, 此方法就不适当. 为此考虑逼近问题:

$$\begin{aligned} &\min \quad f_k(x), \\ &\text{s.t.} \quad Ax \le b, \end{aligned} \tag{2}$$

其中 $f_k(x)$ 为连续的凸泛函 (但可能不可微) 且满足

$$\lim_{k\to+\infty} f_k(x^k) = f(x), \tag{3}$$

$$\lim_{k\to+\infty} v^k = \nabla f(x), \tag{4}$$

* 1992 年 10 月 27 日收到.

对任何 $\{x^k\} \to x, v^k \in \partial f_k(x^k)$.

定义 1(广义 Fritz-John 点). 如果 $u_0 \geq 0, u \geq 0$ 且 $(u_0, u) \neq 0$, 使 $0 \in \partial(u_0 f_k(x) + u^T(Ax-b))|_{x=\bar{x}}$, 则称 $\bar{x}$ 为 (2) 的广义 Fritz-John 点.

定义 2. 给定凸泛函 g, 可行点 x, 可行方向 d, 实数 $c_0 \in (0,1]$, $\lambda_{\max} \geq 0$. 设

$$g^* = \min\{g(x+\lambda d) | 0 \leq \lambda \leq \lambda_{\max}\},$$

定义

$$M_{c_0}(x,d,g) = \{y | y = x + \lambda d, 0 \leq \lambda \leq \lambda_{\max}, \text{且满足} g(y) \leq c_0 g^* + (1-c_0) g(x)\}.$$

注. 由于 $g(x)$ 为凸泛函，故 $M_{c_0}(x,d,g)$ 为非空凸集；当 $c_0 = 1$ 时，则相当于线搜索的精确求解.

在 [6] 中精确求解一维线搜索的条件下，提出了一种形式的逼近可行方向法. 实际上，一维线搜索的精确求解，特别对于随机规划来说，是相当麻烦的. 鉴于此，在 [6] 的基础上，利用定义 2 的提法，给出了一种带不精确线搜索的逼近方法，证明了收敛性. 在一些条件下，还给出了可行的计算方法.

算法 A. 步 0. 给定常数 $c_0 \in (0,1)$, 置 $k = 1$.

步 1. 设有可行点 x^k 及 $v^k \in \partial f_k(x^k)$, 解

$$\begin{aligned} &\min \quad z, \\ &\text{s.t} \quad {v^k}^T d - z \leq 0, \quad Ad - ze \leq b - Ax^k, \quad -1 \leq d_j \leq 1, \quad j = 1, \cdots, n, \end{aligned}$$

其中 $e = (1, \cdots, 1)^T$.

设 (z_k, d_k) 为最优解. 如果 $z_k = 0$, 由 [6] 知 x^k 是 (2) 的广义 Fritz-John 点. 进而如果 $f_k(x)$ 较好的逼近 $f(x)$, 那么 x^k 作为近似解. 否则转步 3.

步 2. 寻求 $x^{k+1} \in M_{c_0}(x^k, d_k, f_k)$, 即存在 λ_k 使得

$$x^{k+1} = x^k + \lambda_k d_k \tag{6}$$

(在此较温和的条件下，我们将证明这一步在有限次可以达到. $k := k+1$, 转步 1.

步 3. 令 $x^{k+1} = x^k$, 用 $f_{k+1}(x)$ 代替 $f_k(x)$, $k := k+1$, 转步 1.

可以看出，算法 A 的主要工作量集中在步 2. 我们将证明，实际上这一步也是容易实现的.

2. 收敛理论

定理 1[6]. 1) 如果 $z_k = 0$, 则 x^k 是 (2) 的广义 Fritz-John 点；

2) 如果 $z_k < 0$, 则 d_k 或者是 $f_k(x)$ 在 x^k 处的下降方向或者 x^k 是下列问题的最优解:

$$\begin{aligned} &\min \quad f_k(x^k + \lambda d_k), \\ &\text{s.t.} \quad A(x^k + \lambda d_k) \le b, \quad \lambda \ge 0. \end{aligned} \tag{7}$$

定理 2[6]. 假设 x^k 是由算法 A 得到的 (2) 的广义 Fritz-John 点, $k = 1, 2, \cdots$, 则 $\{x^k\}$ 的任何聚点 $\bar{x}$ 是问题 (1) 的 Fritz-John 点.

定义 3(ε- 下单调性). 称函数列 $\{f_k(x)\}$ 在集合 D 上为 ε- 下单调的, 如果存在非负序列 $\{\varepsilon_k\}, \varepsilon_k \ge 0$ 且 $\sum_{k=1}^{+\infty} \varepsilon_k < +\infty$, 使得 $f_{k+1}(x) \le f_k(x) + \varepsilon_k, \forall x \in D$ 都成立, $k = 1, 2, \cdots$.

注. 如果对所有的 k 都有 $\varepsilon_k = 0$, 则 $\{f_k(x)\}$ 即为通常的下降函数列, 从而单调下降函数列是 ε- 下单调的.

引理 3. 设 $f(x) \in C^1$ 且 $\{f_k(x)\}$ 为满足 (3) 和 (4) 的 ε- 下单调函数列, 则算法 A 不可能产生满足下列性质的子列 $\{x^{k_i}\}$:

a) $x^{k_i} \to \bar{x}, \quad d_{k_i} \to \bar{d}$;

b) $x^{k_i} + \lambda d_{k_i} \in D = \{x | Ax \le b\}, \quad \lambda \in (0, \delta]$, 其中 δ 是不依赖于 k_i 的正数;

c) $\nabla f(\bar{x})^T \bar{d} < 0$.

证明. 设算法 A 产生了这样的子列 $\{x^{k_i}\}$, 再设 $\nabla f(\bar{x})^T \bar{d} = -2\varepsilon < 0$. 由微分稳定性条件 (4), 存在一正常数 δ', 使得对所有 $\lambda \in (0, \delta^1]$, 当 i 充分大时有

$$v^{k_i}(\lambda)^T d_{k_i} < -\varepsilon,$$

其中 $v^{k_i}(\lambda) \in \partial f_{k_i}(x^{k_i} + \lambda d_{k_i})$. 在算法 A 中, $x^{k+1} \in M_{c_0}(x^k, d_k, f_k)$, 故 $x^{k+1} = x^k + \lambda_k d_k$, 其中 x^{k+1} 满足

$$f_k(x^{k+1}) \le c_0 f_k^* + (1 - c_0) f_k(x^k),$$

这里 f_k^* 是 (7) 的最优值, 故

$$f_{k_i}(x^{k_i+1}) \le (1 - c_0) f_{k_i}(x^{k_i}) + c_0 f_{k_i}^* \le (1 - c_0) f_{k_i}(x^{k_i}) + c_0 f_{k_i}(x^{k_i} + \bar{\delta} d_{k_i}),$$

其中 $\bar{\delta} = \min\{\delta, \delta'\}$.

由函数列 $\{f_k(x)\}$ 的 ε 下单调性, 可得

$$\begin{aligned} &f_{k_i+1}(x^{k_i+1}) \le f_{k_i}(x^{k_i+1}) + \varepsilon_{k_i} \le (1 - c_0) f_{k_i}(x^{k_i}) + c_0 f_{k_i}(x^{k_i} + \bar{\delta} d_{k_i}) + \varepsilon_{k_i}, \\ &f_{k_i+1}(x^{k_i+2}) \le f_{k_i+1}(x^{k_i+1}), \\ &\qquad\vdots \\ &f_{k_{i+1}}(x^{k_{i+1}}) \le f_{k_{i+1}-1}(x^{k_i+1}) + \varepsilon_{k_{i+1}-1}, \end{aligned}$$

故

$$f_{k_{i+1}}(x^{k_{i+1}}) \le (1-c_0)f_{k_i}(x^{k_i}) + c_0 f_{k_i}(x^{k_i}+\bar{\delta}d_{k_i}) + \sum_{j=k_i}^{k_{i+1}-1}\varepsilon_j. \tag{8}$$

由中值定理 [2], 有

$$f_{k_{i+1}}(x^{k_{i+1}}) \le (1-c_0)f_{k_i}(x^{k_i}) + c_0[f_{k_i}(x^{k_i}) + \bar{\delta}v^{k_i}(\tilde{\lambda})^T d_{k_i}] + \sum_{j=k_i}^{k_{i+1}-1}\varepsilon_j, \quad i=1,2,\cdots, \tag{9}$$

其中 $\tilde{\lambda}\in[0,\bar{\delta}]$, 从而有

$$f_{k_{i+1}}(x^{k_{i+1}}) \le f_{k_i}(x^{k_i}) + c_0\bar{\delta}v^{k_i}(\tilde{\lambda})^T d_{k_i} + \sum_{j=k_i}^{k_{i+1}-1}\varepsilon_j, \quad i=1,2,\cdots. \tag{10}$$

在不等式 (10) 两边取极限，得

$$f(\bar{x}) \le f(\bar{x}) - c_0\bar{\delta}\varepsilon + 0.$$

这是不可能的，故完成了引理的证明.

定理 4. 设引理 3 的条件成立，则由算法 A 产生的序列 $\{x^k\}$ 的任何聚点都是问题 (1) 的一个 Fritz–John 点.

证明. 利用 [6] 中的定理 4 及上面的引理 3, 容易证明本定理的结论.

3. 不精确线搜索的计算

可以看出，算法 A 的主要工作量集中在线搜索，这里结合问题的特性给出一种新的不精确线搜索.

假设 A. 对所有 k, 下列问题：

$$\begin{aligned} \min \quad & f_k(x^k+\lambda d_k), \\ \text{s.t.} \quad & 0\le\lambda\le\lambda_{\max}, \quad \lambda_{\max}>0 \end{aligned} \tag{11}$$

的最优解可在有限处达到，其中 $\lambda_{\max}$ 为满足 $A(x^k+\lambda d_k)\le b$ 的λ 的最大值.

考虑 $x^{k+1}\in M_{c_0}(x^k,d_k,f_k)$. 由于 f_k^* 是未知的，所以 $M_{c_0}(x^k,d_k,f_k)$ 无法预先确定. 但我们仅需求得其中一个元素.

定义 4. $\overline{M}_{c_0}(x^k,d_k,f_k)=\{\lambda|0\le\lambda\le\lambda_{\max},\ f_k(x^k+\lambda d_k)\le c_0 f_k^* + (1-c_0)f_k(x^k)\}$.

为了求解 $x^{k+1}\in M_{c_0}(x^k,d_k,f_k)$, 采取如下的步骤：

步 1. 判断 $\lambda_{\max}$ 是否为 $+\infty$. 若 $\lambda_{\max}=+\infty$, 则由假设知 (11) 有有限解，故可用 $\bar{\lambda}_{\max}$ 代替 $\lambda_{\max}$ 且在 $[0,\bar{\lambda}_{\max}]$ 中有 (11) 的最优解；

步 2. 判断 $\lambda_{\max}$(或 $\bar{\lambda}_{\max}$) 是否为最优解. 若否，寻求 $\lambda_k\in\overline{M}_{c_0}(x^k,d_k,f_k)$, 置 $x^{k+1}=x^k+\lambda_k d_k$.

引理 5. 设 $\bar{\lambda}$ 为 (11) 的一个最优解. 若 $0 \le \lambda \le \bar{\lambda}$, 则 $v_{(\lambda)}^T d_k \le 0$; 若 $\bar{\lambda} \le \lambda \le \lambda_{\max}$, 则 $v_{(\lambda)}^T d_k \ge 0$. 如果 $v_{(\lambda)}^T d_k = 0 (0 \le \lambda \le \lambda_{\max})$, 则 λ 为 (11) 的最优解, 其中 $v(x) \in \partial f_k(x^k + \lambda d_k)$.

证明. 由 $f_k(x)$ 的凸性及次微分定义易得证.

引理 6. 设 $\lambda_{\max}$ 有限且 $v_{(\lambda\max)}^T d_k \le 0$, 则 $\lambda_{\max}$ 为 (11) 的最优解, 其中 $v_{(\lambda\max)} \in \partial f(x^k + \lambda_{\max} d_k)$.

证明. 由引理 5 易得证.

下面给出求 $\lambda_k \in \overline{M}_{c_0}(x_k, d_k, f_k)$ 的计算方法:

步 1. 令 $\lambda_1 = 0, \lambda_2 = \min\{\lambda_{\max}, \bar{\lambda}_{\max}\}$;

步 2. 任取 $v_1 \in \partial f_k(x^k + \lambda_1 d_k), v_2 \in \partial f_k(x^k + \lambda_2 d_k)$, 计算 $v_1^T d_k$ 及 $v_2^T d_k$. 若其中之一为零, 则相应的 λ_1 或 λ_2 必为最优解 (引理 5), 于是已求得 $\overline{M}_{c_0}(x_k, d_k, f_k)$ 中的一个值, 停止; 否则由引理 5 和假设 A 必有 $v_1^T d_k < 0$ 及 $v_2^T d_k > 0$. 过 $(\lambda_1, f_k(x^k + \lambda_1 d_k))$, $(\lambda_2, f_k(x^k + \lambda_2 d_k))$ 作两条直线:

$$L_1 : y = f_k(x^k + \lambda_1 d_k) + v_1^T d_k(\lambda - \lambda_1), \quad L_2 : y = f_k(x^k + \lambda_2 d_k) + v_2^T d_k(\lambda - \lambda_2). \tag{12}$$

设其交点为 $(\alpha, y(\alpha))$, 由引理 5 易知 $y(\alpha) \le f_k^*$;

步 3. 判断 λ_1, λ_2 是否满足:

$$\min\{f_k(x^k + \lambda_1 d_k), f_k(x^k + \lambda_2 d_k)\} \le c_0 y(\alpha) + (1 - c_0) f_k(x^k). \tag{13}$$

若满足, 则由于 $y(\alpha) \le f_k^*$, 我们已经求得 $\bar{M}_{c_0}(x^k, d_k, f_k)$ 中的一个值, 从而相应的有 $x^{k+1} \in M_{c_0}(x^k, d_k, f_k)$; 否则, 转步 4;

步 4. 按线搜索的方法如二分法等缩小区间得到新的 $[\lambda_1, \lambda_2]$, 返回步 2.

下面在比较弱的条件下证明上面方法的步骤的有限性.

定理 7. 设 $0 < c_0 < 1, \lambda_{\max} < +\infty, f_k^* < f_k(x^k)$ 且 $\sup\limits_{0 \le \lambda \le \lambda_{\max}} \{v^k(\lambda)^T d_k\} < +\infty$, 其中 $v^k(\lambda) \in \partial f_k(x^k + \lambda d_k)$; 设 (11) 的某一最优解 $\lambda_k^* \in (0, \lambda_{\max})$, 则必存在包含 λ_k^* 的区间 $[\lambda_1, \lambda_2] (\lambda_1 < \lambda_2)$, 使当 $\lambda \in [\lambda_1, \lambda_2]$ 时,

$$f_k(x^k + \lambda d_k) \le f_k(x^k) + c_0(y(\alpha) - f_k(x^k)),$$

其中 $(\alpha, y(\alpha))$ 为 L_1 与 L_2 的交点, 从而有 $[\lambda_1, \lambda_2] \subset \overline{M}_{c_0}(x^k, d_k, f_k)$.

证明. 由于 $\sup\limits_{0 \le \lambda \le \lambda\max} \{v^k(\lambda)^T d_k\} < +\infty$, 故当 $\lambda \to \lambda_k^*$ 时, $(\lambda - \lambda_k^*) v^k(\lambda)^T d_k \to 0$, 从而存在正数 $\delta_1 > 0$, 使得当 $\lambda_k^* - \delta_1 \le \lambda \le \lambda_k^* + \delta_1$ 时, 有

$$\left| (\lambda - \lambda_k^*) v^k(\lambda)^T d_k \right| \le \frac{1}{2} \frac{(1 - c_0)}{c_0} (f_k(x^k) - f_k^*). \tag{14}$$

由 $f_k(x^k + \lambda d_k)$ 关于 λ 的连续性, 存在正常数 $\delta_2 > 0$, 使当 $|\lambda - \lambda_k^*| \le \delta_2$ 时, 有

$$f_k(x^k + \lambda d_k) \le f_k^* + \frac{1}{2}(1 - c_0)(f_k(x^k) - f_k^*). \tag{15}$$

取 $\delta=\min\{\delta_1,\delta_2\},[\lambda_1,\lambda_2]=[\lambda_k^*-\delta,\lambda_k^*+\delta]$. 对如此选取的 λ_1,λ_2, 由于 $\alpha\in[\lambda_1,\lambda_2]$, 故不妨设 $\alpha\in[\lambda_1,\lambda_k^*]$. 由 (14) 及 $(\alpha,y(\alpha))$ 的定义知

$$y(\alpha)\ge f_k^*-\frac{1}{2}\frac{(1-c_0)}{c_0}(f_k(x^k)-f_k^*),$$

即

$$c_0y(\alpha)\ge c_0f_k^*-(1-c_0)f_k(x^k)+(1-c_0)f_k^*+\frac{1}{2}(1-c_0)(f_k(x^k)-f_k^*),$$

故

$$f_k^*+\frac{1}{2}(1-c_0)(f_k(x^k)-f_k^*)\le c_0y(\alpha)+(1-c_0)f_k(x^k)=f_k(x^k)+c_0(y(\alpha)-f_k(x^k)).$$

从而由 (15) 式及上式知

$$f_k(x^k+\lambda d_k)\le f_k(x^k)+c_0(y(\alpha)-f_k(x^k)),\ \forall\lambda\in[\lambda_1,\lambda_2],\tag{16}$$

而 $y(\alpha)\le f_k^*$, 故由 (16) 式知 $[\lambda_1,\lambda_2]\subset\overline{M}_{c_0}(x^k,d_k,f_k)$.

注. 上面的定理 7 保证了线搜索步骤的有限性，通常仅需几步.

4. 对带完备补偿矩阵的随机规划的应用

考虑随机规划模型

$$\begin{aligned}&\min c^Tx+Q(x,\xi),\\&\text{s.t.}\quad Ax\le b,\end{aligned}\tag{17}$$

其中 $A\in I\!R^{m\times n},Q(x,\xi)=EQ(x,\xi(\omega)),\xi(\omega)$ 为概率空间 $(\Omega,\mathcal{F},P)$ 上的一个随机向量， E 指数数学期望，

$$\begin{aligned}&Q(x,\xi(\omega))=\min q(\omega)^Ty,\\&\text{s.t.}\ WY=b(\omega)-T(\omega)x,\quad y\ge0,\\&W\in I\!R^{m_1\times n_1},\quad b(\omega)\in I\!R^{m_1},\quad T(\omega)\in I\!R^{m_1\times n}.\end{aligned}\tag{18}$$

直接用非线性规划技术解 (17) 是相当困难的，特别 $\nabla Q(x,\xi)$ 的计算相当麻烦. 为此，[6] 考虑了一种逼近方法：设 $\{\xi^k\}$ 为一列在 l_2- 意义下收敛到 ξ 的随机向量，在 (18) 中用 $\xi^k(\omega)$ 代替 $\xi(\omega)$, 得到

$$Q^k(x,\xi^k)=EQ(x,\xi^k(\omega)),$$

则 $Q^k(x,\xi^k)$ 将收敛到 $Q(x,\xi)$. 这里的收敛性可有多种理解，例如 Kall[5] 证明了点收敛性；Wets[14] 证明了上图收敛性；Wang[8] 证明了 $Q(x,\xi)$ 关于 x,ξ 的联合 Lipschitz 连续性. Lipschitz 连续性表示

$$\lim_{k\to+\infty}Q^k(x^k,\xi^k)=Q(x,\xi)\tag{19}$$

对任何收敛到 x 的序列 $\{x^k\}$ 都成立.

为验证微分稳定性条件，注意到 $Q^k(x,\xi^k)$ 上图收敛到 $Q(x,\xi)$ 表示

$$\lim_{k\to+\infty} G(\partial Q^k(x,\xi^k)) = G(\partial Q(x,\xi)), \tag{20}$$

其意义如下:

(i) 对任何序列 $\{(x^k,v^k)\}$, 若 $x^k\to x, v^k\to v$, $v^k\in\partial Q^k(x^k,\xi^k)$, 则 $v\in\partial Q(x,\xi)$;

(ii) 对任何 $(x,v), v\in\partial Q(x,\xi)$, 存在序列 $\{(x^k,v^k)\}$ 收敛到 (x,v), 其中 $v^k\in\partial Q^k(x^k,\xi^k)$. 这个结论见 [15], 由 (20) 即可推出 [9] 微分稳定性条件:

$$\lim_{k\to+\infty} v^k = \nabla Q(x,\xi), \quad v^k\in\partial Q^k(x^k,\xi^k). \tag{21}$$

关于 ξ^k 的构造同 [10].

在上面给出的算法中，需要求 $\partial Q^k(x^k,\xi^k)$ 的一个元素 v^k. 这里给出如何求得. 为了得到逼近问题，设 ξ^k 具有离散分布:

$$P(\xi^k = t_{k_j}) = P_{k_j}, \quad j=1,,\cdots,N_k,$$

则

$$Q^k(x,\xi^k) = \sum_{j=1}^{N_k} P_{k_j} Q(x,t_{k_j}),$$

其中 $t_{k_j} = (q_{k_j}, b_{k_j}, T_{k_j})$ 及

$$\begin{aligned} &Q(x,t_{k_j}) = \min q_{k_j}^T y, \\ &\text{s.t. } WY = b_{k_j} - T_{k_j}x, \quad y\ge 0. \end{aligned} \tag{22}$$

由于每个 $Q(x,t_{k_j})$ 为 x 的凸分片线性函数，它在 Clarke[2] 意义下是正则的，故

$$\partial Q^k(x^k,\xi^k) = \sum_{j=1}^{N_k} P_{k_j}\partial_x Q(x^k,t_{k_j}).$$

任取 $\partial_x Q(x^k,t_{k_j})$ 的一个元素 v_{k_j}, 有

$$v^k = \sum_{j=1}^{N_k} P_{k_j} v_{k_j} \in \partial Q^k(x^k,\xi^k). \tag{23}$$

另一方面，当对 $x=x^k$ 求解 (22) 时，可得相应于 $WY=b_{k_j}-T_{k_j}x$ 的 Lagrange 乘子 Π_{k_j}, 则 $\Pi_{k_j}^T(-T_{k_j})$ 是 $\partial_x Q(x^k,t_{k_j})$ 的一个元素，可取

$$v_{k_j} = -\Pi_{k_j}^T T_{k_j}.$$

事实上，也可求 (22) 的对偶问题:

$$\begin{aligned} &\max (b_{k_j} - T_{k_j}x)^T z, \\ &\text{s.t. } \ W^T z \le q_{k_j} \end{aligned} \tag{23}$$

而得到 $v_{k_j} = -\bar{z}^T T_{k_j}$，其中 $\bar{z}$ 为 (23) 的一个最优解.

5. 算法的实施

在 4 中，我们看到算法存在两方面的潜在困难：

1) 需要计算 $P(\xi^k = t_{k_j}) = P_{k_j}, j = 1, \cdots, N_k$. 一般情况下多维随机向量是由密度函数给出的而不是由分布函数给出的，故 P_{k_j} 的计算是高维数值积分，而数值实现存在困难；

2) 在 [6] 中，为保证收敛，需要 $\{Q^k(x,\xi^k)\}$ 的单调下降性，为此必须付出额外的工作量，从计算角度讲也不经济.

为解决上面的问题，假定下列条件成立：(A_1) 集合 $D = \{x|Ax \le b\}$ 有界；

(A_2) $\xi(\omega)$ 分布在一有界集合上，总可设分布在一有界长方体 T 内，其体积记为 $\Delta(T)$，直径 $\delta(T)$;

(A_3) ξ 的分布密度函数 $g(t)$ 为 Lipschitz 连续的，即存在常数 $L > 0$，使得 $|g(t_1) - g(t_2)| \le L|t_1 - t_2|, \forall t_1, t_2$;

(A_4) 设在无限加细剖分过程中对任一小长方体 v 被剖分成 $v_1, \cdots, v_i$，应满足

$$\delta = \sup_v \left\{ \frac{\max[\delta(v_1), \cdots, \delta(v_i)]}{\delta(v)} \right\} < 1, \tag{24}$$

其中 $\delta(v), \delta(v_1), \cdots, \delta(v_i)$ 分别表示 $v, v_1, \cdots, v_i$ 的直径；

(A_5) 对任给的 ω，都有 $Q(x, \xi(\omega)) > -\infty$.

由假设 (A_5) 及 (A_1) 和 (A_2) 知，$Q(x,\xi(\omega))$ 关于 $x \in D$ 及 $\xi(\omega)$ 为一致有界，即

$$|Q(x,\xi(\omega))| \le M_1 < +\infty, \forall x \in D, \xi(\omega) \in T. \tag{25}$$

由假设 (A_2)，ξ 为分布在一有界长方体 T 内的随机向量. 设剖分进行到第 k 步，剖分 T 成为小块长方体 T_j^k，其相应体积为 $\Delta(T_j^k), j = 1, 2, \cdots, N_k$. 任取 T_j^k 中一点 t_{k_j}，定义 ξ^k 如下：

$$P_{k_j} = P(\xi^k = t_{k_j}) = g(t_{k_j})\Delta(T_j^k), \quad j = 1, \cdots, N_k. \tag{26}$$

定义

$$\tilde{Q}^k(x,\xi^k) = \sum_{j=1}^{N_k} P_{k_j} Q(x, t_{k_j}). \tag{27}$$

假设加细剖分 $\{T_j^k\}, (j = 1, \cdots, N_k)$ 中的 $T_{j_1}^k, \cdots, T_{j_{p_k}}^k$，$P_k \le N_k$，进而得到 $T_j^{k+1} (j = 1, \cdots, N_{k+1})$，定义

$$\tilde{Q}^{k+1}(x,\xi^{k+1}) = \sum_{j=1}^{N_{k+1}} P_{(k+1)_j} Q(x, t_{(k+1)_j}). \tag{28}$$

由参数规划理论知 $Q(x,\xi(\omega)) = Q(x, q(\omega), b(\omega), T(\omega))$ 关于$q(\omega)$ 为凹函数，关于 $b(\omega)$ 和 $T(\omega)$ 为凸函数. 对 $i = 1, \cdots, P_k$，考虑 $\xi^k(\omega) \in T_{j_i}^k$ 的情形. 设 $T_{j_i}^k$ 被加细剖分成 $T_{i_1}^{k+1}, \cdots,$

$T_{i_{l_i}}^{k+1}$, 则 $t_{(k+1)_{i_1}},\cdots,t_{(k+1)_{i_{l_i}}}\in T_{j_i}^k$ 且

$$\sum_{jj=1}^{l_i}\Delta(T_{i_{jj}}^{k+1})=\Delta(T_{j_i}^k). \tag{29}$$

设完备补偿矩阵 W 的所有 $m_1\times m_1$ 阶非奇异子矩阵的个数为 r, 则由 $Q(x,\xi(\omega))$ 的凹凸性得 [8]

$$\begin{aligned}&\left|Q(x,q_{k_j},b_{k_j},T_{k_j})-Q(x,q_{k_j},b_{(k+1)_{j_0}},T_{(k+1)_{j_0}})\right|\\&=\left|q_{k_j}^{l_1T}W_{\bar{l}_1}^{-1}(b_{k_j}-T_{k_j}x)-q_{k_j}^{l_2T}W_{\bar{l}_2}^{-1}(b_{(k+1)_{j_0}}-T_{(k+1)_{j_0}}x)\right|\\&\le\max_{1\le\bar{l}\le r}\left\|q_{k_j}^{lT}W_{\bar{l}}^{-1}\right\|[\|b_{k_j}-b_{(k+1)_{j_0}}\|+\|T_{k_j}-T_{(k+1)_{j_0}}\|\|x\|].\end{aligned} \tag{30}$$

由假设 (A$_1$) 及 (A$_2$), 若 $t_{(k+1)_{j_0}}=(q_{(k+1)_{j_0}},b_{(k+1)_{j_0}},T_{(k+1)_{j_0}})\in T_{j_i}^k$, 则

$$\left|Q(x,q_{k_{j_i}},b_{k_{j_i}},T_{k_{j_i}})-Q(x,q_{k_{j_i}},b_{(k+1)_{j_0}},T_{(k+1)_{j_0}})\right|\le M_2\delta(T_{j_i}^k), \tag{31}$$

其中 $\delta(T_{j_i}^k)$ 表示 $T_{j_i}^k$ 的直径, $0<M_2<+\infty$ 为常数. 同理, 若 $t_{(k+1)_{j_0}}=(q_{(k+1)_{j_0}},b_{(k+1)_{j_0}},T_{(k+1)_{j_0}})\in T_{j_i}^k$, 有

$$\left|Q(x,q_{k_{j_i}},b_{(k+1)_{j_0}},T_{(k+1)_{j_0}})-Q(x,q_{(k+1)_{j_0}},b_{(k+1)_{j_0}},T_{(k+1)_{j_0}})\right|\le M_3\delta(T_{j_i}^k), \tag{32}$$

其中 $0<M_3<+\infty$ 为常数. 从而有

$$\begin{aligned}&\left|P_{k_{j_i}}Q(x,t_{k_{j_i}})-\sum_{jj=1}^{l_i}P_{(k+1)_{i_{jj}}}Q(x,q_{(k+1)_{i_{jj}}},b_{(k+1)_{i_{jj}}},T_{(k+1)_{i_{jj}}})\right|\\&=\left|P_{k_{j_i}}Q(x,t_{k_{j_i}})-\sum_{jj=1}^{l_i}P_{(k+1)_{i_{jj}}}Q(x,t_{(k+1)_{i_{jj}}})\right|\\&=\Big|\sum_{jj=1}^{l_i}g(t_{k_{j_i}})\Delta(T_{i_{jj}}^{k+1})Q(x,q_{k_{j_i}},b_{k_{j_i}},T_{k_{j_i}})\\&\quad-\sum_{jj=1}^{l_i}g(t_{i_{jj}}^{(k+1)})\Delta(T_{i_{jj}}^{(k+1)})Q(x,q^{(k+1)_{i_{jj}}},b_{(k+1)_{i_{jj}}},T^{(k+1)_{i_{jj}}})\Big|\\&=\Big|\sum_{jj=1}^{l_i}(g(t_{k_{j_i}})-g(t_{(k+1)_{i_{jj}}})\Delta(T_{i_{jj}}^{k+1})Q(x,q_{k_{j_i}},b_{k_{j_i}},T_{k_{j_i}})\\&\quad-\sum_{jj=1}^{l_i}g(t_{(k+1)_{i_{jj}}})\Delta(T_{i_{jj}}^{(k+1)})[Q(x,q_{(k+1)_{i_{jj}}},b_{(k+1)_{i_{jj}}},T_{(k+1)_{i_{jj}}})\\&\quad-Q(x,q_{k_{j_i}},b_{k_{j_i}},T_{k_{j_i}})]\Big|.\end{aligned} \tag{33}$$

由于 $g(t)$ 是 Lipschitz 连续而 $\xi(\omega)$ 分布在有界集内，故

$$|g(t)| \le M_4 < +\infty, \quad t \in T, \quad M_4 \text{ 为常数}, \tag{34}$$

$$\begin{aligned} &|Q(x, q_{(k+1)_{i_{jj}}}, b_{(k+1)_{i_{jj}}}, T_{(k+1)_{i_{jj}}}) - Q(x, q_{k_{j_i}}, b_{k_{j_i}}, T_{k_{j_i}})| = |Q(x, q_{(k+1)_{i_{jj}}}, b_{k_{j_i}}, T_{k_{j_i}}) \\ &\quad - Q(x, q_{k_{j_i}}, b_{k_{j_i}}, T_{k_{j_i}}) + Q(x, q_{(k+1)_{i_{jj}}}, b_{(k+1)_{i_{jj}}}, T_{(k+1)_{i_{jj}}}) \\ &\quad - Q(x, q_{(k+1)_{i_{jj}}}, b_{k_{j_i}}, T_{k_{j_i}})| \le (M_2 + M_3)\delta(T_{j_i}^k). \end{aligned} \tag{35}$$

结合 (29)-(35) 得

$$\left|P_{k_{j_i}} Q(x, t_{k_{j_i}}) - \sum_{jj=1}^{l_i} P_{(k+1)_{i_{jj}}} Q(x, t_{(k+1)_{i_{jj}}})\right| \le [L \cdot M_1 + M_4(M_2 + M_3]\Delta(T_{j_i}^k)\delta(T_{j_i}^k). \tag{36}$$

记 $M = L \cdot M_1 + M_4(M_2 + M_3)$, 则

$$\left|\tilde{Q}^k(x, \xi^k) - \tilde{Q}^{k+1}(x, \xi^{k+1})\right| \le M \sum_{i=1}^{P_k} \Delta(T_{j_i}^k)\delta(T_{j_i}^k). \tag{37}$$

定理 8. 函数列 $\{\tilde{Q}^k(x, \xi^k)\}$ 在 $x \in D$ 上是 ε- 下单调的.

证明. 令

$$\varepsilon_k = M \sum_{i=1}^{P_k} \Delta(T_{j_i}^k)\delta(T_{j_i}^k), \quad k = 1, 2, \cdots. \tag{38}$$

考虑前 n 项和

$$S_n = \sum_{k=1}^{n} \varepsilon_K. \tag{39}$$

设剖分到第 n 次，所有 $T_i^n (i = 1, 2, \cdots, N_n)$ 最多经过 d_n 次剖分而得到，再由假设 (A_4) 可得

$$S_n \le M \cdot \Delta(T) \cdot \delta(T) \sum_{j=1}^{d_n} \delta^j, \quad 0 < \delta < 1. \tag{40}$$

由于 $0 < \delta < 1$, 故 $\sum_{j=1}^{+\infty} \delta^j < +\infty$, 从而可得结论.

利用 $Q(x, \xi)$ 关于 x, ξ 的 Lipschitz 连续性 [8], 不难得到下述定理.

定理 9. $\{\tilde{Q}^k(x, \xi^k)\}$ 上图收敛于 $Q(x, \xi)$.

上面的定理保证了条件 (3) 和 (4) 是满足的，故算法可以进行下去.

关于剖分过程中如何保证 (A_4) 成立，实际上只要使小长方体的边长的比例不要太小 (或太大) 即可.

本文考虑的是带不等式约束的，原则上等式可化为不等约束，也可以利用本文思想及 [9] 的结果考虑.

6. 数值实验

例 1. 考虑

$$
\begin{aligned}
& z=\min 2x_1-x_2+Q(x,\xi), \\
& \text{s.t.}\quad x_1+x_2\leq 10,\quad x_1,x_2\geq 0 \\
& Q(x,\xi)=EQ(x,\xi(\omega)),Q(x,\xi(\omega))=\min q^{+T}y_++q^{-T}y_- \\
& \text{s.t.}\quad y_+-y_-=b(\omega)-Ax,\quad y_+,y_-\geq 0,
\end{aligned}
$$

其中 $q^+=(1,2)^T,q^-=(0.6,1)^T,A=\begin{bmatrix}2 & -1\\ -1 & 0.5\end{bmatrix}$, $b(\omega)$ 的分布密度函数

$$
f(t_1,t_2)=\frac{6}{\pi}e^{(-6(t_1^2-t_1t_2+t_2^2))}.
$$

由于 $P\{t|t\in[-1.2,1.2]\times[-1.2,1.2]\}\geq 0.9998$, 故可设

$$
T=[-1.2,1.2]\times[-1.2,1.2].
$$

数值结果见表 1. 由于这个例子的特殊性，可以把解析式写出来，得到最优值 $F^*\approx 0.3957491$. 由表 1 知，本文的算法是有效的.

表 1

初始点	迭代次数	剖分块数	最优解	最优值	概率值
$x_1=1.0$ $x_2=8.0$	22	455	$x_1^*=2.7130065$ $x_2^*=5.9141846$	$F^*=0.4002215$	$P^*=0.9981698$
$x_1=5.0$ $x_2=2.0$	22	455	$x_1^*=1.2625115$ $x_2^*=3.0124393$	$F^*=0.4002002$	$P^*=0.9981698$
$x_1=10.0$ $x_2=0.0$	24	505	$x_1^*=1.2917189$ $x_2^*=3.0707946$	$F^*=0.4000045$	$P^*=0.9986605$

例 2. 考虑

$$
\begin{aligned}
& z=\min 2x_1-x_2+x_3+Q(x,\xi), \\
& \text{s.t.}\quad x_1+x_2+x_3\leq 10,\quad x_1-x_2\leq 4,\quad x_1,x_2,x_3\geq 0, \\
& Q(x,\xi)=EQ(x,\xi(\omega)),Q(x,\xi(\omega))=\min q^{+T}y_++q^{-T}y_-, \\
& \text{s.t.}\quad y_+-y_-=b(\omega)-Ax,\quad y_+,y_-\geq 0,
\end{aligned}
$$

其中 $q^+ = (1,2,1.5,1,1)^T, q^- = (0.6,1,-1,0.5,-0.5)^T$,

$$A^T = \begin{bmatrix} 2 & -1 & 1 & -0.5 & 2.0 \\ -1 & 0.5 & 0.5 & 1 & 0.5 \\ 0 & 0 & 1 & 2.0 & -1.0 \end{bmatrix},$$

$b(\omega)$ 的分布密度函数

$$f(t_1,t_2,\cdots,t_5) = \frac{186.624}{\pi\sqrt{2\pi}} e^{(-25.92(t_3^2+t_4^2+t_5^2))}, \quad T = [-0.5,0.5]^5.$$

计算结果见表 2:

表 2

初始点	迭代次数	剖分块数	最优解	最优值	概率值
$x_1 = 0.0$ $x_2 = 0.0$ $x_3 = 0.0$	105	1520	$x_1^* = 2.7514025$ $x_2^* = 5.5493350$ $x_3^* = 0.8367474$	$F^* = -6.9713684$	$P^* = 1.0009783$
$x_1 = 5.0$ $x_2 = 2.0$ $x_3 = 2.0$	103	1490	$x_1^* = 2.7513075$ $x_2^* = 5.5500176$ $x_3^* = 0.8370262$	$F^* = -6.9712021$	$P^* = 1.0009421$

参 考 文 献

[1] M.Bazaraa, Shetty, Nonlinear Programming, Theory and Algorithms, John-Wiley, New York, 1979.

[2] F.Clarke, Optimization and Nonsmooth Analysis, John-Wiley, New York, 1983.

[3] J.E.Dennis, R.B.Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, New Jersey, 1983.

[4] Yu Ermoliev, R.J-B Wets, Numerical Techniques for Stochastic Programming, 1988.

[5] P.Kall, Approximation to Stochastic Programs with Complete Recourse, *Numer. Math.*, 22(1974), 333.

[6] X.Li, J.Wang, Approximate feasible direction method for stochastic programming problems with recourse: linear inequality deterministic constraints, *optimization* vol. 21, No.3, 1990.

[7] M.J.D. Powell, Minimization without exact line searches", in Nonlinear Programming Siam-Ams Proceedings, V.IX, Edited by R.W.Cottle and C.E.Lemke. 1976.

[8] J.Wang, Lipschitz continuity of objective functions in stochastic programs with fixed recource and it's applications", *Math. Progr. Study*, 27(1986).

[9] J.Wang, Approximation reduced gradient algorithm for stochastic programs with recourse: linear equality deterministic constraints", Tchnical Report, Dept. of Math., Nanjing University, China, 1988.

[10] R.Wets L.Nazareth, Algorithms for stochastic programs: the case of nonstochastic tenders, *Math. Progr. Study*, 28(1986).

[11] J.Wang, Approximate nonlinear programming algorihtms for solving stochastic programs with recourse", *Annals of Operations Research*, 31(1991), 371–384.

[12] J.Wang, Approximation method for probalistic constrained programs, preprint, Institute of Math., Nanjing Univ. China.

[13] 王金德，随机规划，南京大学出版社， 1990.

[14] R.Wets, Stochastic programming: solution techniques and approximation schemes, in Math. Progr., the State of the Art, ed. by A. Bachem, M.Groschel and B.Korte, Springer-Verlag, 1983, 566.

[15] R.Wets, Convergence of convex functions, variational inequalities and convex optimization problems, in: Variational Inequalities and Complementarity Problems, ed. by R.Cottle et., John Wiley, 1980, 375.