

ビッグデータに基づく機会制約問題の 適応型差分進化による解法

田川 聖治^{1,a)}

概要：機会制約問題の定式化において、不確実性を正規分布など一般的な確率分布でモデル化した場合、モデル化誤差による解の信頼性の低下が懸念される。本稿では、従来の確率分布によるモデルに代えて、膨大な数のシナリオ（ビッグデータ）を用いた機会制約問題の定式化と、その差分進化に基づく解法を提案する。提案法によれば、ビッグデータのクラスタリングにより、解の信頼性の向上と計算負荷の軽減を両立できる。また、提案法の有効性を降雨量のビッグデータを用いた貯水池の設計問題で検証する。

Chance Constrained Problem based on Big Data With Solution by Adaptive Differential Evolution

TAGAWA KIYOHARU^{1,a)}

1. はじめに

現実の世界において、様々な意思決定に関わる諸問題の多くは、予測が難しい不確実性を含む最適化問題である。このため、不確実性を含む最適化問題が長年に渡って研究されてきた。それらの最適化問題の記述法は幾つもあるが、非確率論的なものと確率論的なものに大別される。

機会制約問題 (CCP: Chance Constrained Problem) は、確率論的な定式化の1つであり、非確率論的なロバスト最適化問題 [1] の制約条件を確率的に緩和したものである。CCP では、不確実性に対するリスクが確率で保証された現実的な解が得られるため、様々な実問題への応用が期待され、確率計画法の分野で研究されてきた [2], [3]。

確率計画法における CCP の定式化では、不確実性を正規分布など一般的な確率分布でモデル化した確率モデルを使用する。著者らも確率モデルに基づく CCP の定式化と解法の研究に取り組んできた [4], [5], [6], [7], [8]。確率モデルを CCP で使用する利点として、モンテカルロ法 [9] により膨大な数のシナリオ（疑似シナリオ）を確率変数とし

て生成できる。それらの膨大な数の疑似シナリオを用いれば、CCP に含まれる確率の計算精度を高め、得られる解の信頼性を保証できると思われる [10]。また、確率モデルに既知の確率分布を仮定することで、確率の計算を陽に含まない等価問題に CCP を変換できる場合もある [6]。

一方、確率モデルの問題点として、疑似シナリオに含まれるモデル化誤差が挙げられる。すなわち、モンテカルロ法で生成した膨大な数の疑似シナリオから確率を計算しても、「Garbage In, Garbage Out」に陥る恐れがある。

近年、センサネットワーク等の発達により、様々な分野でビッグデータが蓄積され、その有効な活用法が注目されている。本稿では、膨大な数のシナリオ（ビッグデータ）を用いた CCP の定式化と、その差分進化 (DE: Differential Evolution) [11] に基づく解法を提案する。CCP の定式化で確率モデルに代えてビッグデータを用いれば、モデル化誤差の問題が解消される。提案法では、ビッグデータのクラスタリングと重み付けから、新たに加重シナリオを生成し、CCP に含まれる確率の計算に使用する。これにより、解の信頼性の向上と計算負荷の軽減を両立させる。また、標準的な DE [11] は制約条件のない最適化問題を対象としている。そこで、上記の CCP の解を求めるため、著者らが考案した適応型 DE (JADE2G) [12] を採用する。

¹ 近畿大学 理工学部
School of Science and Engineering,
Kindai University, Higashi-Osaka 577-8502 Japan
^{a)} tagawa@info.kindai.ac.jp

2. 機会制約問題

機会制約問題 (CCP) は個別機会制約問題と同時機会制約問題に大別される。以下に, CCP の記述に用いる記号を定義する。本稿では, シナリオを実数値ベクトル $\xi \in \mathfrak{R}^K$ とし, 全シナリオの集合をビッグデータ $\Omega \subseteq \mathfrak{R}^K$ とする。また, 充足水準 $\alpha \in (0, 1)$ は用途ごと任意に与える。

決定変数 $x = (x_1, \dots, x_D) \in X = [\underline{x}_j, \bar{x}_j]^D \subseteq \mathfrak{R}^D$

シナリオ $\xi = (\xi_1, \dots, \xi_K) \in \Omega \subseteq \mathfrak{R}^K$

可測関数 $g_m : X \times \Omega \rightarrow \mathfrak{R}, m = 1, \dots, M$

目的関数 (最小化) $g_0 : X \rightarrow \mathfrak{R}$

充足水準 (確率) $\alpha \in (0, 1)$

2.1 個別機会制約問題 (SCCP : Separate CCP)

SCCP では, M 個の各制約条件が満たされる確率が, 与えられた α 以上であり, 下記のように定式化される。

$$\begin{cases} \min_{x \in X} & g_0(x) \\ \text{sub. to} & \Pr(\forall \xi \in \Omega : g_m(x, \xi) \leq 0) \geq \alpha, \\ & m = 1, \dots, M \end{cases} \quad (1)$$

ただし, $\Pr(A)$ は事象 A が起きる確率である。

2.2 同時機会制約問題 (JCCP : Joint CCP)

JCCP では, すべての制約条件が満たされる確率が, 与えられた α 以上であり, 下記のように定式化される。

$$\begin{cases} \min_{x \in X} & g_0(x) \\ \text{sub. to} & \Pr \left(\begin{array}{l} \forall \xi \in \Omega : g_m(x, \xi) \leq 0, \\ m = 1, \dots, M \end{array} \right) \geq \alpha \end{cases} \quad (2)$$

2.3 シナリオ集合による確率の推定

以下のように指示関数 $\mathbb{1}$ を定義する。

$$\mathbb{1}(A) = \begin{cases} 1 & \text{if 事象 } A \text{ が真である} \\ 0 & \text{if 事象 } A \text{ が偽である} \end{cases} \quad (3)$$

式 (2) の JCCP の場合, 制約条件の確率の計算は,

$$\begin{aligned} & \Pr(\forall \xi \in \Omega : g_m(x, \xi) \leq 0, m = 1, \dots, M) \\ &= \frac{1}{|\Omega|} \sum_{\xi \in \Omega} \mathbb{1}(g_m(x, \xi) \leq 0, m = 1, \dots, M) \end{aligned} \quad (4)$$

となる。ただし, $|\Omega|$ は全シナリオ数である。

ビッグデータは巨大であり, 式 (4) の計算は容易でない。そこで, ビッグデータ Ω から N 個のシナリオをランダムに選択し, そのシナリオ集合 $\Xi = \{\xi^1, \dots, \xi^N\} \subseteq \Omega$ から, 以下のように式 (4) の確率の値を推定する。

$$\begin{aligned} & \Pr(\forall \xi \in \Omega : g_m(x, \xi) \leq 0, m = 1, \dots, M) \\ &\simeq \frac{1}{N} \sum_{n=1}^N \mathbb{1}(g_m(x, \xi^n) \leq 0, m = 1, \dots, M) \end{aligned} \quad (5)$$

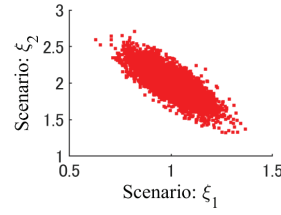


図 1 ビッグデータ Ω

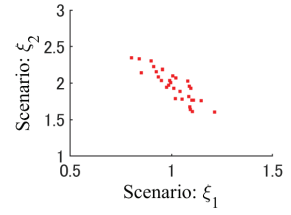


図 2 シナリオ集合 Ξ

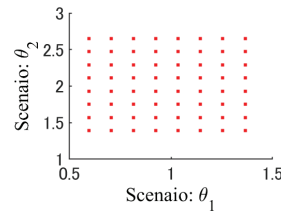


図 3 標準シナリオ

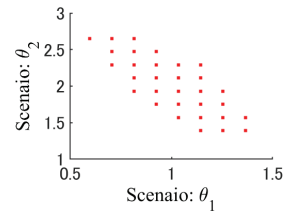


図 4 加重シナリオ

3. ビッグデータの活用

3.1 加重シナリオの生成

ビッグデータから N 個のシナリオ $\xi^n \in \Xi$ をランダムに選択すると, ビッグデータに含まれる大半のシナリオが使われず無駄になる。そこで, すべてのシナリオの情報を集約した加重シナリオ $\theta^n \in \Theta$ を以下の手順で生成する。

手順 1 ビッグデータに含まれる全シナリオから K 次元のヒストグラム [13] を作る。これにより, 全シナリオがヒストグラムの階級の数にクラスタリングされる。

手順 2 各階級の中心点を階級の標準シナリオとする。

手順 3 各階級の相対度数を標準シナリオの重みとする。

手順 4 重みが非零の標準シナリオを加重シナリオとする。ただし, 階級の区間と数は適切に与えるものとする。

3.2 加重シナリオによる確率の推定

加重シナリオ $\theta^n \in \Theta$ の重みを $w_n \in (0, 1]$ とし, 加重シナリオの集合 $\Theta = \{\theta^1, \dots, \theta^N\} \subseteq \mathfrak{R}^K$ から, 以下のように式 (4) に示した JCCP の確率の値を推定する。

$$\begin{aligned} & \Pr(\forall \xi \in \Omega : g_m(x, \xi) \leq 0, m = 1, \dots, M) \\ &\simeq \sum_{n=1}^N w_n \mathbb{1}(g_m(x, \theta^n) \leq 0, m = 1, \dots, M) \end{aligned} \quad (6)$$

3.3 加重シナリオの具体例

シナリオの次元を $K = 2$ とし, ビッグデータを以下の平均 μ , 分散共分散行列 C の正規分布で模擬する。

$$\xi = (\xi_1, \xi_2) \sim \mathcal{N}(\mu, C) \quad (7)$$

分散共分散行列 C は標準偏差の σ_1 と σ_2 から,

$$C = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \quad (8)$$

となる。ただし, $\rho \in [-1, 1]$ は相関係数である。

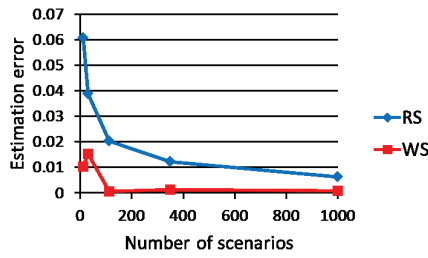


図5 確率の推定誤差

図1にシナリオ数5000のビッグデータ Ω を示す。図2に $N = 31$ 個のシナリオ集合 $\Xi \subseteq \Omega$ を示す。図3にビッグデータから作成した標準シナリオを示す。図4に標準シナリオから選んだ $N = 31$ 個の加重シナリオを示す。

図2と図4から、ランダムに選択したシナリオ $\xi^n \in \Xi$ に比べて、加重シナリオ $\theta^n \in \Theta$ の方が広範囲に分布している。さらに、図1のビッグデータの分布範囲と比較し、加重シナリオでは情報の損失が少ないことがわかる。

上記のビッグデータ Ω について、式(9)の確率をシナリオ集合 Ξ と加重シナリオの集合 Θ を用いて推定する。

$$\begin{cases} g(\mathbf{x}, \boldsymbol{\xi}) = x_1 \xi_1 + x_2 \xi_2 - 3.15 \\ \Pr(\forall \boldsymbol{\xi} \in \Omega : g_m(\mathbf{x}, \boldsymbol{\xi}) \leq 0) \end{cases} \quad (9)$$

ただし、式(8)の相関係数は $\rho = -0.8$ 、 $\mathbf{x} = (1, 1)$ とする。

図5にランダムに選んだシナリオ(RS) $\xi^n \in \Xi$ と、加重シナリオ(WS) $\theta^n \in \Theta$ 、 $n = 1, \dots, N$ による式(9)の確率の推定誤差を示す。横軸はシナリオ数 N である。RSの結果は100回の試行の平均値である。図5から、RSに比べてWSは推定誤差が遙かに小さいことがわかる。

4. 適応型差分進化

著者らが考案したJADE2G[12]は、制約条件のない最適化問題を対象としたJADE[14]を、制約条件付き最適化問題に適用するために拡張したものである。JADE2GではCCPに対する解候補の集団を、実行可能なグループ G_1 と実行不可能なグループ G_2 に分け、各グループ内で独立にJADEの制御パラメータを適応的に調整する。

例えば、JCCPにJADE2Gを適用する場合、解候補 $\mathbf{x} \in G_2$ と加重シナリオの集合 Θ から式(6)で推定された確率を $P(\mathbf{x}, \Theta)$ とし、 $\mathbf{x} \in G_2$ の制約違反量 $\phi(\mathbf{x})$ を

$$\phi(\mathbf{x}) = \max\{\alpha - P(\mathbf{x}, \Theta), 0\} \quad (10)$$

とする。JADE2Gは $\mathbf{x} \in G_1$ については式(2)の目的関数 $g_0(\mathbf{x})$ をJADEで最小化し、 $\mathbf{x} \in G_2$ については式(10)の制約違反量 $\phi(\mathbf{x})$ をJADEで最小化する。また、次世代の集団の選択では、実行可能な $\mathbf{x} \in G_1$ を優先する。

5. 貯水池の設計問題

近年、地球温暖化の影響であろうか、短期間の集中豪雨

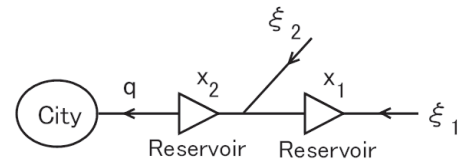


図6 河川と貯水池のモデル

(ゲリラ豪雨)による水害が都市部でも増えている。本稿では、提案法の応用例として、河川の氾濫による水害から都市を守るための貯水池の設計問題[15]を取り上げる。

5.1 設計問題の定式化

図6は河川のモデルであり、水の流れる方向を矢印で示している。ゲリラ豪雨により河川に流れ込む水量 ξ_1 と ξ_2 は、過去の観測からビッグデータとして得られているものとする。ただし、本稿では、水量 $\boldsymbol{\xi} = (\xi_1, \xi_2) \in \Omega$ のビッグデータ Ω を、式(7)の正規分布で模擬する。

図6の川下には都市があり、都市を水害から守るために建設を計画している2つの貯水池を三角印で示す。ここで、各貯水池の貯水量 x_1 と x_2 を決定変数とする。各貯水池の建設費は貯水量 x_j に比例する。貯水池の設計問題では、都市を水害から守って貯水池の建設費を最小としたい。貯水池を満杯にして都市に流れ込む水量 $q(\mathbf{x}, \boldsymbol{\xi})$ は、

$$q(\mathbf{x}, \boldsymbol{\xi}) = \begin{cases} ((\xi_1 - x_1) + \xi_2) - x_2 & \text{if } (\xi_1 > x_1) \wedge (\xi_1 + \xi_2 > x_1 + x_2) \\ \xi_2 - x_2 & \text{if } (\xi_1 \leq x_1) \wedge (\xi_2 > x_2) \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

となる。そこで、以下のような関数を定義する。

$$\begin{cases} g_1(\mathbf{x}, \boldsymbol{\xi}) = \xi_1 + \xi_2 - x_1 - x_2 \\ g_2(\mathbf{x}, \boldsymbol{\xi}) = \xi_2 - x_2 \end{cases} \quad (12)$$

式(11)の水量が $q(\mathbf{x}, \boldsymbol{\xi}) \leq 0$ となる確率を $\alpha = 0.9$ とし、貯水池の設計問題を以下のJCCPに定式化する。

$$\begin{cases} \min_{\mathbf{x}} & g_0(\mathbf{x}) = 3x_1 + x_2 \\ \text{sub. to} & \Pr \left(\begin{matrix} g_1(\mathbf{x}, \boldsymbol{\xi}) \leq 0 \\ g_2(\mathbf{x}, \boldsymbol{\xi}) \leq 0 \end{matrix} \right) \geq 0.9 \\ & 0 \leq x_1 \leq 0.8, 0 \leq x_2 \leq 2.5 \end{cases} \quad (13)$$

5.2 実験結果と考察

水量 $\boldsymbol{\xi} \in \Omega$ の相関係数 ρ を変えて、式(13)のJCCPの最適解 $\mathbf{x}^* \in \mathbf{X}$ をJADE2Gで求めた。加重シナリオ $\theta^n \in \Theta$ 、 $n = 1, \dots, N$ の個数は $N \approx 100$ とした。

表1に最適解 $\mathbf{x}^* = (x_1^*, x_2^*)$ とその目的関数値 $g_0(\mathbf{x}^*)$ 、式(6)による確率の推定値 $P(\mathbf{x}^*, \Theta)$ 、式(4)による真の確率 $P(\mathbf{x}^*, \Omega)$ を示す。表1から、 $P(\mathbf{x}^*, \Theta)$ と $P(\mathbf{x}^*, \Omega)$ の値は近く、JCCPの解が実行可能か否かも正しく判定できている。これより、提案法の有効性が確認できる。

表 1 JADE2G による式 (13) の JCCP の最適解

ρ	x_1^*	x_2^*	$g_0(\mathbf{x}^*)$	$P(\mathbf{x}^*, \Theta)$	$P(\mathbf{x}^*, \Omega)$
-0.8	0.675	2.500	4.525	0.928	0.902
0.0	0.791	2.500	4.873	0.911	0.904
+0.8	0.800	2.500	4.900	0.851	0.849

表 1 によれば, 目的関数値 (建設費) $g_0(\mathbf{x}^*)$ は水量 ξ_1 と ξ_2 の相関係数 ρ に比例して増加している. また, 水量の相関係数が負の場合, 建設費が割高な貯水池の容量 x_1^* を減らして建設費を抑制している. 一方, 水量の相関係数が正の場合, 両貯水池の容量を最大限に増やしても, 式 (13) の JCCP の実行可能解は得られない. すなわち, 水量 ξ_1 と ξ_2 に正の相関がある場合, 計画している 2 つの貯水池だけでは, 水害から都市を守れる確率は 90% 未満である.

6. おわりに

本稿では, ビッグデータから生成した加重シナリオと JADE2G を用いた CCP の解法を提案した. また, 小規模な貯水池の設計問題で提案法の有効性を確認した.

今後の課題として, 以下の事項が挙げられる.

- ヒストグラムの階級の区間と数の決め方を検討する.
- 現在の加重シナリオは確定的に生成しているが, ランダム性を加えた確率的な生成法を考案する.
- 標準シナリオの生成に階層的なクラスタリングの技法 [13] を導入する. これにより, 重点サンプリング [9] のように加重シナリオを特定の領域に集中させる.
- 国土交通省の高性能レーダ雨量計ネットワーク (XRAIN) [16] など, 実際のビッグデータを利用した現実的な CCP で提案法の有効性を検証する.

謝辞 本研究は, JSPS 科学研究費補助金 (科研費) 課題番号 17K06508 の助成を受けたものである.

参考文献

- [1] 武田朗子: ロバスト最適化法とその動向, 電気学会論文誌 C, Vol. 134, No. 6, pp. 760–764 (2014).
- [2] Prékopa, A.: *Stochastic Programming*, Kluwer Academic Publishers (1995).
- [3] 椎名孝之: 確率計画法, 朝倉書店 (2015).
- [4] 田川聖治, 宮永 峻: 個別機会制約条件を含む最適化問題の経験分布と差分進化による解法, 情報処理学会研究報告, Vol. 2017-MPS-112, No. 10, 情報処理学会, pp. 1–6 (2017).
- [5] 田川聖治: 重み付き経験分布と差分進化による機会制約問題の解法, 情報処理学会研究報告, Vol. 2017-MPS-116, No. 14, 情報処理学会, pp. 1–6 (2017).
- [6] 田川聖治, 綿谷剛至: 銀行預金と銀行融資を利用した機会制約ポートフォリオ最適化, 情報処理学会研究報告, Vol. 2018-MPS-120, No. 10, 情報処理学会, pp. 1–6 (2018).
- [7] 田川聖治, 宮永 峻: 重み付き経験分布と差分進化による機会制約問題の解法と治水計画への応用, 電気学会論文誌 C, Vol. 138, No. 10, pp. 1260–1268 (2018).
- [8] Tagawa, K. and Miyanaga, S.: An approach to chance constrained problems using weighted empirical distribution and differential evolution with application to flood control planning, *Electronics and Communications in Japan*, Vol. 102, No. 2, pp. 1–11 (2019).
- [9] 津田孝夫: モンテカルロ法とシミュレーション, 電子計算機の確率論的応用, 三訂版, 培風館 (1995).
- [10] Tempo, R., Calafiore, G. and Dabbene, F.: *Randomized Algorithms for Analysis and Control of Uncertain Systems*, Springer (2005).
- [11] Price, K. V., Storn, R. M. and Lampinen, J. A.: *Differential Evolution - A Practical Approach to Global Optimization*, Springer (2005).
- [12] Tagawa, K.: Extension of weighted empirical distribution and group-based adaptive differential evolution for joint chance constrained problems, *Proc. GECCO'18 Companion*, pp. 223–224 (2018).
- [13] 野口博司: 図解と数値例で学ぶ多変数解析入門: ビッグデータ時代のデータ解析, 日本規格協会 (2018).
- [14] Zhang, J. and Sanderson, A. C.: JADE: Adaptive differential evolution with optional external archive, *IEEE Trans. on Evolutionary Computation*, Vol. 13, No. 5, pp. 945–958 (2009).
- [15] Prékopa, A.: Flood control reservoir system design using stochastic programming, *Mathematical Programming Study*, Vol. 9, pp. 138–151 (1978).
- [16] 国土交通省: 川の防災情報, <http://www.river.go.jp>.