

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 8 日現在

機関番号：17102

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24560499

研究課題名(和文)変動の主観的リスクを考慮した多段決定問題の解法に関する研究

研究課題名(英文) Solution of multistage decision problems considering subjectively evaluated risks of changes in problems

研究代表者

村田 純一 (MURATA, Junichi)

九州大学・システム情報科学研究科(研究院・教授)

研究者番号：60190914

交付決定額(研究期間全体)：(直接経費) 2,400,000円

研究成果の概要(和文)：変動を含む多段決定問題の解法を開発した。多段決定問題は、街路網における最短経路発見のように、選択肢集合の中から一つを適切に選択することを繰り返すことで、結果として得られる評価を最適化する問題である。対象とする変動は、街路網の区間の通行止めに相当する選択肢集合の変動である。

変動の発生確率と変動による損失に基づいて主観的なリスク評価を得る方法、主観的リスクを考慮して解を求める強化学習法、複数の解を求め変動に応じて適切に切替えて使用する学習分類子システムを開発し、さらに、出力が不確定に変動する太陽光発電を含む電力システムの計画・運用問題の解法を開発した。

研究成果の概要(英文)：Methods have been developed for multistage decision problems with changes. A multistage decision problem finds a sequence of appropriate alternatives such that their overall evaluation is optimized. The typical evaluation is finding a shortest path in a network of streets. The changes considered are those in the set of alternatives which corresponds to street closures in street networks.

The following have been developed: a method for subjective evaluation of risks based on the probability of and the damages caused by changes, a reinforcement learning algorithm that solves the problem considering the subjective risks, and learning classifier systems that acquire multiple solutions and switch one from another in response to the changes. Also, operation planning schema have been developed for power systems with photovoltaic generation units whose outputs changes non-deterministically.

研究分野：システム工学

キーワード：最適化 不確定性 リスク 機械学習 再生可能エネルギー

1. 研究開始当初の背景

(1) 課題

最適化は、意思決定やシステムの設計、計画、運用、制御において重要な役割を果たす。最適化問題は、制約条件を満たしつつ目的関数の値を最適（最大または最小）にする変数の値を求める問題として定式化される。多段決定問題は複数回にわたって行う選択を、その総合的な結果が最良になるように行う問題である。選択肢が変数であり、選択肢の集合が制約条件を規定する。たとえば、街路網における目的地までの最短時間経路を求める問題では、総所要時間が最小となる街路区間の系列を求める。選択可能な街路区間が制約条件に、総所要時間が目的関数に、街路区間の系列が変数に相当する。

現実問題においては、制約条件や目的関数が時間とともに変動することも多い。街路網の問題では、渋滞や通行止めという変動が発生する。現実問題を扱う上では、このような変動を適切に考慮することができる最適化の方法が必要である。

(2) 研究の状況

多段決定問題の解法

多段決定問題の代表的解法は動的計画法である。また、対象とする問題が未知の場合に問題を推定しながら最適ないしそれに近い解を求める方法として強化学習（別名適応動的計画法、近似動的計画法）が提案されている。これらの方法は対象問題中に存在する確率的な変動を考慮することができる。しかし、確率的な変動は基本的に期待値によって評価されており、リスク評価の本質的指標である変動発生確率や変動による損失については、それらの一部を考慮した解法が提案されているに過ぎない。また、多くの場合、リスク評価は意思決定者の主観に依存するが、従来リスクを考慮した解法では、主観を自然な形で取り入れることが困難であった。

変動リスクの評価法

変動が発生すると、通常、目的関数の値が劣化する。このリスクは変動の発生確率と変動による損失の2指標に基づいて評価する必要がある。これを、通常の強化学習における変動の取り扱いに準じて、損失の単純な期待値として1指標に統合して評価すると、元来の2指標固有の情報失われ、適切な評価が行えなくなる。そこで、確率は低いが損失の大きいケースへの感度を高くして評価する方法 [1] や、期待値と分散の2種の統計量を用いた方法 [2][3] などが提案されている。しかし、これらの方法では、変動による損失が直接的に評価できる報酬すなわち目的関数の値の変動のみを考慮している。

一方で、現実には制約条件が変動する問題も多く存在する。街路網の例における通行止めは、利用可能な選択肢の集合の変動であり、これは制約条件の変動である。この他にも、

電力会社が、想定される電力需要を満たしつつ発電費用が最小となるように各発電機の時間帯ごとの起動・停止を決定する問題においては、実際の需要は想定した値から変動する可能性がある。これも、発電機出力が満たすべき制約条件の変動である。このような制約条件が変動する多段決定問題を対象とした解法は見受けられない。また、上記の既存の方法では、感度を表す係数や、期待値と分散の重みづけを調整することによって、意思決定者のリスク評価の主観的傾向を取り入れることができる。しかし、感度や重みは意思決定者の考えと直結したのではなく、その適切な調整は容易であるとは言い難い。特に、カーナビのように一般ユーザーが意思決定者である場合では、職務として運用者が意思決定を行う電力系統運用などと異なり、感度や重みの調整は意思決定者にとって大きな負担となる。

2. 研究の目的

(1) 概要

本研究は、多段決定問題に発生する変動（街路網の区間の通行止めなど）のリスクを、意思決定者の主観を考慮に入れやすい形で評価し、これを組み入れて多段決定問題の解（例では出発前に定める経路）を求める方法および、解のオンライン修正（例では出発後の経路変更）を行うことができる方法を考案することを目的とした。

(2) 詳細

まず、対象とする変動のリスクを評価する方法を確立する。リスク評価においては変動の発生確率と変動による損失の2指標を考慮する必要がある。通常の動的計画法や強化学習では1目的最適化問題しか取り扱えないため、これら2指標に元来の目的関数を加えた3指標を、適切に一つの目的関数に統合する必要がある。本研究では、意思決定者の主観を従来手法よりも自然に反映させることができるリスクの新しい観点を提案し、それに基づいてこの統合を行い、リスクを評価する方法を明らかにする。

上述のリスクの評価を用いて、変動を考慮した多段決定問題の解を得る方法を考案する。同時に、得られた解を実行中に変動が生じた場合、より適切な解へ変更するオンライン修正機能を開発する。

3. 研究の方法

研究は以下の項目に分けて行った。

- (1) 変動のリスクの評価方法の確立
 - (2) 変動のリスクを考慮した強化学習の開発
 - (3) 変動に応じてオンラインで解を修正する強化学習の開発
 - (4) 電力系統における変動を考慮した最適化問題への適用
- 項目(1)については、他の研究が取り扱っ

ていない、通行止めのように選択肢が選択不能になる変動に着目し、変動の発生確率および変動による損失と主観的リスクとの関係を考察し、この関係を表現する関数を考案した。

項目(2)については、(1)で考案した主観的リスクによって、目的関数値を目減りさせて評価し、解を求める方法を開発した。さらに、求解と並行して、変動の発生確率を推定する方法も開発した。

項目(3)については、複数変動間の依存関係を条件付き確率で把握し、これに基づいて解をオンラインで変更するアプローチと、さまざまな変動に応じた解候補を発見し、変動に応じて適切に切替えるアプローチの2種の強化学習法を構築した。

項目(4)については、再生可能エネルギー発電出力の変化を変動と捉え、この変動に対処できる発電機運転計画および配電系統運用構成決定を、最適化によって得る方法を考案し、例題に適用した。

4. 研究成果

(1) 変動のリスクの評価方法の確立

変動の発生確率に基づいた主観的リスク評価

選択肢が選択不能になる変動の発生確率のみに基づいて主観的リスクを表現する関数を考案した[雑誌論文]。

降水確率が何パーセントになったら本当に雨が降りそうだと思って傘を持っていくかは個人によって異なる。このような好ましくない現象が発生すると判断する「主観的」な確率を主観的リスクと捉え、客観的確率と主観的リスクとの対応関係を表現する関数を構築した。この関数は、主観を表現する2種類のパラメータを持つ。一つのパラメータは、客観的確率がどの程度の値であれば「主観的」にも発生すると判断するかを表し、もう一つのパラメータは、客観的確率の変化に伴う「主観的」確率の変化の傾きを表現する。いずれも当事者に意味が理解しやすく値の設定が容易なパラメータである。

選択肢が選択不能になる変動の発生確率と変動による損失に基づいた主観的リスク評価

変動の発生確率に加えて、変動によって生ずる損失も考慮して主観的リスクを表現する関数を考案した[学会発表]。

変動によって生ずる損失を、選択不能となった選択肢に対応する目的関数値と他の選択肢に対応する目的関数値との差によって評価する。この損失と変動の発生確率とから、主観的リスクを導く関数を構築した。この関数は、変動の発生確率がどの程度の値になれば主観的リスクが発生するか、変動による損失がどの程度の値になれば主観的リスクが発生するか、および、主観的リスクの変化の傾きをそれぞれ表現するパラメータを含む。

これらのパラメータによって主観の相違を表現することができる。

(2) 変動のリスクを考慮した強化学習の開発

強化学習の代表的手法であるQ学習において、上記(1)で開発した主観的リスクを考慮する方法を開発した[雑誌論文, 学会発表]。

強化学習では、ある状態においてある選択肢を実行すると、その実行結果の評価値に相当する報酬が与えられる。報酬の総和が目的関数であり、これを最大化する各状態における最適な選択肢を見出すことにより、最適な選択肢の系列を発見する。

Q学習は、状態と選択肢の価値を推定する方法である。選択肢を実行するたびに、得られた報酬を用いて価値の推定値が更新される。この更新の際に、主観的リスクに応じて更新量を「目減り」させることにより、主観的リスクを反映した価値を推定する方法を開発した。このようにして、主観的リスクに含まれる変動発生確率と変動による損失の2要素と、報酬で表現される目的関数の合計3要素を一つに統合して取り扱うことができる。

また、強化学習は、学習(最適化)とその結果の利用を並行して行うオンライン型の学習として実施することができる。この場合、学習の過程で、選択肢が選択不能になる事象に遭遇することがある。これは、変動発生確率をオンラインで推定することができることを意味しており、それを実施する方法も開発した。さらに、川が増水して橋が通行止めになった場合、その川にかかる他の橋も通行止めである可能性が高いように、ある選択肢が選択不能になる事象は、他の選択肢が選択不能になる事象に依存することがある。このことを把握するために、他の変動が発生した場合の条件付き確率を推定する方法も考案した。

(3) 変動に応じてオンラインで解を修正する強化学習の開発

条件付き確率を活用したオンライン修正解として得た選択肢系列を実行中に変動を検知した場合、その後の選択肢系列をオンラインで修正する方法を開発した[雑誌論文, 学会発表]。

これは、上記(2)で考案した複数変動間の相互関係を表す条件付き確率の推定法を活用したものである。リスク評価を考慮した価値と考慮しない価値の両方を学習によって推定する。選択肢系列を実行中に、変動を検知した場合、その変動発生を条件とする条件付き確率と、リスク評価を考慮しない価値推定値を用いて、その時点以降に予定されていた行動選択肢のリスク評価込みの価値を評価し直し、新しい価値を用いて選択肢の選択をやり直す。これによって、実行中のオンラ

インでの変動対処を可能とする。

変動に応じた解の切り替え [雑誌論文]

予め多くの解を用意し、変動に応じて切り替えて使用する強化学習法を開発した [雑誌論文]。

開発した方法は、強化学習と進化計算を統合した学習分類子システム (Learning Classifier Systems) と呼ばれる方法の一種である。ある状態においてどの選択枝が最適かを記述したルールの集まりを、進化計算によって構築する。提案手法では、状態の区分を、状態がとり得る値の集合をまんべんなくカバーできるように自動的に定める方法を提案している。さらに、各状態において、適切な選択枝を決定するルールを一つだけでなく、複数用意し、変動が起こった際には、ルールを自動的に適切に切り替えて使用する機構を提案した。この方法は変動を経験するたびに、その変動に応じた解を獲得し、蓄積することができ、同種の変動が発生した際に迅速に対応することができる。

(4) 電力系統における変動を考慮した最適化問題への応用

再生可能エネルギーの変動に対応した発電機運転計画問題

太陽光発電などの再生可能エネルギーによる発電と、複数台のガスエンジン発電機などを併用する小規模電力系統において、再生可能エネルギー出力の変動に対処可能な、ガスエンジン発電機等の運用計画の最適決定問題の解法を提案した [学会発表]。

太陽光発電出力は日射に依存して変動する。電力系統においては消費電力と発電電力は一致していなければならない。したがって、太陽光発電の出力変動に対応して、ガスエンジン発電機はその出力を変動させなければならない。複数台の発電機の運転計画は事前にある一定の範囲の時間帯分を立案する。各時点での発電出力の決定を繰返すため、これは多段決定問題である。計画立案時点では、日射量の正確な予測値は得られないため、日射量の予測値からの変動を想定し、それに合わせたガスエンジン発電機の出力調整余裕を持たせつつ、発電機の燃料費を最小にする必要がある。特に、太陽光発電電力は予測値から増加することも減少することもあるため、その両方に対応可能とする必要がある。このような運転計画を立案する方法を提案し、モデル電力系統を用いたシミュレーションによってその有効性を確認した。

太陽光発電出力の変動に対応した配電系統運用構成決定問題

配電系統において事故が発生した後の、復旧時の運用構成を決定する際に、太陽光発電出力の変動を適切に考慮する方法を提案した [雑誌論文, 学会発表]。

太陽光発電装置が配電系統に多く接続されるようになっている。配電系統において事故が発生した場合、安全性の観点から太陽光発電装置は配電系統から切断される。事故後、配電系統内の開閉器 (スイッチ) の開閉を変更することにより、電力供給経路を変更し、事故によって電力供給が停止していた地域に電力供給を再開する。この開閉器の操作によって決定される電力供給経路は運用構成と呼ばれ、できるだけ多くの地域に電力供給を行うべく決定される。電力供給再開時には、需要と等しい電力を供給する必要があるが、太陽光発電装置が切断されているため、事故前にこれら装置が供給していた電力も合わせて、変電所から供給する必要がある。しかし、太陽光発電出力は実時間で計測されていない。実時間で計測されるのは変電所から供給している電力のみである。このため、供給再開時に供給すべき電力の大きさが不明である。

本研究では、事故前の太陽光発電出力の最大値と最小値を定め、この範囲で太陽光発電出力が変動していても、支障なく電力供給が再開できる運用構成決定法を提案した。ここでは、ロバスト最適化の考えを導入し、太陽光発電出力の変動幅全域で制約条件を満たし、目的関数を最適にする解を求めている。この方法を実規模の配電系統モデルに適用し、太陽光発電出力の変動に対処できることを確認している。

<引用文献>

- [1] R. A. Howard and J. E. Matheson, Risk-Sensitive Markov Decision Processes, Management Sciences, Vol.18, No.7, 356-369 (1972)
- [2] 佐藤, 木村, 小林, 報酬の分散を推定する TD アルゴリズムと Mean-Variance 強化学習の提案, 人工知能学会論文誌, 16 巻, 3 号 F, 353-362 (2001)
- [3] K. Sladký and M. Sitař, Risk Sensitive and Mean Variance Optimality in Markov Decision Processes, Proc. 26th Int. Conf. Mathematical Methods in Economics 2008 (2008)

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計5件)

Danilo V. Vargas, Hiroataka Takano, Junichi Murata, Novelty-Organizing Team of Classifiers in Noisy and Dynamic Environment, Proc. 2015 IEEE Congress on Evolutionary Computation, 査読有, 2015, 2937-2944

Danilo V. Vargas, Hiroataka Takano, Junichi Murata, Novelty-Organizing Team of Classifiers - a

Team-Individual Multi-Objective
Approach to Reinforcement Learning,
査読有, Proc. SICE Annual Conference
2014, 2014, 1785-1792

Hiroataka Takano, Junichi Murata,
Yukino Maki, Makoto Yasuda, Improving
the Search Ability of Tabu Search in
the Distribution Network
Reconfiguration Problem, Journal of
Advanced Computational Intelligence
and Intelligent Informatics, 査読有,
Vol.17, No.5, 2013, 681-689

Takuya Etoh, Hiroataka Takano, Junichi
Murata, Reinforcement learning
approach to multi-stage decision
making problem with changes in action
sets, Artificial Life and Robotics が
査読有, Vol.17, No.2, 2012, 293-299

高野 浩貴 (TAKANO, Hiroataka)
九州大学・大学院システム情報科学研究院
・助教
研究者番号 : 5 0 4 3 5 4 2 6

〔学会発表〕(計 12 件)

高野 浩貴, 大賀 博文, 村田 純一, 樺澤
明裕, PV の不確実性に対応した配電ネッ
トワーク復旧問題の一解法, 平成 26 年電
気学会電力技術・電力系統技術合同研究
会, 2014 年 9 月 25 日, 大阪府立大学(大
阪府・堺市)

大賀 博文, 高野 浩貴, 村田 純一, 飯坂
達也, 樺澤 明裕, 太陽光発電の時間毎最
大可能出力に基づくロバスト最適化を用
いた配電系統事故復旧構成決定法, 平成
26 年電気学会電力・エネルギー部門大会,
2014 年 9 月 12 日, 同志社大学(京都府・
京田辺市)

江藤 拓也, 大坪 良介, 高野 浩貴, 村田
純一, 確率的変動を伴う意思決定問題に
対する主観的リスク評価を考慮した強化
学習法, 第 23 回インテリジェント・シス
テム・シンポジウム, 2013 年 9 月 25 日,
九州大学(福岡県・福岡市)

高野 浩貴, 村田 純一, 張 鵬, 橋口 卓
平, 合田 忠弘, 飯坂 達也, 中西 要祐,
再生可能エネルギーの不確実性に対応し
た小規模グリッドの電源運用最適化に関
する基礎検討, 平成 24 年電気学会電力・
エネルギー部門大会, 2012 年 9 月 12 日,
北海道大学(北海道・札幌市)

〔その他〕

ホームページ等

[http://cig.ees.kyushu-u.ac.jp/~murata/M
SDP.html](http://cig.ees.kyushu-u.ac.jp/~murata/MSDP.html)

6. 研究組織

(1) 研究代表者

村田 純一 (MURATA, Junichi)

九州大学・大学院システム情報科学研究院
・教授

研究者番号 : 6 0 1 9 0 9 1 4

(2) 連携研究者