

## 論文の内容の要旨

論文題目    Online Optimization with Limited Information  
                  (限られた情報を用いたオンライン最適化)

氏    名    伊藤 伸志

組合せ最適化や凸最適化などの最適化技法は幅広い領域に適用されている一方、現実の問題では、意思決定の時点で具体的な評価関数が与えられないことがあり、この場合は最適化の技法を直接適用できない。その具体例として、オンライン予測の問題が挙げられる。オンライン予測問題は、逐次的に得られるデータの標本とラベルの情報をもとに次の標本のラベルを予測する問題であり、スパムメールの分類や商品レコメンデーションなど幅広い応用をもつ。ここで、ラベルの予測においては各標本の特徴量ベクトルと呼ばれる数値的な情報を利用でき、この特徴量とラベルの間の関係性を学習することが重要である。教師あり学習問題の多くは凸最適化問題として定式化できる一方、オンライン予測問題では逐次的にデータが与えられるため、目的関数は常に変化し、予測値の出力時点では目的関数を知ることができない。同様に、逐次的に資産配分を決定するオンラインポートフォリオ選択問題では、目的関数を定義する価格指数(投資の費用対効果)が投資を実施したあとにはじめて観測されるため、解を出力する前に目的関数を知ることができず、目的関数は投資のラウンドごとに変化する。

上述のように、評価関数が与えられる前に意思決定しながら逐次的に戦略を改善していく問題を扱う枠組みとして、オンライン最適化(逐次的最適化)が知られている。この枠組みでは、複数の意思決定ラウンドにわたって解の出力と目的関数の観測を繰り返しながら、目的関数値の総和の最小化を目指す。ここで目的関数の情報は解を出力した直後に与えられる。アルゴリズムの性能は、最善の固定戦略(すべてのラウンドで一貫した戦略)で達成される目的関数値の総和と、達成した目的関数値の総和の差で評価される。この差をリグレットとよぶ。オンライン最適化の研究領域における中心的な目標は、リグレットをできる限り小さくする方法を構築することである。この領域の重要な結果として、目的関数が凸の特殊ケースに対しては、ラウンドあたりのリグレットの値が  $0$  に収束する(劣線形なリグレット上界をもつ)アルゴリズムが存在する。このことは、ラウンドを重ねるごとにアルゴリズムの出力が最善の固定戦略と同等の水準に近づいていくことを意味する。

オンライン最適化は、意思決定の段階で利用可能な情報が限られている状況に対応するための有望なアプローチといえるが、現実世界の応用においては、利用できる情報がさらに制限されることがある。たとえば、オンライン予測を医療診断に適用するとき、各サンプルは患者に対応し、ラベルが疾病リスク、特徴量ベクトルが医学的検査の結果に対応する。このとき、ひとつの標本の特徴量をすべて観測すること、つまり一人の患者にすべての検査を行うことは患者への負担が大きく、現実的に観測可能な情報は一部の医学的検査の結果に限られる。このことは、各ラウンドにおいて特徴ベクトルの一部の成分のみ観測できる状況に対応し、目的関数に関する情報の一部しか得られないことを意味する。同様の問題はオンラインポートフォリオ選択においてもあらわれる。たとえば、研究開発費への投資や広告費への投資を考えると、投資しなかった施策についてのフィードバックは得られないため、意思決定の後にも目的関数の全体像を知ることはできない。このように、現実の応用においては、意思決定のあとにも目的関数の情報が完全には得られないことがあり、このとき既存のオンライン最適化の手法を直接適用することはできない。

本論文では、観測可能な情報が制限されている状況に対応するいくつかの問題設定、とくに (1) 組合せ最適化問題を含む線型関数の最適化、(2) 劣モジュラ関数最小化、および (3) 凸最適化のそれぞれに対し、オンラインアルゴリズムの構築とその効率性の解析、また困難性の解析を与える。アルゴリズムの効率性は、目的関数値のよさに対応するリグレット上界と、計算効率のよさに対応する計算量上界のふたつの観点から評価する。困難性についても同様に、最悪ケースリグレット下界と計算複雑性のふたつの観点から評価する。

### (1) バンディット線型最適化問題

目的関数が線型のオンライン最適化問題で、選択した解における目的関数値のみがフィードバックとして得られる問題を扱う。このようなフィードバックをバンディットフィードバックとよび、対応する問題をバンディット最適化問題とよぶ。この問題設定は、バンディット最短経路問題などの、実行可能領域が離散的な集合の場合も含む。例えばバンディット最短経路問題では、事前に枝重みが未確定の有向グラフと始点・終点が与えられたうえで、複数のラウンドにわたって、始点から終点への経路の選択と、その経路長のフィードバックが繰り返される。ここでグラフの枝重みはラウンドごとに変化する。通常のオンライン最適化では各ラウンドにおいて全ての枝の重みがフィードバックされるのとは対照的に、バンディットフィードバック設定では選択した経路の枝の重みのみしか観測できないため、バンディット設定の方が相対的に難しい問題設定といえる。

この問題に対して、リグレット上界の観点では最適なオーダーを達成するアルゴリズムが知られている一方で、そのようなアルゴリズムは最悪時に指数時間の計算を要し、多項式時間アルゴリズムで最適なオーダーを達成できるかは未解決だった。

本論文では、最適なりグレット上界を達成する初の効率的なアルゴリズムを構築することで、上述の未解決問題に対して解を与える。ここで、効率的とは、対応するオフライン

最適化問題(バンディット最短経路問題の場合は、通常の最短経路問題が対応)のアルゴリズムを呼び出す回数が問題サイズの多項式で上から抑えられることを意味する。この結果は、オフライン最適化が多項式時間で解けるという、ある意味で必要最小限の仮定のみでバンディット最適化問題も多項式時間で解ける、つまり、ある意味でオフライン最適化とバンディット最適化の計算複雑性が多項式時間帰着の意味で等価であることを示唆する。

## (2) 確率的な関数値評価オラクルを用いた劣モジュラ関数最小化

劣モジュラ関数最小化問題は組合せ最適化における重要な問題であり、強多項式時間可解であることが示されて以降も、現在までに様々なアルゴリズムが提案されてきた。これらの既存研究は、具体的な変数の値を指定したときに目的関数値を返すオラクルにアクセスできるという仮定の計算モデルに基づく。一方で一部の応用例では、厳密な評価オラクルの存在を仮定できず、実際に観測できる目的関数値にノイズが含まれる状況が存在する。そこで本研究では、確率的なノイズを伴う評価オラクルを用いた劣モジュラ関数最小化問題を考える。

この問題設定は、バンディットフィードバック・オンライン・劣モジュラ関数最小化問題の特殊ケースと見なすことができる。このより一般的な問題に対しては、アルゴリズムと複雑性の解析が与えられている一方で、達成可能な最適ナリグレットのオーダーは未だ未解決である。

本研究では、確率的な関数値評価オラクルを用いた劣モジュラ関数最小化に対する多項式時間アルゴリズムを提案し、アルゴリズムが出力する解の誤差評価を与える。加えて、任意のアルゴリズムに対する誤差下界を示すことで、提案アルゴリズムの誤差の一定以上の改善がある意味で不可能である、つまり提案アルゴリズムが最適に近いことを示す。

## (3) 限られたフィードバックを用いたオンライン凸最適化

バンディットフィードバックを用いたオンライン凸最適化(バンディット凸最適化)問題に対する新しいアルゴリズムを提案する。提案手法は、任意の凸領域上の問題に適用でき、強凸性と平滑性の仮定のもとで最適ナリグレット上界を達成する初のアルゴリズムである。既存手法は、制約なしの問題でしか最適ナリグレットを達成できない、また必ずしも多項式時間で実行できないという課題点をもっていたが、本研究はこれらの課題を解決している。

加えて、オンライン凸最適化の特殊ケースであるオンラインポートフォリオ選択問題で、一度に投資できる資産の組合せと観測可能な情報が制限されている状況を考え、これに対するアルゴリズムを提案する。アルゴリズムが達成するナリグレット上界と、任意アルゴリズムに対する最悪ケースのナリグレット下界を与えることで、提案アルゴリズムの性能が最適に近いことを示す。さらに、計算複雑性クラスNPがBPPに含まれないという仮定のもとで、提案アルゴリズムの時間計算量を一定以下に削減できないことを示す。