

NTTコミュニケーション科学基礎研究所
上席特別研究員

澤田 宏 Hiroshi Sawada

「相関」と「複素数」を駆使したアルゴリズムで、「音源分離」と「ハードウェア向けのニューラルネットワークの訓練方法」の異なる分野の研究に挑む

解決したい問題を数式で表し、問題に対する最適な解を求める最適化問題は、従業員の出勤の頻度や時間帯が均等になるシフトの作成や、利益を最大化するための商品の生産量の決定など、あらゆる方面で応用されており、さまざまなアプローチがあります。AI（人工知能）の普及に伴い、最適化問題の解決にAIを活用する機会も増えています。こうした応用だけではなく、AIの機械学習そのものにおいても最適化が活用されています。概念は古くからある最適化のアプローチを、「音源分離」「ハードウェア向けのニューラルネットワークの訓練方法」の研究に取り入れて新たな成果を創出する、NTTコミュニケーション科学基礎研究所 澤田宏上席特別研究員に、「音源分離」「ハードウェア向けのニューラルネットワークの訓練方法」の研究アプローチとその成果、自分の技術の軸足をもって他分野との接点をとることの重要性を伺いました。



「音源分離」の研究で取り組んだ「自然勾配法」を「ハードウェア向けのニューラルネットワークの訓練方法」の研究に適用

現在、手掛けていらっしゃる研究について教えていただけますでしょうか。

「音源分離」と「ハードウェア向けのニューラルネットワークの訓練方法」に関する研究に取り組んでいます。

音源分離に関しては約20年にわたり取り組んできています。前回（2022年3月号）のインタビューでは、データや信号などの情報源の構造や特徴をうまくとらえる非負値行列因子分解（NMF：Nonnegative Matrix Factorization）と、データや信号が観測系を通じてどのようにセンサで観測されたかを推定する独立成分分析（ICA：Independent Component Analysis）を統合したILRMA（Independent Low-Rank Matrix Analysis：独立低ランク行列分析）による音源分離を紹介し、分野ごとに毎年5名選出されるIEEE（Institute of Electrical and Electronics Engineers）Signal Processing Society Distinguished Lecturer（期間2年）のうちの1人として選出された話をしました。Lecturerについてはコロナ禍により実質的な講演が開催され

たのが2023年で、オーストラリア、ニュージーランド、ドイツの研究機関でライブデモを交えてILRMAを紹介し、特にドイツでは音響信号処理に関して権威のある研究者が多数参加しており、30分を超える質疑応答が飛び交うなど、大好評でした。

さて、ILRMAは音源の数と同数かそれより多いマイクで拾った音で、部屋の残響が比較的短い場合の音源分離をするものですが、その後、音源数がマイクの数より多く、かつ部屋の残響が長い状況に有効な新たな音源分離手法、mfFCA（multi-frame Full-rank spatial Covariance Analysis）の提案を行い、国際会議ICASSP2022で発表し、IEEE/ACM Transactions on Audio, Speech, and Language Processingに採録されました。図1に示すとおり、私はそれまでICAやILRMA、FCAなど残響が比較的短い状況で有効な音源分離手法に取り組んできました。一方、同じ研究所には、残響除去WPE（Weighted Prediction Error）を

	音源の数 ≤ マイクの数 	音源の数 > マイクの数 
残響が比較的短い	ICA ILRMA	FCA
残響が長い	WPE	mfFCA

図1 ILRMAとmfFCAの比較

研究開発した中谷智広上席特別研究者らがいますので、彼らと相談しながら残響が長い問題にもチャレンジし、異なる時間のフレーム間での相関に着目してモデル化するmfFCAを提案できました。

「ハードウェア向けのニューラルネットワークの訓練方法」はどのような研究でしょうか。

ニューラルネットワークの訓練（機械学習）は、一般的にバックプロパゲーション（誤差逆伝播法）という学習アルゴリズムで行われています。入力層、隠れ層、出力層から構成されるニューラルネットワークにおいて、出力と目標値の誤差をネットワークの各層を逆向きに伝播させることで、これらの層間の重みを調整し、効率的に目標値に近い出力を生成できる（性能向上）ようにするのがバックプロパゲーションで、これにより、ニューラルネットワークは複雑なパターンを認識し、予測や分類などの機能を実現することができます。

さて、最近では、ニューラルネットワークの電力消費が増大していることが話題になっていますが、光デバイスのようなハードウェアそのものにニューラルネットワークを実装することでこれに対応することが期待されています。一方、例えば製造におけるばらつきや、わずかな温度変化でも特性が変わるといった、ハードウェア固有の特徴があり、こうしたばらつきや特性の変化により、バックプロパゲーションがどうしても不正確なものになってしまいます。

私の研究テーマは、このように一般的な訓練方法が通用しない状況に対して、その課題を解決する「ハードウェア向けのニューラルネットワークの訓練方法」です。1つのアプローチとして、Zeroth-Order Optimization (ZOO: ゼロ次最適化) という、訓練においてパラメータを初期状態から少しずつ変化させる（摂動を与える）中で、損失関数を減少させるという良い結果が出た方向にパラメータを変化させていく方法があります。ZOOの手順を図2に示します。はじめに摂動を何個か（Q個とします）生成します。次に各摂動について、実際にハードウェアに実装されたニューラルネットワークを用いて、損失関数の変化量を評価します。こうして摂動とその評価値がそろると、近似勾配が計算でき、それに基づいてパラメータを更新します。このループの繰り返しで訓練を進めます。

通常、ニューラルネットワークのパラメータは行列の形式であり、その行列に途中経過のベクトルを掛けてさらにベクトルを得るといった計算をしています。ところが、ニューラルネットワークをハードウェアで直接実装するとき、例えば光デバイスに実装する場合、光の位相をパラメータとして連続的にずらすことで、任意のユニタリ行列（実数による直交行列を複素数へ拡張したもの）を実装し、それぞれのユニタリ行列を面的かつ層状に広げることでより大きなユニタリ行列を計算します。これは明らかに通常の行列形式のパラメータとは異なっています。特に課題となるのは、パラメータが層状に配置されるため、光の経路において多くのパラメータが寄与し、それらが互いに絡み合ってしまうという状況です（図3）。

この研究に関する最初の成果として、自然勾配法（Natural Gradient Method）の考え方に基づく新たな近似勾配の計算方法を考案し、2024年6月に米国サンフランシスコで開催された、電子設計自動化分野の世界トップカンファレンスである、Design Automation Conference (DAC) 2024で発表しました。自然勾配法は最適化のアルゴリズムの1つで、どのようにパラメータが配置されるかという構造の影響を受けにくい方法であるため、光デバイスに実装されたニューラルネットワークの訓練において、効率化や精度向上を図ることができます。私は、音源分離におけるICAに取り組む中で自然勾配法を標準的な手法として用いていましたし、近年ではバックプロパゲーションを用いる通常のニューラルネットワークの訓練でも使われていて、今回のZOOによる訓練でも使えるのではないかと発想し研究に取り組みました。通常のニューラルネットワークの訓練で自然勾配法を用いようとす

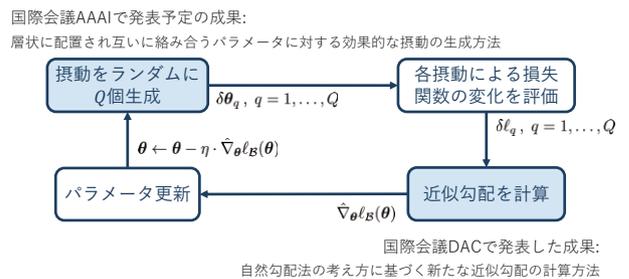


図2 ゼロ次最適化ZOOの手順と成果の位置付け

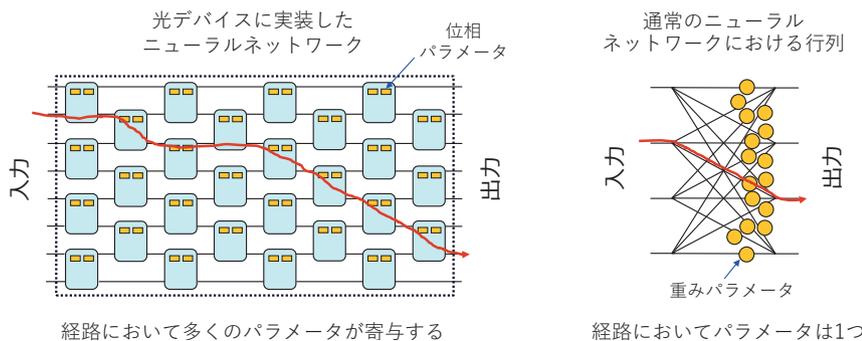


図3 光の経路において互いに絡み合うパラメータ（左）と通常の行列形式のパラメータ（右）

ると、パラメータ数×パラメータ数の巨大な行列の逆行列を計算する必要があるので、現実的には数千以上のパラメータ数でその計算は不可能になり、いくつかの近似手法が提案されています。一方、本成果では、摂動数 Q ×摂動数 Q の行列の逆行列を計算する手法を提案しましたので、これにより摂動数 Q が現実的には数百程度となるため、少ない計算オーバーヘッドで実行できるものとなっています。

音源分離についてはmfFCAで一区切りが付き、現在は研究の軸足を「ハードウェア向けのニューラルネットワークの訓練方法」にシフトしてきているのですが、前回紹介した、次世代の研究者育成をめざして新入社員や入社3年目あたりの社員向けの技術講座「機械学習」も継続しています。始めて8年目となりますが、ここ数年はほぼ同じ内容〔最初の導入部分で最近の生成AI（人工知能）ブームなどには触れますが〕のため、説明が安定して分かりやすいためか、「本質的な説明や踏み込んだ内容を聞けた」と好評です。こうした取り組みを通じて、私自身もニューラルネットワークについてかなり理解が深まり、今回の研究においてもこれが活かしていることはいうまでもありません。

最近新しい成果が出て国際会議で発表するというホットな話題があるそうですね。

最新の情報なのですが、次の成果としてAI分野の最大級の学会である、AAAI (Association for the Advancement of Artificial Intelligence : 米国人工知能学会) に投稿していた論文が2024年12月に採録されました。そして、2025年2月末から始まる国際会議、AAAI-25で発表予定です。本成果の位置付けは図2において摂動をランダムに生成する部分にあたります。ZOOにおいて通常は、多変量の標準正規分布からランダムに摂動を生成します。これはパラメータの独立性が高い場合には有効な手法です。しかし、すでに述べたとおり、光ニューラルネットワークではパラメータが互いに絡み合うため、どのようなランダムから摂動を生成すべきか、ということにも配慮すべきであると思いつきました。本成果として、面的かつ層状に構成されるユニタリ行列に対して、すべてのパラメータ間の相関を計算し、その相関を打ち消すように摂動を与える方法を考案しました。

自分の技術の軸足を定め、他の分野の人との接点をとることで成果を生み出す

研究者として心掛けていることを教えてください。

私の研究は、デジタル計算機的设计自動化、音源分離、ハードウェアを中心とした光ニューラルネットワークといったように、分野が変わってきています。途中で研究部長や企画部長も経験しており、そのときは自身の専門とは異なるほかの研究の話聞く機会もありました。こうした中で、専門とする技術を軸足として、極力その軸足となる技術に何らかのかたちでかかわってみたいと

の思いを持っています。全く異なる分野で研究を進める場合、本当に成果を出せるのか不安に思うこともあります。しかし、軸足の技術が共通であれば、こうした不安も払拭されます。軸足となる共通の技術という意味では、例えば前述の設計自動化、音源分離、光ニューラルネットワークいずれにおいても、共通的にコアとなっているところは、「相関」と「複素数」を駆使したアルゴリズムなのです。

興味を持った分野をテーマにしようか考えるときに、共通の技術を確認して、それをベースに次のテーマを設定するというアプローチをしている、つまり興味を持つ分野の入り口が共通の技術となっているため、これまでの研究テーマは結果として共通の技術でつながってきているのです。

そして、異なる分野の人との接点を意識し、できれば連携をしていきたいと思っています。ここ10年ぐらい、AIや、ビッグデータ・機械学習等が注目されるようになり、さまざまな分野で利用されるようになりました。私自身、信号処理や機械学習を中心に研究をする中で、研究所をまたいだコロキウム等に参加してきました。信号処理や機械学習の研究の中だけでは、「相関」と「複素数」等に関する数式を操作するだけの範囲にとどまりそうだったのですが、応用分野の方々とコロキウムで意見交換や議論を進めることで、研究の出口に対して多くの知見を得ることができました。最近は音声音響、言語、ハートタッチング、人間研究等テーマのコロキウムが研究所をまたぐかたちで同時多発的に発生し、議論の成果とともに収束しており、非常に良いことだと思います。私は、過去にはコロキウムを仕掛ける側でもありましたが、常に仕掛けるわけではなく、興味のある分野について積極的に参加することが重要だと思います。

得意分野を軸足に補完関係、協力関係を構築し、研究を良い方向へ導く

後進の研究者へのメッセージをお願いします。

国際会議などにおける論文採択は非常に難度が高いのですが、挑戦し続けていくことは大切だ、と前回話しました。それに加えて、長期間研究者を続けていると常に1つのテーマで研究しているわけではなく、テーマが変わることが一般的です。また、ほかの分野の人との共同研究も今後ますます重要になるばかりではなく、その機会は増えていくと思います。こうしたときに、自分の軸足となる技術を持つことが重要で、これにより新しい分野で自分が活躍する場を見出すことができるのではないのでしょうか。今後技術分野はさらに細分化されていくことは十分予想されるのですが、こうした環境下ですべての分野において実力を十分発揮できるスーパー研究者がいるわけではありません。だからこそ、「これだけは」というような自分の得意分野を軸足として、それをそれぞれが示すことで補完関係や協力関係を構築することができ、それが相乗効果を生んで研究を良い方向へ進めることができるのではないのでしょうか。