

ベイズ最適化とオートエンコーダを用いた最適化手法の開発

DEVELOPMENT OF OPTIMIZATION METHODS USING BAYESIAN OPTIMIZATION AND AUTOENCODERS

森田泰之^{#, A)}, 福田光宏^{B)}, 依田哲彦^{B)}, 中島悠太^{C)}, 鷲尾隆^{D)}

Yasuyuki Morita^{#, A)}, Mitsuhiro Fukuda^{B)}, Tetsuhiko Yorita^{B)}, Yuta Nakashima^{C)}, Takashi Washio^{D)},

^{A)} RIKEN Nishina Center for Accelerator-Based Science

^{B)} Research Center for Nuclear Physics, Osaka University

^{C)} Osaka University Institute for Dataability Science

^{D)} Sanken Osaka University

Abstract

One of the major challenges in accelerator operation is the development of highly reproducible operation control methods. In recent years, automatic control methods using machine learning have been developed, and results have already been achieved in automatic control using the Gaussian process regression model. However, Gaussian process regression models are limited to optimizing 15-20 parameters, and it is impossible to comprehensively tune the entire accelerator. To solve this problem, we have developed a method that combines a dimensionality reduction method with an autoencoder. By utilizing a Gaussian process regression model in the latent variable space after dimensionality reduction, the number of parameters to be adjusted can be reduced to a smaller number than the actual number of parameters, which is expected to allow adjustment of a larger number of parameters. In this study, some parameters of the ECR ion source were optimized by applying this method to the actual operation of the source. As a result, the expected performance of the ECR ion source could not be achieved in the actual operation, mainly because the parameters are very sensitive. Based on this result, we will continue to verify whether this method can be used effectively for ECR ion sources by taking measures such as excluding parameters that are sensitive to the objective from the dimensionality reduction.

1. はじめに

近年機械学習による自動制御が進んでおり、SACLAでのビーム輸送系の最適化[1]や RCNP での ECR イオン源の最適化[2]のようにベイズ最適化[3]を用いた最適化が成果を上げている。しかし、本手法では 15 パラメーター程度の最適化が限界であり、ビームラインの包括的なチューニングをするとなると扱えるパラメーター数が足りない。その為、ビームラインを診断機事ごとに分割するなどして部分的な最適化調整を繰り返している。

ビームラインを分割して最適化する手法の場合、ある部分の最適化が全体の最適化と一致しているとは限らない。したがってローカルマキシマムにとらわれる可能性が高いことが課題といえる。そこでオートエンコーダ[4]による次元削減とベイズ最適化を併用した最適化手法の開発を行っている。本研究ではその一環として 10 GHz の ECR イオン源'NANOGANTM'[5]を用いた実機試験の結果と課題に関して報告を行う。

2. 最適化手法

本研究ではオートエンコーダによって次元削減を行い、削減した次元内で最適化を行うことで膨大なパラメーター数の最適化の実現を目指す。

2.1 オートエンコーダの性質

オートエンコーダーとは入力と出力が一致するように学習を行う機械学習の一種である。入力層と出力層の間

に‘潜在変数’を出力する層が設けられている。入力層から潜在変数出力層までの部分をエンコーダーと呼び、潜在変数出力層から出力層までの部分をデコーダーと呼ぶ(Fig. 1)。この潜在変数は入力層の値を復元するために必要な情報を残した状態で、別空間で表現するものである。この潜在変数の次元数を入力層や出力層の次元数よりも下げることで、次元削減を行うことができる。ただしこれは潜在変数が入力層の値を復元するために必要な情報を残していることが前提となるため、過剰に次元を減らすと出力層で入力層の値を正しく復元できなくなる。その為、設計時には出力層で入力層の値が再現されていることを損失関数と呼ばれる関数で評価を行い、損失関数が十分に小さくなる範囲で次元数を減らすことで入出力の次元を下げるができる。

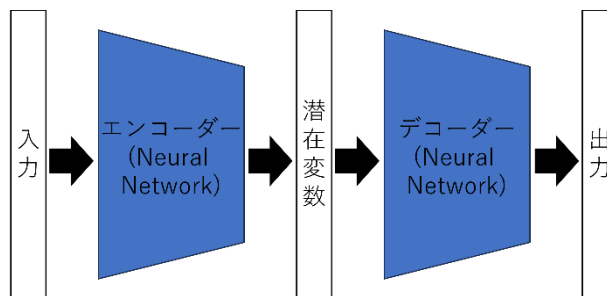


Figure 1: The conceptual diagram of autoencoder.

2.2 最適化手法

事前にオートエンコーダを学習させることで潜在変数から元の入力値を復元することができる。この性質を活

[#] yasuyuki.morita@riken.jp

用し、潜在変数空間を最適化パラメーター空間として (Fig. 2) ベイズ最適化による最適化を行う。これにより本来設定するパラメーターの次元数よりも少ないパラメーター空間での最適化することができる。

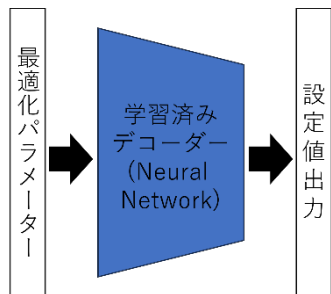


Figure 2: The conceptual diagram of tuning method using latent variable.

本手法を実機で自動制御実験するためには、安全対策は必須である。最適化パラメーターを学習データとは全く異なった範囲で設定してしまった場合、デコーダーによる出力はどのような値になるかわからない。そこで最適化パラメーターの範囲は、学習データの潜在変数空間内での分布をもとに決定を行う。また、設定値の出力が安全に運転可能な範囲を逸脱している場合には設定値を変更しないように設計した。

3. オートエンコーダの設計

3.1 学習データの作成

本研究を行うにあたり、まずはオートエンコーダの学習を行うためのデータが必要である。そこで先行研究[2]と同じ手法で 10 GHz ECR イオン源 'NANOGAN' の最適化を行い、学習用のデータを作成した。この時最適化したパラメーターは RF パワー、Rf 周波数、ガスバルブ、中間電極電圧の 4 つである。その後最適化の嘉永で得られたデータのうち、ビーム輝度が $1E-5(\text{mA}/(\text{mm})^2)$ のデータのみを学習データとした。ただしビーム輝度はビームの強度(I)と x-x'位相空間及び y-y'位相空間のエミッタンス(ϵ_x, ϵ_y)を用いて式 1 であらわされる値である。

$$\text{Brightness} = \frac{I}{\epsilon_x \epsilon_y} \quad (1)$$

今回の学習データでビーム輝度の下限値を設けた理由はパラメーター間に相関を持たせるためである。[2]の最適化では、4 パラメーター間の相関はなく、無作為に選ばれたパラメーターが存在する。そのようなパラメーターでオートエンコーダを学習させた場合、元のパラメーター間に相関がないので次元を減らすことができない。そこでビーム輝度に制限をかけることでパラメーター間に相関を持たせることが目的である。

学習用データ取得は 2023 年 4 月 26 日、27 日、28 日の三日間にわたり行った。その内 26 日に行った計 400 回のパラメーター調整と 28 日に行った計 300 回のパラメーター調整のうち、ビーム輝度が $1E-5(\text{mA}/(\text{mm})^2)$ となったデータ 173 個を学習データとした。27 日に行った計 100 回のパラメーター調整のうち、ビーム輝度が $1E-5(\text{mA}/(\text{mm})^2)$ となった 22 個のデータは学習には使用せずにテスト用のデータとして用いた。

3.2 オートエンコーダの学習

今回、オートエンコーダでは 4 パラメーターを 2 次元の潜在変数に変換し、4 次元に復元するモデルの作成を行った。ライブラリとして Tensorflow[6]を活用し、エンコーダーとして 5 層、デコーダーとして 5 層の中間層をもつモデルを作成した。また、入力層では実際の各パラメーターの最適化可能範囲が 0 から 1 にするために、式 (2) に示すように規格化を行った。

$$\text{Input} = \frac{\text{set value} - \text{minimum value}}{\text{maximum value} - \text{minimum value}} \quad (2)$$

ただし、set value は設定値であり、maximum value、minimum value はそれぞれ各設定値の最大値と最小値である。出力層でも同様に規格化した値が出力されるため、実際に設定する際には式(3)によって設定すべき値に戻したうえで設定する。

$$\begin{aligned} \text{set value} &= \text{Output} \\ &\times (\text{maximum value} - \text{minimum value}) \\ &+ \text{minimum value} \end{aligned} \quad (3)$$

学習後に学習データおよびテストデータの入力値と出力値の関係をプロットした図を Figs. 3-6 に示す。Figure 3 はガスバルブの設定値、Fig. 4 は RF 信号源のパワー Fig. 5 は RF 周波数、Fig. 6 は中間段電極の電圧に関するプロットである。これらの図は全て横軸がオートエンコーダの入力層での値、縦軸が出力層での値である。ガスバルブ及び中間段電極の設定可能範囲はバッファータンクのガス残量やイオン源の状態によって変動しうるため、すべてをカバーできる範囲で入力層の規格化を行った。今回のイオン源やバッファータンクの容量では設定値を上限いっぱいまで上げられなかったため、Fig. 3 の 11600 steps 以上の領域のデータが不足する結果となった。

各設定値の復元に関して誤差の RMS を算出すると Table 1 のようになり、完全に問題がないとは言えないまでもいづれも設定可能範囲に比べ小さい。

Table 1: The Error of Prediction Value for Each Parameters

Parameter	RMS
Gas Valve	96
RF Power	0.10
RF Frequency	0.013
Intermediate electrode	1.0

本来はオートエンコーダの層数や各層内部の計算式などのモデルの最適化を行い、より小さな誤差を目指すべきである。しかし今回は実験時間の都合上、設定可能範囲に比べ小さいという状態で実験を行った。

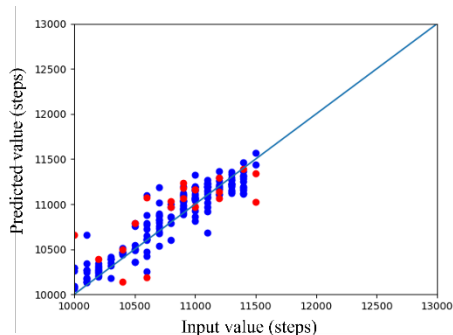


Figure 3: Relationship between input and predicted value of gas valve.

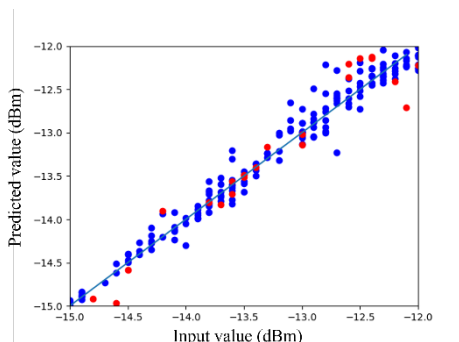


Figure 4: Relationship between input and predicted value of RF power.

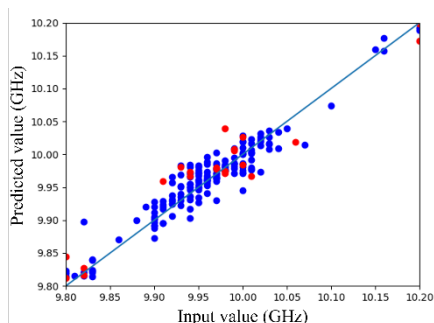


Figure 5: Relationship between input and predicted value of RF frequency.

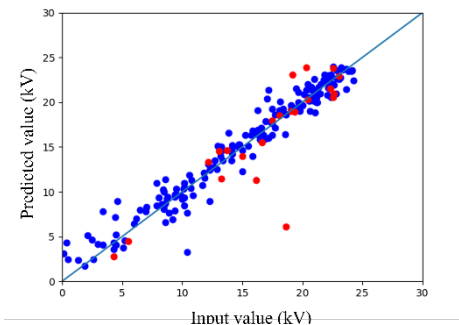


Figure 6: Relationship between input and predicted value of intermediate voltage.

4. ECR イオン源の最適化の結果

4.1 次元削減と併用した最適化

まずはオートエンコーダによる次元削減とベイズ最適化を併用した最適化を行った。本最適化ではパラメーターの変更を最大で 80 回行い、ビーム輝度が最大となったパラメーターに戻して終了とした。本手法ではパラメーター設定値が設定可能範囲を超える可能性があり、装置の破損を防ぐために設定可能範囲を超えた場合はパラメーターを設定せずに輝度を 0 として評価する安全策が施されている。その為先行研究と同様にパラメーターの変更回数の最大値を 100 回とすると先行研究よりも最適化調整時間が長くなるのが想定される。そこで 8 割に当たる 80 回を最大値とした。

本手法で最適化された後のビーム輝度を Table 2 に示す。今回は全く同様の最適化を 2 度行い、再現性の確認も行った。この 2 度の結果はおおよそ一定であり、十分再現性があることが確認された。

Table 2: The Optimization Results of ECR Ion Source using Dimensionality Reduction

	Brightness (mA/(mm) ²)
AE+BO : Test 1	1.23×10^{-5}
AE+BO : Test 2	1.36×10^{-5}

4.2 次元削減を行わない最適化

続いて次元削減を行わない従来のベイズ最適化を用いた最適化手法を 2 度行った。この最適化でのパラメーター変更回数は、先行研究と同様に最大 100 回としている。100 回のパラメーター変更後、最大の輝度を記録したパラメーターに設定しなおした際のビーム輝度を Table 3 に示す。

Table 3: The Optimization Results of ECR Ion Source without Dimensionality Reduction

	Brightness (mA/(mm) ²)
Only BO : Test 1	2.17×10^{-5}
Only BO : Test 2	2.21×10^{-5}

次元削減を用いずに最適化した結果、オートエンコーダによる次元削減を併用した手法よりも高輝度となった。

4.3 考察

上記の結果より、次元削減を用いない方が最適化の結果が良いことが分かった。その原因として学習モデルの最適化が不十分であることが考えられる。特に ECR イオン源の最適化では RF 周波数のようにわずかな変化に敏感に反応し、ビーム輝度を落とす原因になりうるパラメーターが存在する。そこで次元削減を使わない最適化での最終設定値にどこまで近づけることができるかを 2 回の最適化結果それぞれに対して検証した。その結果をそれぞれ Table 4, 5 に示す。

Table 4: Closest Value to the Setting Value of Only BO Test 1 in the Latent Variable Space

	Gas valve (steps)	RF Power (dBm)	RF Frequency (GHz)	Intermediate electrode (kV)
Only BO Test 1	10600	-12.8	10.05	3.40
latent variable space	10726	-12.293	10.011	5.104

Table 5: Closest Value to the Setting Value of Only BO Test 2 in the Latent Variable Space

	Gas valve (steps)	RF Power (dBm)	RF Frequency (GHz)	Intermediate electrode (kV)
Only BO Test 1	10900	-14.0	10.09	3.51
latent variable space	10874	-14.198	9.996	3.478

なお、潜在変数空間内をベイズ最適化によって最大 1000 回探索し、復元された各パラメーターを規格化した際の差分の二乗和が最も小さい値を最も近い値とした。これらの結果より、オートエンコーダを採用したことで意図せず離散的な最適化になっており、その影響でビーム輝度が最大化できなかつたと考えられる。その為、ECR イオン源のようにわずかなパラメーター変更でビームへ多大な影響を与える機器への応用には向かない可能性が高い。また、ECR イオン源への応用を考えると正確なシミュレーションができないため、学習データが不足することも懸念点である。そこで本手法はビーム輸送系などのシミュレーションが可能で且つ微小な誤差による影響を強く受けにくい領域で、最適解近傍を短時間で探索するに留め、その後別の手法で微調整するような使い方が望ましいと考えられる。

5. 結論

本研究では、オートエンコーダによる次元削減とベイズ最適化を用いた最適化手法を ECR イオン源へと応用した。その結果次元削減を用い手法に性能的に劣るということが分かった。これは ECR イオン源がパラメーター設定に非常に敏感であることが主な原因であると考えられる。その為、本手法は過敏なパラメーターのない領域で活用する必要があると考えられる。そこで今後は長いビーム輸送系に対してシミュレーションで検証を行い、活用可能性を探る。

参考文献

- [1] Eito Iwai *et al.*, Proceedings of PASJ in 2021. WEOB02.
- [2] Morita, Y. *et al.*, “Developments of control system for ion source using machine learning”, Journal of Physics: Conference Series. Vol. 2244, No. 1. IOP Publishing, 2022.
- [3] Močkus, Jonas, “On Bayesian methods for seeking the extremum”, Optimization Techniques IFIP Technical Conference: Novosibirsk, July 1–7, 1974. Springer Berlin Heidelberg, 1975.
- [4] Zhai, Junhai *et al.*, “Autoencoder and its various variants”, 2018 IEEE international conference on systems, man, and cybernetics (SMC). IEEE, 2018.
- [5] T Itahashi *et al.*, ‘Performance of the NANOGUN™ electron cyclotron resonance ion source applied for nuclear astrophysics’, Rev. Sci. Instrum. 71, 1075 (2000)
- [6] Martín Abadi *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.