

# 制約付き高次元パラメータにおけるベイズ最適化: 粉体計量への応用

宮川 翔貴<sup>1,a)</sup> 矢野 敦仁<sup>1,b)</sup> 澤田 奈生子<sup>1,c)</sup> 小川 勇<sup>1,d)</sup>

**概要:** ベイズ最適化はブラックボックスなモデルに対して少ない試行回数でパラメータを最適化できる手法だが、一般に高次元なパラメータでは機能しない。低次元空間への非線形な埋め込みにより効率的な探索を行う方法があるが、非線形変換により制約条件を陽に表現できない問題がある。本研究では、パラメータの分割および Disentangled Representation Learning を導入することで等式制約と不等式制約をそれぞれ考慮する方法を提案する。さらに、高次元かつ制約があるパラメータを最適化する問題の応用例として粉体計量タスクを導入し、提案手法の有効性を評価した。実験の結果、提案手法により制約条件を考慮することができ、手動でのパラメータ調整にくらべて試行回数を約 1/3 まで減らすことができた。

## High-Dimensional Bayesian Optimization with Constraints: Application to Powder Weighing

### 1. はじめに

ベイズ最適化はブラックボックスなモデルに対して少ない試行回数で効率的に適切なパラメータを推定することができるが、10次元以上の高次元なパラメータに対してうまく機能しないことが知られている。そのため、高次元なパラメータに対してよりベイズ最適化の効率性を向上するための手法が研究されている。とくに、非線形な次元圧縮により獲得した低次元な潜在空間を探索する方法 [1][2][3] がある。非線形な次元圧縮は、線形な次元圧縮を扱う方法 [4][5] や高次元空間の部分空間を扱う手法 [6] に比べて効率的な探索が可能であり、また離散値を連続値に変換することができる。しかし、非線形変換によりパラメータの制約条件を明示的に表現することができない問題がある。

そこで本研究では、非線形な次元圧縮を行っても制約条件、とくに不等式制約を考慮する方法を提案する。具体的には、Disentangled Representation Learning (DRL) とよばれる潜在変数のもつれを解く表現学習を導入することで不等式制約を暗黙的に考慮する。一般に DRL は人間が

制御しやすい生成モデルを学習するために用いられるが、我々はデータセットから不等式制約を学習する用途に応用する。この方法ではユーザが不等式制約を明示的に与える必要はなく、画像や自然言語のように制約条件を定式化しづらい非構造化データに対して適用できる。なお、等式制約についてはパラメータを可変パラメータと固定パラメータに分割することで明示的に考慮することができる。

提案手法は汎用的な設計であり、高次元かつ制約条件のあるパラメータを最適化する問題に対して広く適用できる。本研究では具体的なユースケースとして、粉体計量タスクとよばれる所定の分量の粉体を精度よく量り取る問題への適用を検討した。従来では人手で時間をかけて粉体計量装置に入力するパラメータをチューニングしており、本研究は初めて機械学習を適用した事例となる。

我々は粉体計量タスクに関して二種類の実験を行った。ひとつは、DRL によって不等式制約を考慮することができるかを評価する実験であり、我々はこの実験から潜在空間の次元数と DRL による正則化係数を適切にチューニングすることで不等式制約を考慮できることを明らかにした。もうひとつは実用的な試行回数で適切なパラメータを探索することができるかを評価する実験であり、我々は人間が設定する場合に比べて約 1/3 の試行回数で適切なパラメータを探索できることを示した。

<sup>1</sup> 三菱電機 情報技術総合研究所  
Kamakura, Kanagawa 247-8501, Japan  
a) Miyagawa.Shoki@ds.MitsubishiElectric.co.jp  
b) Yano.Atsumyoshi@dn.MitsubishiElectric.co.jp  
c) Sawada.Naoko@df.MitsubishiElectric.co.jp  
d) Ogawa.Isamu@ah.MitsubishiElectric.co.jp

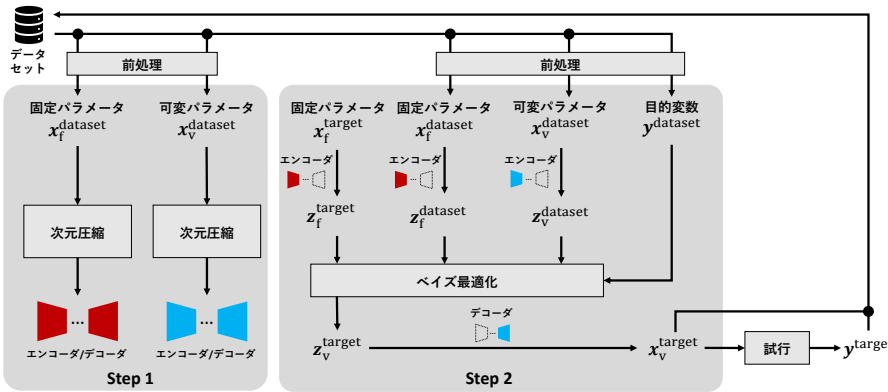


図 1: 提案手法の全体像.

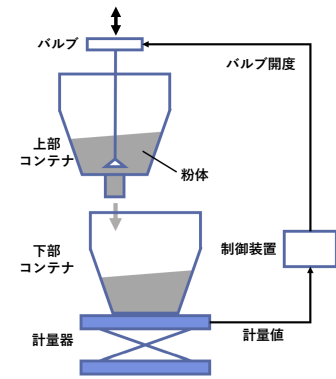


図 2: 粉体計量装置.

## 2. 提案手法

提案手法の全体像を図 1 に示す. はじめに, ベイズ最適化において等式制約を考慮するために, ユーザはすべてのパラメータを等式制約のない可変パラメータと等式制約のある固定パラメータと目的変数に振り分ける. 一般には, ユーザが最適化したいパラメータを可変パラメータとし, 生産対象や温度など外的要因に依存するパラメータを固定パラメータとする. つぎに, データセットから抽出した各パラメータに対して正規化や外れ値除去などの前処理を実施し, 後述の各ステップに入力する. 最後に, 可変パラメータ  $x_v^{\text{target}}$  を用いて試行し, 新たな評価値  $y^{\text{target}}$  を得る. この対応関係  $((x_v^{\text{target}}, x_f^{\text{target}}), y^{\text{target}})$  は新たにデータセットに追加され, 次の試行における可変パラメータの生成に利用される.

### 2.1 Step1: 次元圧縮

Step1 では, 可変パラメータと固定パラメータのそれぞれについて, 次元削減するエンコーダおよび次元復元するデコーダを学習する. ここで, 可変パラメータについては効率的な圧縮を行うために非線形な深層学習モデルである Variational Auto-Encoder (VAE) [7] を用いた. 固定パラメータは探索の対象ではないため, 線形なモデルである Principle Component Analysis (PCA) [8] を用いた.

VAE の学習に DRL を導入することで不等式制約を暗黙的に考慮することができる. DRL により低次元空間を人間が解釈可能な特徴の組み合わせで表現することができる. つまり, 低次元空間のあらゆる探索領域について制約条件を満たすかどうかを判定する代わりに, 各軸に対応する特徴が制約条件を満たすかどうかのみ判定すればよい. 特定の軸について制約条件をみたさない特徴を学習している場合は, ベイズ最適化時にそのような軸方向の探索を制限することで暗黙的に不等式制約を考慮することができる.

VAE に DRL を導入した  $\beta$ -VAE [9] の損失関数  $L$  を以下に示す. 第一項は再構成誤差を表し, 第二項は近似事後

分布と事前分布間のカルバック・ライブラー情報量を表す. 第二項の係数  $\beta$  を大きくするほど不等式制約を考慮した表現学習となるが, 一方で再構成誤差が大きくなってしまい, 試行回数を増やしても適切なパラメータを生成できない可能性がある点に注意する.

$$L(x_v, z_v) = \mathbb{E}_{q(z_v|x_v)}[\log p(x_v|z_v)] - \beta D_{\text{KL}}[q(z_v|x_v)||p(z_v)]$$

### 2.2 Step2: ベイズ最適化

Step2 ではユーザが事前に決定した固定パラメータ  $x_f^{\text{target}}$  を入力として次の試行に用いる可変パラメータ  $x_v^{\text{target}}$  を生成する. はじめに, Step1 で学習したエンコーダを用いてデータセットの可変パラメータと固定パラメータの潜在変数  $z_v^{\text{dataset}}, z_f^{\text{dataset}}$  を得る. つぎに, これらの潜在変数に対応する評価値  $y^{\text{dataset}}$  を用いて, ガウス過程回帰モデルなどの期待値  $\mu(z_v, z_f)$  および分散  $\sigma(z_v, z_f)$  を推定するベイズモデルを学習する. 最後に, 等式制約  $z_f = z_f^{\text{target}}$  のもとで獲得関数  $a(z_v, z_f)$  を最大化する可変パラメータの潜在変数  $z_v^{\text{target}}$  を生成する. 本研究では獲得関数は Upper Confidence Bound [10] として活用重視, 探索重視, あるいはその中間に対応する可変パラメータを生成し, ユーザはいずれか一つを選べるようにした.

$$z_v^{\text{target}} = \arg \max a(\mu(z_v, z_f = z_f^{\text{target}}), \sigma(z_v, z_f = z_f^{\text{target}}))$$

## 3. 粉体計量

図 2 の粉体計量装置は, 粉体を収めた上部のコンテナからバルブを介して下部のコンテナに所定の分量を量り取る装置である. 制御装置はバルブの開度を段階的に 9 回変更することで粉体の投入量を制御できる. 具体的には下部コンテナの底にある計量器の値を読み取り, 一定の値に達したらバルブ開度を変更する. このときの計量値 (切替重量とよぶ) およびバルブ開度が可変パラメータに該当し, バルブ開度の初期値を含めて可変パラメータは 19 次元であ

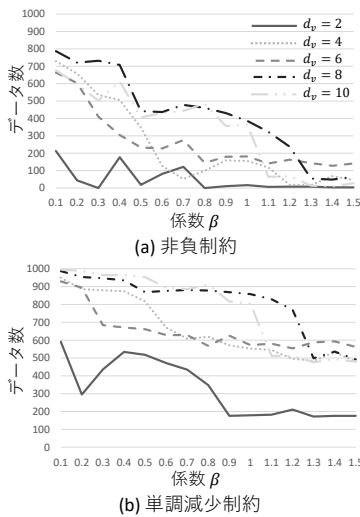


図 3: 各不等式制約を満たさないデータ数。

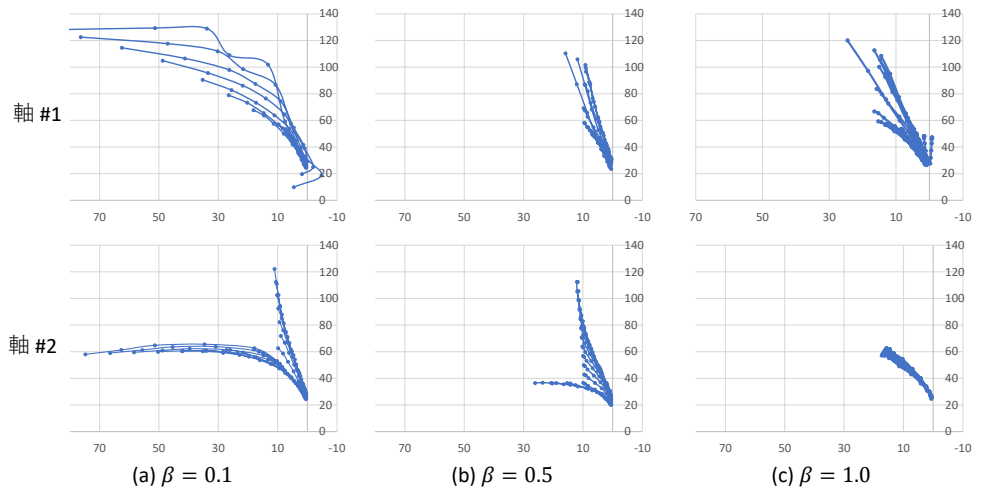


図 4: 生成された可変パラメータの可視化結果。横軸と縦軸はそれぞれ切替重量とバルブ開度を表す。

る。可変パラメータの適切な値は粉体の種類や計量したい分量によって異なり、従来では 20 回ほどの試行回数を重ねて人手で決定している。

バルブ開度と切替重量の両方について、非負制約と単調減少制約の 2 種類の不等式制約がある。一般に計量値の小さい初期の段階ではバルブの開度を大きくし、計量が進むにつれてバルブの開度を徐々に小さくするように設定パラメータを入力する必要がある。なぜなら、計量初期の段階でバルブ開度が小さいと計量に時間がかかり、また計量末期の段階でバルブ開度が大きいと既定の分量よりも多く粉体を計量してしまうためである。

さらに、固定パラメータと目的変数を定義する。固定パラメータには、粉体の物理的な特性値\*1やユーザが事前に決めた条件\*2が含まれ、合わせて 17 次元となる。また、装置が実際に計量した値と必要分量の誤差を計量誤差と定義し、計量誤差を二乗した値を目的変数とした。

データセットには 60 種類の粉体含まれており、総計 1792 回分のデータが存在する。各粉体の試行回数の平均および標準偏差は  $31.33 \pm 19.48$  回である。データセットには各粉体ごとに最終的に実務で採用された試行が含まれており、このときの採用基準は計量誤差が必要分量に対して 1%以下となることである。

## 4. 評価

### 4.1 実験 1: 次元圧縮の評価

#### 4.1.1 実験内容

実験 1 では DRL の導入により不等式制約を考慮できて

\*1 平均粒径 [μm], 見かけ比重 (ゆるみ)[-], 見かけ比重 (固め)[-], 圧縮度 [%], 安息角 [°], スパチュラ角 (平均)[°], 流動性指数 [-], 崩壊角 [°], 差角 [°], 分散度 [%], 噴流性指数 [-] の 11 種類

\*2 必要分量, バルブ径, 入力分量, バルブの揺動の有無, バイブレーションの有無, 事前バイブレーションの有無の 6 種類

いるかを評価する。具体的には、潜在空間の次元数  $d_v$  および係数  $\beta$  をハイパーパラメータとして、潜在空間において原点を中心とする半径 2 のバウンディングボックスから 1000 個のデータを一様にランダムサンプリングしたときの不等式制約を満たさないデータ数を定量評価した。また、定性的な評価として、潜在空間の各軸上を等間隔にサンプリングしたデータをデコードし、DRL によって学習した特徴を可視化する。これらの結果を通して潜在空間の適切な次元数  $d_v$  と係数  $\beta$  の値を決定する。

#### 4.1.2 結果

図 3 に次元数  $d_v$  および係数  $\beta$  を変化させたときの不等式制約を満たさないデータ数を評価した結果を示す。この結果から以下のことがわかる。

- 係数  $\beta$  が大きくなると不等式制約を満たさないデータの割合が減る。これは DRL によって不等式制約を考慮できたことを意味している。
- 潜在空間の次元数を大きくすると不等式制約を満たさないデータの割合が増える。これは、次元数が大きいとより細かい特徴の学習が可能であり、原点から離れた探索領域においてはその細かい特徴が強調されて不等式制約を満たさなくなるためである。

上記の結果より、不等式制約を満たしやすいという理由から実験 2 で用いる VAE の次元数を  $d_v = 2$  とした。

### 4.2 実験 2: ベイズ最適化の評価

#### 4.2.1 実験内容

実験 2 では、提案手法を用いて実用的な試行回数で適切なパラメータを探索できるかどうかを評価する。ここで、実用的な試行回数とは計量誤差が必要重量に対して 1%以下になるまでの試行回数と定義した。本実験ではデータセットに含まれない 3 種類の未知の粉体を用いて評価を

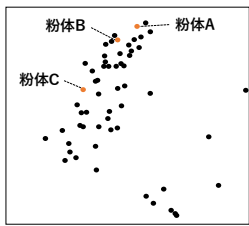


図 5: 固定パラメータの分布を可視化した結果。

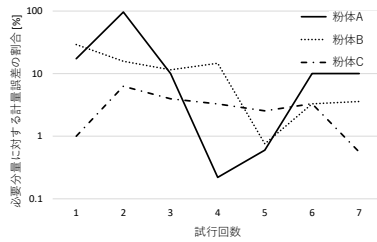


図 6: 各試行における計量誤差。

行った。以降ではこれらの粉体をそれぞれ粉体 A, B, C とよぶ。固定パラメータの分布を PCA により 2 次元へ可視化した結果を図 5 に示す。いずれの粉体も外れ値でないため、提案手法の汎用性を評価するうえで適切な粉体であることがわかる。

不等式制約を満たさないパラメータが生成された場合には、不等式制約を満たすような最小の変形を与えて装置に入力した。ただし、変形が大きい場合や明らかに装置が正しく機能しないとユーザが判断した場合には、装置に入力せず計量誤差を 10% として可変パラメータを再探索した。

#### 4.2.2 結果

3 種類の粉体について 7 回試行を続けたときの必要分量に対する計量誤差の割合を示した結果を図 6 に示す。いずれの粉体についても 7 回以内に誤差 1% を下回っており、人手でのチューニング回数 (20 回以上) に対して 1/3 程度の試行回数で目標性能を達成することができた。これは我々のフレームワークが有効に機能したことに加え、今回学習した VAE およびガウス過程モデルの両方について未知の粉体に対する汎化性能が高いことを意味している。

さらに、粉体 B, C についてはすべての試行において不等式制約を満たす可変パラメータの生成に成功した。ただし、粉体 A については 7 回目の探索において非負制約を満たさない可変パラメータが生成された。これは、図 5 に示すように粉体 A は粉体 B, C に比べて性質が類似した粉体が少なく、データの少ない領域を探索したことによる影響だと考えられる。

## 5. 議論

実験 1 の結果から、 $\beta$ -VAE を用いる場合は潜在空間の次元数  $d_v$  と係数  $\beta$  の 2 種類のハイパーパラメータの調整が重要であることがわかる。次元数  $d_v$  が大きいと再構成誤差が小さくなるが、ベイズ最適化によって生成された可変パラメータは不等式制約を満たさない可能性が高くなる。係数  $\beta$  が大きいと事後分布崩壊が起きて再構成誤差が大きくなり、一方で可変パラメータは不等式制約を満たしやすくなる。つまり、次元数  $d_v$  と係数  $\beta$  の影響は対照的であり、また係数  $\beta$  の適切な値は次元数  $d_v$  に依存することを意味している。そのため、次元数  $d_v$  と係数  $\beta$  の適切な値

を決定するうえで、図 3 や図 4 のような可視化は有用である。実際に、このような可視化に基づいてハイパーパラメータを決定したことで実験 2 では実用的な試行回数でパラメータを最適化することができた。

ただし、実験 2 では次元数  $d_v$  と係数  $\beta$  の特定の組み合わせについてのみ検証しており、提案手法が他の組み合わせに対してどのくらい頑健に機能するかは明らかでない。今後は、他のハイパーパラメータの組み合わせについても実験し、提案手法の頑健性を向上する方法を検討する。

## 6. おわりに

本研究では高次元かつ制約条件のあるパラメータに対し次元圧縮とベイズ最適化を組み合わせた汎用的なフレームワークを設計した。とくに、次元圧縮において DRL を導入することでユーザが不等式制約を明示的に与えなくても考慮できることを明らかにした。応用例として粉体計量を考え、提案手法によって非負制約と単調減少制約の両方を暗黙的に考慮でき、かつ実用的な試行回数でパラメータを最適化できたことを示した。ただし、提案手法では  $\beta$ -VAE における潜在空間の次元数および係数  $\beta$  のチューニングが重要である。今後はこれらハイパーパラメータに対する提案手法の頑健性を検討していく。

**謝辞** 本研究では月島マシンセールス株式会社の新田博文様、中込真人様、瀧澤憲治様および月島機械株式会社の吉村知規様にご提供いただき、さらに実験にご協力いただきました。この場を借りて深く御礼申し上げます。

## 参考文献

- [1] Rafael Gómez-Bombarelli, et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, Vol. 4, No. 2, pp. 268–276, 2018.
- [2] Pascal Notin, et al. Improving black-box optimization in vae latent space using decoder uncertainty. Vol. 34 of *NeurIPS '21*, pp. 802–814.
- [3] Chaochao Yan, et al. Re-balancing variational autoencoder loss for molecule sequence generation. *BCB '20*.
- [4] Benjamin Letham, et al. Re-examining linear embeddings for high-dimensional Bayesian optimization. *NeurIPS '20*.
- [5] Ziyu Wang, et al. Bayesian optimization in high dimensions via random embeddings. *IJCAI '13*, p. 1778–1784.
- [6] Johannes Kirschner, et al. Adaptive and safe Bayesian optimization in high dimensions via one-dimensional subspaces. Vol. 97 of *ICML '19*, pp. 3429–3438.
- [7] Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. *ICLR '14*.
- [8] I.T. Jolliffe. *Principal Component Analysis*. Springer Verlag, 1986.
- [9] Irina Higgins, et al. beta-vae: Learning basic visual concepts with a constrained variational framework. *ICLR '17*.
- [10] Peter Auer. Using confidence bounds for exploitation-exploration trade-offs. Vol. 3, p. 397–422, 2003.