

新規医療知識発見を目的とした量子因果探索アルゴリズムの開発と実装 —未知の因果を検知する、医療課題への処方箋—

1. 背景

近年我が国の高齢化が急速に進行しており、解決すべき医療課題もますます増加していくと想定される。こうした背景の中、医療の情報化・電子化に伴い増加する医療データの活用が注目を集めている。医療分野でのエビデンス構築としては臨床試験が中心だが、費用・時間的制約などの課題が知られている。そこで、昨今注目されているのが、実現場から収集される 2 次データである医療リアルワールドデータである。例えば、疾患レジストリ、電子カルテデータ、レセプトデータが挙げられる。このように、医療分野においても、ビッグデータを利活用することで新規の知見を蓄積する環境が整備されつつある。

現状、医療リアルワールドデータ分析手法としては、ある因果関係の方向性を既知とした上でその因果効果を推定する、「因果推論」が中心である。しかしながら、因果関係の方向をあらかじめ決定するのが難しい場面が、医療という不確実性の多い領域では少なくない。因果関係の事前決定が難しいケースでは、データのみから因果関係を検知し、新規知識の発見につなげる「因果探索」が重要である。

因果探索アルゴリズムは、あるデータが得られた場合、そのデータに数理的仮定を置くことで、事前知識なくデータのみから変数間の因果関係の方向性を決定し、変数間の因果関係を表現する因果グラフを同定する手法である。

昨今注目されている因果探索手法の一つに、linear non Gaussian acyclic model (LiNGAM) が挙げられる。LiNGAM は、データに非正規性の仮定を置き、誤差変数に仮定した独立性を利用して因果グラフの構造を一意的に特定する手法である。LiNGAM は選択する独立性指標によっては適切に因果構造を推定できないことが指摘されているが、カーネル法を利用した独立性指標を用いることで、より精度よく因果構造を推定できることが知られている。

本プロジェクトでは、量子コンピューティングとカーネル法の、ヒルベルト空間で効率的に計算を実行するという共通点に着目し、カーネル法を用いた LiNGAM の独立性指標に量子カーネルを用いれば、因果探索アルゴリズムを改良できるのではないかと、という着想に至った。量子カーネルを用いた因果探索アルゴリズムは、私の知る限り、本プロジェクトが初の試みである。このような新規性を有する一方で、既に有効性が示されているアルゴリズムの一部に限定して量子コンピューティングを応用する点で、実現可能性が高いと考えた。

2. 目的

量子カーネルを用いた因果探索アルゴリズムを開発し、実際に医療データに適用して精度を検証することで、新規医療知識発見に資するか否かの検討を行う。

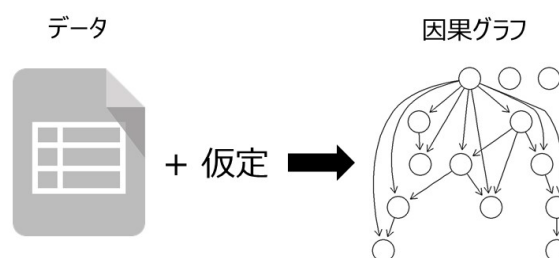


図 1 因果探索アルゴリズムのイメージ

3. ソフトウェア開発内容

量子因果探索アルゴリズムを開発するにあたり、まず、骨格となる LiNGAM アルゴリズムを決定した。LiNGAM には主に、独立成分分析を用いる ICA-LiNGAM と、回帰分析をベースとした Direct LiNGAM があるが、本プロジェクトでは Direct LiNGAM を選択した。Direct LiNGAM では、単回帰分析をベースとして、その説明変数と残差に関する独立性指標が最も低い外生変数を順次求める操作を行い、因果構造を推定していく (図 2)。

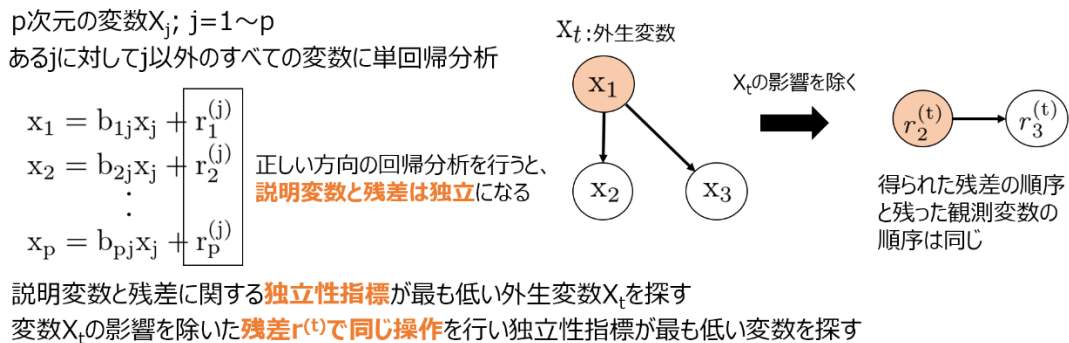


図 2 Direct LiNGAM の大まかな流れ

Direct LiNGAM の精度の向上には、独立性指標の選択が重要である。本プロジェクトでは、カーネルの形状に依存しない独立性指標である、INOCCO が量子カーネルには、より適切であると考え、INOCCO を選択した。INOCCO の計算において、グラム行列を計算する箇所に量子回路を用いた。

量子カーネルの計算のため、回路を前半/後半に分け、後半に前半のダガーを取って回路とする inversion test を採用した。量子回路を設計するにあたり、古典コンピュータでは再現困難な量子カーネルの模索というモチベーションの達成のため、出力される確率分布を古典コンピュータで精度よく推定することは困難であることが知られている Instantaneous Quantum Polynomial time circuit (IQP 回路) をベースにした量子回路を設計した (詳細は図 3)。

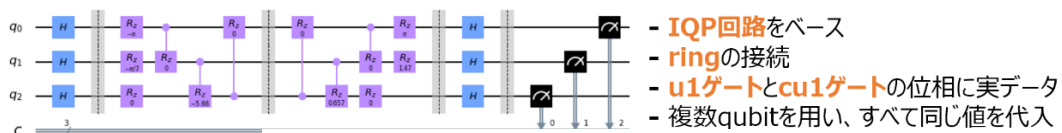


図 3 量子カーネルの回路構成のまとめ

実装には、python ライブラリである cirq を用いて回路実装を行った。より具体的には、cirq で量子カーネルを計算する量子回路の前半を構築し、cirq-qulacs でその回路の final state vector を算出した。その後、量子回路の後半については、final

state vector のダガーを取り、両者を内積計算することで、量子カーネルを形成するグラム行列の計算を行った。

4. 新規性・優位性

まずは疑似データを用いたシミュレーションを行った。サンプル数 100、変数 3 で、ラプラス分布から誤差項を発生させたデータセットを合計 100 セット用意した。変数間の関係性については、図 4 の関係性を設定した。量子カーネルを用いた因果探索アルゴリズムと、ガウシアンカーネルを用いた既存の因果探索アルゴリズムをそれぞれ適用し、図 4 の正しい因果構造を推定できたデータセット数をカウントした。

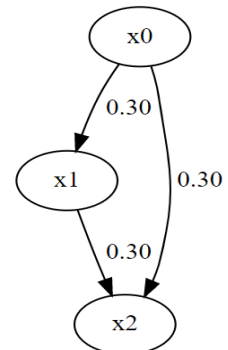


図 4 変数間関係

表 1 が実験の結果である。ガウシアンカーネルを用いた既存の因果探索アルゴリズムでは特定できなかった因果構造を、量子カーネルを用いた量子因果探索アルゴリズムでは、14 データセット特定することができた。

表 1 疑似データを用いたシミュレーションの実験結果

両方正解	両方不正解	量子カーネルのみ正解	ガウシアンのみ正解
39	38	14	9

続いて、医療データに量子因果探索アルゴリズムを適用した。用いたデータは、UCI Heart Disease Data Set という、循環器に関するオープンデータである。欠損値を削除した 297 例の全例データセットと、そこからランダムに 100 例を選んだショートバージョンのデータセットを準備し、量子因果探索アルゴリズムと、ガウシアンカーネルを用いた既存の因果探索アルゴリズムをそれぞれ実装し、臨床的に妥当な因果構造が特定できるかを検討した。UCI Heart Disease Data Set の中で使用した変数は、age という年齢の変数と、cp という 4 種類の胸痛タイプを表す変数と、exang という運動が誘発する狭心症の有無を表す変数である。exang は cp の分類に寄与するため、exang から cp に矢印が引かれるのが、臨床的に妥当な関係性であると想定される。

実験の結果、量子カーネルを用いた量子因果探索アルゴリズムを用いた場合、297 例の全例を用いたデータセットでも、100 例しかないショートバージョンのデータセットでも、図 5 で描かれている、臨床的に妥当な因果関係を特定した。一方で、ガウシアンカーネルを用いた既存の因果探索アルゴリズムでは、全例データセットを用いれば図 5 の臨床的に妥当な因果関係を特定できたが、100 例のショートバージョンのデータセットでは、全ての変数が独立という、臨床的に妥当ではない結果を提示した (図 6)。

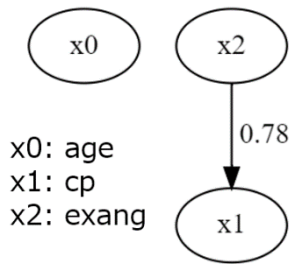


図 5 臨床的に妥当な因果関係

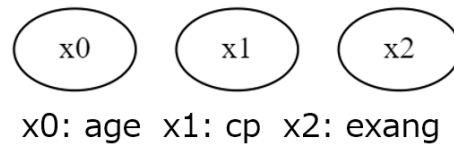


図 6 臨床的に妥当でない因果関係

5. 期待されるユーザー価値と社会へのインパクト

ユーザー価値として、本プロジェクトで開発したアルゴリズムを用い、大規模医療データから新規知識を検知し、医療者が発見しきれない因果関係を抽出することで、現場の医療者をサポートすることを想定している。

具体的な医療分野への応用先としては、精神科領域では、統合失調症患者に関する新規薬剤反応の探索に応用できると期待される。薬効判断が難しく、経験則からは予測できない副反応が検出できれば、新たな創薬のヒントになり得る。将来的に実現し得るより一般的な課題解決として、ゲノムデータと臨床データの統合データベースを用いた因果探索が挙げられる。例えば、新たな腫瘍マーカーの発見や、遺伝要因と生活因子との関連性についての新規知識の発見が、未病予防につながる可能性も想定される。

また、医療・介護・福祉領域という、分野横断的な課題解決も期待される。医療分野は専門が細分化されているため、経験則だけでは分野横断的知識を体系化するのが困難なこともあり、データのみから新規知識を検知するアルゴリズムを用いれば、既存にはない新しい医療・介護連携体制の発見につながる可能性も考えられる。

さらに、産業領域に関しては、ヘルステック関連の課題解決に寄与できると期待される。例えば、高血圧・糖尿病などの生活習慣病患者からデバイスを通じて生活関連データを取得し、採血データと紐づけた上でアルゴリズムを適用すれば、日常生活と臨床症状の間の新たな因果関係を検知でき、既存にはない生活指導の開発に寄与することが期待される。

このように、因果探索アルゴリズムは医療に関する様々な分野の課題を解決し得ると期待され、データの蓄積が進むにつれて、ユーザーの使用範囲も拡大し、より一層本プロジェクトの成果の価値が高まると期待される。

6. 氏名（所属）

川口 英明（東京大学医学部附属病院 企画情報運営部）