

審査の結果の要旨

氏 名 伊藤 陽一

本研究は、測定誤差を伴うマイクロアレイデータにおける **Feature selection** を、誤分類確率を考慮しつつデータをカテゴリー化した上で行う方法論を提案したものであり、その方法に関して以下の結果を得ている。

1. 誤分類確率をある値以下に抑えてカテゴリー化した際に、単位変化量あたり何カテゴリー設定できるかという **Resolution** という指標を定義した。この指標により、誤分類確率を明示的に制御したカテゴリー化が可能となった。
2. データをカテゴリー化したことにより、全対象者の遺伝子発現量変化を変化パターンとして捉えることができた。このことにより、変化パターンが同一となる遺伝子をグループ化することができ、検討すべき変数の数を減らすことが可能となった。減少の程度に関して、異なる変化パターンの数を総遺伝子数で割った **Dimension Reduction Proportion** という指標を定義した。この指標により、カテゴリー化による変数減少の効果を表わすことができた。
3. 小さな標準誤差が観測された遺伝子は偶然に選択されてしまう傾向があるが、カテゴリー化によってそのような遺伝子を除外することができた。**SAM** では、**fudge factor** を導入することによって、そのような遺伝子が選択されにくくしているが、本研究によるアプローチは、除外される基準をより明示的に指定できるという利点がある。
4. 一般にカテゴリー化に伴い検出力が減少することが知られているが、平均遺伝子発現量変化 0.3、標準偏差 0.3 の正規乱数を発生させてシミュレーションを行った結果、カテゴリー数 3 の場合、検討すべき遺伝子数を劇的に減らすことができるものの、検出力がほとんどなくなってしまうことが明らかとなった。また、平均遺伝子発現量変化が 0.3 以上ある遺伝子であれば、サンプル数 50 のカテゴリー数 5 で十分に検出可能であることが明らかとなった。

以上、本論文はマイクロアレイの測定誤差に着目し、誤分類確率を制御する方法論を提案した。本研究は、マイクロアレイの測定誤差に着目した点において独自性があり、マイクロアレイデータ解析の発展に重要な貢献をなすと考えられ、学位の授与に値するものと考えられる。