

Rによる並列計算とスーパーコンピュータ

中野純司（統計数理研究所），中間栄治（株式会社 COM-ONE）

現在のノートパソコンの性能は十数年前のスーパーコンピュータの性能に匹敵する。またその上でRをはじめとする多くの優秀な統計解析ソフトウェアが利用できる。そのため、これまでの標準的な統計計算に関してはたいていの場合、それらで十分である。

ただ、計算機、ネットワーク、センサー技術などの急速な発達に伴って、得られるデータはますます多量になり（ビッグデータ）、また、複雑な関係を持つようになっている。さらに統計手法としても計算機集約型のものが新しく開発されている。そのようなデータや手法を処理するためにはより高速に多量の統計計算を行う必要がある。そのためには（スーパーコンピュータ上の）並列計算を利用する必要がある。

本発表では、まず、スーパーコンピュータと並列計算の基礎を述べる。特に統計数理研究所のスーパーコンピュータと汎用の並列計算手法 MPI (Message passing interface) について外観する。そしてもっとも利用しやすい統計並列計算技術として R による並列計算の方法を示す。特に MPI を効率的に利用する並列計算パッケージ Rhpc について詳しく述べる。