RESEARCH PAPER | Computer Science | Volume : 5 | Issue : 1 | Jan 2015 | ISSN - 2249-555X

Role and applications of Scientific Computing and Applied Statistics in Life Sciences

KEYWORDS: Scientific Computing, Computational Life Sciences, Applied Statistics, Bioinformatics

Bhagavati Parekh

BCA & PGDCA College,Junagadh

ABSTRACT *The trio - Scientific computing, Life Sciences and Applied Statistics altogether help provide a fundamental understanding of complex biological systems and offers the potential to significantly impact a wide variety of technologies. Scientific computing is the cross-disciplinary field at the intersection of modeling scientific processes, and the use of computers to produce quantitative results from these models. It is what takes a domain science and turns it into a computational activity. Computational statistics is somewhat recent discipline which has emerged out of the increasing applications of computers to sort and simulate the statistical problems. Statistical computing includes numeric analysis, database methodologies, computer graphics, software engineering and the computer-human interface. This paper elaborates the role of the Scientifc computing, Applied statistics, their scope in the field ; the methods of implementation and various applications.*

Introduction

Scientific Computing and Life Sciences have gained the most important place in the field of science and technology in the last decade. They are positioned at the intersection of the high performance computing, quantitative modeling and modern biology. The trio - Scientific computing, Life Sciences and Applied Statistics altogether help provide a fundamental understanding of complex biological systems and offers the potential to significantly impact a wide variety of technologies, including drug discovery, novel therapies for human, animal and plant diseases, metabolic engineering and efficient production of traditional and high-value foodstuffs.

Scientific Computing:

Computational science (also **scientific computing** or **scientific computation)** is concerned with constructing mathematical models and quantitative analysis techniques and using computers to analyze and solve scientific problems. In practical use, it is typically the application of computer simulation and other forms of computation from numerical analysis and theoretical computer science to problems in various scientific disciplines.[1]

The essence of computational science is numerical algorithm[2] and/or computational mathematics. In fact, substantial effort in computational sciences has been devoted to the development of algorithms, the efficient implementation in programming languages, and validation of computational results.

Scientific computing is the cross-disciplinary field at the intersection of modeling scientific processes, and the use of computers to produce quantitative results from these models. It is what takes a domain science and turns it into a computational activity. As a definition, we may describe it: "The efficient computation of constructive methods in applied mathematics."[3].

2.1 Methods and Algorithms:

There are numerous methods which prove vital to the application of Computational Science in various fields. Such applications include pattern recognition, machine learning, multistage modeling of biological data, metrology, geometric modeling, tomography, signal and image processing etc.

Various Methods/Mathematical Models used in Scientific Computing:

The methods used in various applications of Scientific Computing are varied. The methods which are commonly applied include[4]:

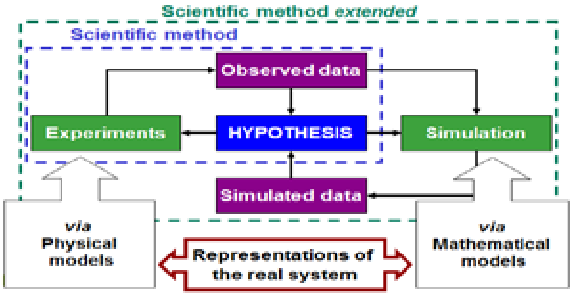
- Numerical analysis
- Application of Taylor series as convergent and asymptotic series
- Computing derivatives by Automatic differentiation (AD)
- Computing derivatives by finite differences
- Graph theoretic suites High order difference approximations via Taylor series and Richardson extrapolation
- Methods of integration on a uniform mesh: rectangle rule (also called midpoint rule), trapezoid rule, Simpson's rule
- Runge Kutta method for solving ordinary differential equations

Use of Various Algorithms in Scientific Computing:

Real-world changing conditions are modeled and implemented by the Scientific Computing application programs. Such application include weather forecasting, flow and pattern of air around a plane, automobile body distortions in a crash, the motion of stars in a galaxy, an explosive device, etc. Each specific condition/problem contains a set of vague information, which are identified, refined, defined and recorded in the form of various absolute as well as discrete data sets. Each data set corresponds to an area in space and contains information about that space relevant to the model.

2.2 Computer Modeling and Simulation:

A model is a physical, mathematical, or logical representation of a system entity, phenomenon, or process. A simulation is the implementation of a model over time. A simulation brings a model to life and shows how a particular object or phenomenon will behave. It is useful for testing, analysis or training where real-world systems or concepts can be represented by a model [5]



Various classes of computer based modeling and simulation:
The three classes of models and simulations are: virtual, constructive, and live:

Virtual simulations represent systems both physically and electronically. Examples are, the, Close Combat Tactical Trainer, certain modules for the delivery of built-in training, various scientific models which function on computational powers and appear physically for the further gesture.

Constructive simulations represent a system and its employment. They include computer models, analytic tools, mockups, IDEF, Flow Diagrams, and some engineering software tools like Computer-Aided Design/ Manufacturing (CAD/CAM).

Live simulations are simulated operations with real operators and real equipment. Examples are fire drills, operational tests, and the systems that work at laboratories to aid to the medical functionalities.

3. Computational Life Sciences:
Rapid advances in computer science, biology, chemistry, and other disciplines are enabling powerful new computational tools and models for simulating the applications of various faculties of applied sciences into the field. These computational tools hold tremendous promise for advancing applied and basic science. They provide the environment for simplifying the process from streamlining drug efficacy and safety testing to increasing the efficiency and effectiveness of risk assessment for environmental chemicals.

3.1 Computational life sciences task categories:
In silico drug discovery:
In silico drug discovery involves identifying potentially biologically active substances by modeling large numbers of complex virtual molecular interactions; it is generally considered to be an extremely computationally intensive task.

Clinical Research Informatics:
It will fundamentally change the process of moving beyond relatively small numbers of parameters and predictors, towards large scale genomic, metabolomic and proteomic datasets which will enable far more detailed associations between patient and disease progression, for improved outcomes.

Biomedical Image Analysis:
Rapid increases in the resolution and usage patterns of, and data generated by, modern imaging technologies like Functional MRI, PET, multislice CT), both clinical and research, necessitate a corresponding increase in computational resources to process and analyse the image data.

Biomolecular Structure Prediction:
Structure prediction from physical principles is very computationally demanding. More computing power will immediately allow the simulation of more complex biomolecules.

Computational biomodelling envelops a large amount of simulation technologies spanning cell, tissue, organ and organ system modelling, physiological state modelling, through to social interactions and population dynamics.

Bioinformatics (or Computational 'omics'): Bioinformatics referes to the computational analysis of the information encoded in the biological data and their expression patterns.

3.2 Major computational methods for life sciences computation:

There are a number of computational methods which support the life sciences computations. Depending on the computational requirements, single or the combination of methods can be put to use for varied tasks. They are:

Data Mining:
It is an interdisciplinary subfield of computer science[6]. It can be regarded as the computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems.

Mathematical Modelling:
It involves describing the system using mathematical concepts and language. The mathematical modeling in life sciences generally includes the problems that thrive upon the mathematical solutions for the given problems. Eg. Differential equations, fitting the curves to lines etc

Image analysis and visualization:
Image analysis generally deals with the extraction of important information from the images. Digital image processing techniques can be used for the purpose of segregating vital information from mainly digital images. These techniques are inspired by human visualization perception models.

Sequence Analysis:
Various Bioinformatics methodologies including sequence alignment, searching biological databases and other such techniques are used to understand the features, functions, structures or evolution of DNA, RNA or such other peptide sequences[7]. They are used for the pattern matching, gene identification application etc.

Data Indexing:
It can be defined as quite large datasets which are likely to be commonplace soon because of the next-gene sequencing and image technologies. Data indexing would include scalable techniques and related indexing technologies for these datasets.

4. Scientific computing and applied statistics:
Statistics is fundamentally concerned with the understanding of structure in data. One of the effects of the information-technology era has been to make it much easier to collect extensive datasets with minimal human intervention. Fortunately, the same technological advances allow the users of statistics access to much more powerful 'calculators' to manipulate and display data.

Volume : 5 | Issue : 1 | Jan 2015 | ISSN - 2249-555X

Computational statistics is somewhat recent discipline which has emerged out of the increasing applications of computers to sort and simulate the statistical problems. It basically aims at the design of algorithm for implementing statistical methods on computers, especially those, which otherwise could not have been implemented without computers, per say bootstrap, simulations as well as some other analytically intractable problems[8]

4.1 Various application areas of statistics:
Statistical techniques are used in a wide range various scientific and social researches, including:

- biostatistics
- computational biology
- computational sociology
- network biology
- Social science, sociology and social research.

Some fields of inquiry use applied statistics so extensively that they have specialized terminology.

4.2 Computational statistics:
Computational statistics is somewhat recent discipline which has emerged out of the increasing applications of computers to sort and simulate the statistical problems. It basically aims at the design of algorithm for implementing statistical methods on computers, especially those, which otherwise could not have been implemented without computers, per say bootstrap, simulations as well as some other analytically intractable problems[8].

The cross currents of the Computational Statistics:
The Computational Statistics is regarded now as a separate discipline in itself. In fact, it is a crux of various fundamental disciplines, and in a true sense actually, it is a multidisciplinary paradigm. Among any other inclusions, it is more closely related to statistics, though. It directly relates itself to the applied statistics, where computationally intensive methods are becoming commonly used in various application areas. Developments in other areas like numeric science and statistical analysis are also related to this area in a relevant manner.

Statistical Bioinformatics:
Technological advances in all the fields including applied statistics as well as Bioinformatics are providing levels and magnitudes of life science related data that were unimaginable even five years ago. This leads every component of what scientists do to be put forward by stretching, changing, and projecting forward in anticipation of further more development, both in the research area as well as in the practical implementation. The largest shift has been in the way we do science, it is no longer single laboratory science, it is now multidisciplinary efforts that bridge many disciplines and many species.

Conclusion:
The life sciences comprise the fields of science that involve the scientific study of living organisms – such as microorganisms, plants, animals, and human beings – as well as related considerations like bioethics. While biology remains the centerpiece of the life sciences, technological advances in molecular biology and biotechnology have led to a burgeoning of specializations and interdisciplinary fields. The largest shift has been in the way we do science, it is no longer single laboratory science, it is now multidisciplinary efforts that bridge many disciplines and many species; and computers have a major role to play in the field.

REFERENCE

[1] “Scientific computing definition by Babylon’s free dictionary.” [Online]. Available: http://dictionary.babylon.com/scientific computing/. [Accessed: 08-Jul-2014]. | [2] J. H. Rowland, “Computational Mathematics—An Introduction to Numerical Approximation (T. R. F. Nonweilerz),” SIAM Rev., vol. 28, no. 3, pp. 437–438, Sep. 1986. | [3] V. Eijkhout, Introduction to High Performance Scientific Computing. Lulu.com. | [4] T. Kauppinen and G. M. de Espindola, “Linked Open Science-Communicating, Sharing and Evaluating Data, Methods and Results for Executable Papers,” Procedia Comput. Sci., vol. 4, pp. 726–731, Jan. 2011. | [5] “Scientific modelling - Wikiquote.” [Online]. Available: http://en.wikiquote.org/wiki/Scientific_modelling. [Accessed: 08-Jul-2014]. | [6] “data mining (computer science) -- Encyclopedia Britannica.” [Online]. Available: http://www.britannica.com/EBchecked/topic/1056150/data-mining. [Accessed: 22-Jul-2014]. | [7] Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. Cambridge University Press, 1998, p. 356. | [8] C. Lauro, “Computational statistics or statistical computing, is that the question?,” Comput. Stat. Data Anal., vol. 23, no. 1, pp. 191–193, Nov. 1996. | |