

ANÁLISE NUMÉRICA DOS INDICADORES ESTATÍSTICOS DESCRITIVOS INTERVALARES

FURLAN, Vinícius Signori¹; GREHS, Érico Alves¹; FINGER, Alice Fonseca²; LORETO, Aline Brum³

¹Graduando em Ciência da Computação; ²Mestranda do PPGCC- CDTec/UFPEL; ³Profa. Adjunta CDTec/UFPEL, Pelotas-RS-Brasil. E-mail para correspondência: vsfurlan@inf.ufpel.edu.br

1 INTRODUÇÃO

Algoritmos convencionais, chamados de algoritmos pontuais, computam uma estimativa para uma resposta, e, talvez, um erro estimado. Quando se trabalha com números de ponto-flutuante o resultado é apenas uma aproximação de um valor real e erros gerados por arredondamentos ou por instabilidade dos algoritmos ou programas, podem levar a resultados incorretos.

Não se pode afirmar a exatidão da resposta estimada sem o auxílio de uma análise de erro, que é extensa, dispendiosa e nem sempre viável.

Técnicas intervalares, em contraste, computam um intervalo, com a garantia de que a resposta pertence a este intervalo. A matemática intervalar tem como objetivo proporcionar um controle de erros numéricos. Utilizando-se intervalos para representação dos números reais, é possível controlar a propagação de erros de arredondamento ou truncamento, entre outros, em procedimentos numéricos computacionais. Portanto, resultados intervalares carregam consigo a segurança de sua qualidade (MOORE, 1966).

O objetivo principal do presente trabalho é analisar a qualidade dos intervalos dos indicadores estatísticos intervalares Média, Variância, Desvio Padrão, Coeficiente de Variação, Covariância e Coeficiente de Correlação em três diferentes ambientes intervalares: XSC (*eXtended for Scientific Computation*), Maple Intervalar e IntLab, os quais possibilitam a programação utilizando operações definidas na matemática intervalar (MOORE, 1966). Através da aplicação dos cálculos dos erros Absolutos e Relativos nos intervalos solução, verifica-se a qualidade do intervalo e define-se qual dos ambientes retorna um melhor resultado.

2 METODOLOGIA (MATERIAL E MÉTODOS)

Considerando que o controle do erro numérico é feito através do uso de intervalos ao invés de números reais, Kulisch (2008) e Kulisch e Miranker (1981) propuseram que a implementação da aritmética intervalar seja realizada através da chamada aritmética de exatidão máxima, o que significa a busca para que resultados numéricos ou sejam um número de ponto flutuante ou estejam entre dois números de ponto flutuantes consecutivos. No caso de intervalos o suporte para a realização das operações aritméticas, com exatidão máxima, é usar os arredondamentos direcionados, ∇ e Δ , nos extremos dos intervalos: o limite inferior é arredondado para baixo (∇) e o superior para cima (Δ). A aritmética intervalar com os

arredondamentos direcionados fornece limites rigorosos para o contradomínio de operações e funções.

Cálculos numéricos em computadores devem ser realizados por meio das linguagens ou bibliotecas que tenham definidos o tipo intervalo e as operações sobre o tipo, usualmente denominadas de linguagens *XSC* (*eXtended for Scientific Computation*).

Existem diversas bibliotecas computacionais para Matemática Intervalar, dentre as quais três foram escolhidas para o presente trabalho. Dois ambientes que estão sendo utilizados nesse trabalho fazem o uso da metodologia de aritmética intervalar nos cálculos dos intervalos, é o caso dos ambientes *Maple Intervalar* e o *IntLab*.

O Maple Intervalar (MAPLE) é um ambiente interativo, com uma interface amigável que, para muitas finalidades, dispensa a programação. Bibliotecas Maple, uma vez carregadas, disponibilizam os comandos e operadores necessários para cálculos específicos, ou seja, intervalares. Possui uma linguagem de programação fundamentada no conceito de linguagem interpretada e um mecanismo de construção e distribuição de pacotes de programas e funções.

IntLab (INTLAB), é um pacote desenvolvido para a ferramenta *Matlab*[®] (MATLAB). Contém tipos de dados básicos e operadores para aritmética intervalar, bem como uma variedade de métodos numéricos usando intervalos.

O último ambiente considerado neste trabalho é o *C-XSC*. O ambiente de programação *C-XSC* (*C++ for eXtended Scientific Computing*) é uma ferramenta de programação poderosa e de fácil uso, especialmente para aplicações científicas e de engenharia. *C-XSC* é particularmente adequado para o desenvolvimento de algoritmos numéricos que proporcionam resultados altamente precisos e verificados automaticamente (KLATTE *et al*, 1993). Todas as operações básicas (intervalo) são de exatidão máxima (KRÄMER, 2011).

Para a validação das soluções intervalares dos indicadores estatísticos, realiza-se um comparativo entre as soluções dos cálculos reais dos indicadores, realizados no software NetBook (CAMPOS *et al.*, 2004), e as soluções dos cálculos intervalares dos mesmos indicadores estatísticos intervalares, realizados nos três ambientes intervalares supra citados.

O NetBook é uma ferramenta gratuita desenvolvida no CIn/UFPE, com o objetivo de suportar análise de desempenho de sistemas de comunicação, em particular, redes de computadores. É composto por quatro módulos: Estatística, Geração, Transformação e Gráfico. Exceto pelo módulo de geração de tráfego auto-similar, que foi implementado em C++, todos os demais módulos do NetBook foram implementados em Java.

Segundo Ratschek (1988), os computadores utilizam uma aritmética chamada aritmética de ponto flutuante. Nesta aritmética, números reais são aproximados por um subconjunto de números reais chamados representação numérica da máquina. Devido esta representação são gerados dois tipos de erros: quando uma entrada de valor real é aproximada por um número de máquina e quando o erro é causado por resultados intermediários aproximados pelos números de máquina. A aritmética intervalar fornece uma ferramenta para estimar e controlar esses erros *automaticamente*. No lugar de aproximar um valor real x por um número de máquina, o valor real x é aproximado por um intervalo x tendo número de máquina nos extremos inferior e superior. O intervalo x contém o valor real x . O comprimento (ou diâmetro) deste intervalo pode ser usado como medida para

qualidade da aproximação da resposta real exata. Os cálculos são executados usando intervalos ao invés de números reais e, conseqüentemente, a aritmética real é substituída pela aritmética intervalar.

Quando se realiza computação com número de máquina \tilde{x} não existe estimativa do erro $|\tilde{x} - x|$. A computação com utilização de intervalos fornece as seguintes estimativas para o erro:

$$\text{- Erro Absoluto: } |x - m(\mathbf{x})| < w(\mathbf{x}) / 2,$$

onde $m(\mathbf{x})$ é o ponto médio do intervalo \mathbf{x} e $w(\mathbf{x}) = \bar{x} - \underline{x}$ é o diâmetro do intervalo \mathbf{x} ;

$$\text{- Erro Relativo: } \left| \frac{x - m(\mathbf{x})}{x} \right| \leq \frac{w(\mathbf{x})}{2 \min|x|} \text{ se } 0 \notin \mathbf{x}, \text{ onde } |x| = \{x : x \in \mathbf{x}\}.$$

Aplicam-se estas medidas de erros com o objetivo de verificar a qualidade do intervalo solução em diferentes ambientes intervalares, verificando qual ambiente retorna o intervalo com melhor qualidade, obtido para os indicadores estatísticos intervalares conforme definidos em Loreto (LORETO, 2006).

3 RESULTADOS E DISCUSSÃO

A qualidade de um intervalo está relacionada com o diâmetro do mesmo (RATSCHEK; ROKNE, 1988). Quanto menor for o tamanho do diâmetro, melhor será a qualidade do intervalo.

Através das medidas de erros verifica-se que o método da extensão intervalar fornece como resposta intervalos que englobam a resposta real exata, e ainda, com qualidade na solução intervalar.

O presente trabalho encontra-se parcialmente executado. Até o momento, foram pesquisados e definidos os ambientes intervalares, bem como as medidas de erros para verificação da qualidade dos intervalos solução, obtidos para os indicadores estatísticos intervalares da média, variância, desvio padrão, coeficiente de variação, covariância e coeficiente de correlação.

Para a finalização, os indicadores estatísticos intervalares serão implementados nos ambientes C-XSC, Maple Intervalar e IntLab. Exemplos serão aplicados para calcular os valores reais, os intervalos solução e, em decorrência, os cálculos de erros para verificar qual ambiente retorna melhor solução intervalar com melhor qualidade.

4 CONCLUSÃO

As opções pelos ambientes intervalares C-XSC, Maple Intervalar e IntLab, para a realização deste trabalho, fundamenta-se na potencialidade de aplicações tanto na pesquisa quanto no ensino.

Com o sistema de ponto flutuante F(10, 14, -10, 10) (ou com quatorze casas decimais) espera-se verificar, através das medidas de erros (absoluto e relativo), que todos os valores reais dos indicadores estatísticos descritivos estejam contidos nos intervalos solução.

A verificação da qualidade de aproximação nos intervalos ocorre através de cálculos destas medidas de erros.

Garantindo a qualidade do intervalo solução e conhecendo o ambiente intervalar que retorna o intervalo com melhor qualidade, tem-se uma ferramenta intervalar confiável para ser utilizada em diversas aplicações que utilizem os indicadores estatísticos descritivos intervalares.

Destaca-se, como pontos positivos inerentes aos resultados obtidos, a possibilidade de aplicação da estatística descritiva intervalar em diversas áreas, como por exemplo, a da Bioinformática. Na Bioinformática existem muitos dados, a maioria com erros de medição, devido os processos serem imprecisos.

5 REFERÊNCIAS

CAMPOS, M. A. et al. Netbook: uma ferramenta para avaliação de desempenho em redes de comunicação. In: **SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES**, SBRC, 22., 2004, Gramado. **Anais**. . . Gramado: II/UFRGS, 2004. p. 967-974.

KLATTE, R, KULISCH, U., WIETHOFF, A., LAWO, C., RAUCH, M. **C-XSC - A C++ Class Library for Extended Scientific Computing**. Springer-Verlag, Heidelberg, 1993.

KRÄMER, Walter. **Multiple/arbitrary precision interval computations in C-XSC**. Springer-Verlag, 2011.

KULISCH, U. and MIRANKER, L. **Computer Arithmetic in Theory and Practice**, 1st ed., Academic Press, 1981.

KULISCH, U. W. (2008, apr) Complete interval arithmetic and its implementation on the computer. [Online]. Disponível em: <http://www.math.kit.edu/iwrm/seite/preprints/media/preprintn%20nr>

LORETO, Aline B. **Análise da Complexidade Computacional de Problemas de Estatística Descritiva com Entradas Intervalares**. Tese de Doutorado em PPGC, Instituto de Informática/UFRGS, Porto Alegre, 06/01/2006.

MAPLE Power Interval Arithmetic. Disponível em: www.math.uni-wuppertal.de/wrswt/software/intpakX.

MATLAB Intlab. Disponível em: www.mathworks.com.

MOORE, R. E. **Interval Analysis**. Englewood Cliffs: Prentice Hall, 1966.

RATSCHEK, H.; ROKNE, J. **New Computer Methods for Global Optimization**. Ellis Horwood, 1988.