

Probabilidade Intervalar e Cadeias de Markov Intervalares no Maple¹

M.A. CAMPOS², Centro de Informática, UFPE, 50740-540 Recife, PE, Brasil

G.P. DIMURO, A.C.R. COSTA, J.F.F. ARAÚJO, A.M. DIAS³, Escola de Informática, UCPel, 96010-000 Pelotas, RS, Brasil.

Resumo. O objetivo deste trabalho é, inicialmente, apresentar uma implementação do cálculo de probabilidades intervalares utilizando o Maple. Foram desenvolvidos duas bibliotecas: **Mat-Int**, que contempla procedimentos para os operadores intervalares básicos da matemática intervalar, incluindo a álgebra matricial intervalar, e **Prob-Int**, para a implementação da probabilidade intervalar, contendo procedimentos para o cálculo de probabilidades intervalares para variáveis aleatórias discretas. A seguir, apresenta-se a noção de cadeia de Markov intervalar e sua implementação no Maple. Ao final, descreve-se uma aplicação utilizando os conceitos propostos.

1. Introdução

No desenvolvimento de aplicações numéricas em Ciência e Tecnologia surgem erros de computação, os quais originam-se, primordialmente, da impossibilidade de se modelar e representar grandezas contínuas em uma entidade de natureza finitária, como o computador. O sistema de ponto flutuante da máquina não é capaz de representar exatamente os números reais, nem os resultados de operações com esses números. Além disso, como um sistema algébrico, suas características e propriedades algébricas são extremamente pobres quando comparadas com as do sistema de números reais [6].

A matemática intervalar [10] é uma teoria matemática, com origem na década de 60, que teve por objetivo inicial responder a questão da exatidão e da eficiência que aparece na prática da computação científica. Desde então, a utilização de técnicas intervalares tem sido uma alternativa para alcançar limites garantidos para resultados de computações, através do controle rigoroso e automático da propagação dos erros dos dados e parâmetros iniciais ao longo do processo computacional, assim como dos erros de arredondamento e truncamento. Algoritmos intervalares, em contraste com os algoritmos pontuais, computam um intervalo como solução, com a garantia de que a resposta exata pertence a este intervalo.

¹Este trabalho foi parcialmente financiado pela FAPERGS e CNPq.

²mac@cin.ufpe.br

³{liz,rocha, fadriq,amd}@atlas.ucpel.tche.br

Atualmente, a Matemática Intervalar ultrapassou as fronteiras das aplicações numéricas sendo também utilizada em Soft Computing (veja mais sobre esse tema em <http://link.springer.de/link/service/journals/00500/index.htm>).

A probabilidade intervalar foi proposta por Campos [3] como uma metodologia para estender a probabilidade real de forma a possibilitar seu cálculo automático através do uso de intervalos. Esta probabilidade intervalar tem uma conceituação similar à probabilidade usual, sendo então definida como uma função satisfazendo a um conjunto de axiomas. Esta função assume valores no espaço dos intervalos [10], e, portanto, é fundamentada na análise intervalar, justificando o nome *probabilidade intervalar*. Adicionalmente, as operações aritméticas são realizadas segundo a aritmética de exatidão máxima [7]. Esta probabilidade intervalar controla os erros decorrentes da representação dos reais por números de ponto flutuante [6].

Na literatura são mencionadas outros tipo de probabilidades intervalares [12, 8, 13], que diferem da abordagem por nós adotada. Nestes trabalhos, um intervalo de probabilidades é definido de tal forma que inclua as diversas opiniões de especialistas sobre um mesmo dado problema. Um intervalo representando uma probabilidade intervalar nesta outra abordagem deve dar conta de todas as possíveis variações que ocorrem nos julgamentos humanos sobre um mesmo objeto.

Uma aplicação de interesse da probabilidade intervalar é no cálculo das probabilidades de transição de uma cadeia de Markov. O estudo das cadeias de Markov tem sido importante para as ciências sociais, biológicas e administrativas. Os modelos de Markov escondidos [2], por exemplo, que surgiram originalmente no domínio de reconhecimento do discurso, atualmente têm sido empregados em de visão computacional e reconhecimento de manuscritos, de formas, gestos e expressões faciais, entre outros (veja em <http://www-sig.enst.fr/~cappe>).

Este trabalho mostra uma aplicação da probabilidade intervalar no Maple [9], pelo desenvolvimento de uma biblioteca, **Prob-Int**, que associada à biblioteca intervalar **Mat-Int** [5], implementa os principais conceitos da probabilidade intervalar. O artigo introduz também o conceito de cadeias de Markov intervalares e apresenta uma aplicação utilizando o Maple.

O artigo está organizado conforme descrito a seguir. A seção 2 apresenta a biblioteca intervalar **Mat-Int**. A seção 3 é dedicada à probabilidade intervalar e à biblioteca **Prob-Int**, que foi implementada, principalmente, para calcular probabilidades intervalares para variáveis aleatórias discretas. A seção 4 introduz as cadeias de Markov intervalares e sua implementação no Maple. Na última seção tem-se as conclusões.

2. Implementação da Matemática Intervalar

2.1. O Software Maple

O software matemático Maple [9] 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. Possui uma linguagem de programação fundamentada no conceito de linguagem interpretada e um mecanismo de construção e distribuição de pacotes de programas e funções. É largamente utilizado em computação científica, e também na implementação de protótipos de sistemas de grande porte que, após testados no Maple, são posteriormente implementados em outras linguagens, que proporcionem um processamento mais rápido do que sistemas de computação algébricos. Assim, a opção pelo software Maple para a realização deste trabalho fundamenta-se na sua potencialidade de aplicações tanto na pesquisa quanto no ensino.

2.2. O Pacote Intervalar Mat-Int

Uma biblioteca no Maple é uma tabela de procedimentos. Na biblioteca `Mat-Int`, o tipo `intervalo` é definido por uma lista de dois elementos do tipo real, onde o primeiro elemento desta é menor ou igual ao segundo. O tipo `matriz-int` é definido como um `array` de intervalos. A biblioteca `Mat-Int` [5], na versão atual, apresenta os operadores intervalares: (i) básicos (aritmética, interseção, união, inclusão, distância, módulo, valor absoluto, etc), (ii) de funções (extensões intervalares para as principais funções reais), (iii) da álgebra matricial intervalar, (iv) de métodos intervalares para solução de sistemas intervalares lineares e (iv) para os arredondamentos direcionados, onde o extremo inferior do intervalo é arredondado por falta e o superior, por excesso (isto permite o encapsulamento do erro, principalmente na realização de uma seqüência de operações intervalares).

Exemplo 1. Na biblioteca `Mat-Int`, o operador `&pot`, para o cálculo de potências intervalares de expoente n , é definido como:

```
>mat_int ['&pot']:= (X::intervalo,n::integer)->
>   'if'(type(X,list),
>       'if'(type(n,nonnegint),
>           'if'(X[1] > 0 or type(n,odd),
>               [X[1]^n,X[2]^n],
>               'if'(X[2] < 0 and type(n,even),
>                   [X[2]^n,X[1]^n] ,
>                   [0, &mod(X)^n])),
>           'if'(type(n,negint) and X[1]<>0 and X[2]<>0,
>               'if'(X[1]>0,
>                   [X[2]^n,X[1]^n],
>                   'if'(X[2]<0,
>                       'if'(type(n,odd),
>                           [X[2]^n,X[1]^n],
>                           [X[1]^n,X[2]^n]),
>                       ERROR("division by zero"))),
>               ERROR("division by zero or invalid exponent"))),
>       X^n);
> end;
```

A aplicação do operador `&pot` no cálculo de potências do intervalo $[-1, 2]$ é apresentada abaixo. O comando `with(mat_int)` carrega a biblioteca `Mat-Int`. O operador `avalia` é utilizado para avaliar o resultado com a precisão indicada, respeitando os arredondamentos direcionados.

```
> with(mat_int):
> [-1,2] &pot 4;
                                [0, 16]
> [-1,2] &pot 3;
                                [-1, 8]
> [-1,2] &pot (-6);
Error, (in &pot) division by zero
> [-3,-1] &pot (-6}
                                [1/729, 1]
> &avalia(%,3);
                                [0.0137, 1]
> [-3,-1] &pot (-6}
                                [-1, -1/243]
> &avalia(%,4);
                                [-1, -.004115]
```

3. A Probabilidade Intervalar e a Biblioteca Prob-Int

3.1. A Probabilidade Intervalar

A probabilidade intervalar [3] é realizada através de uma peculiar composição de funções envolvendo extensões intervalares de funções reais. Inicialmente a um evento A é associada uma probabilidade $P(A) = p$, que, por sua vez, é associada a um intervalo que a contenha. Do ponto de vista da implementação, p sendo não representável em um específico sistema de ponto flutuante, é substituído pelo menor intervalo de máquina que o contém propriamente.

Definição 1. *Seja \mathbb{IR} o conjunto de todos os intervalos de máquina, cujos extremos são números de um sistema de ponto flutuante \mathbb{F} . A função de aproximação $F_{Id} : \mathbb{R} \rightarrow \mathbb{IR}$ associa cada $x \in \mathbb{R}$ ao menor intervalo $X \in \mathbb{IR}$ que aproxima $x \in \mathbb{R}$ em \mathbb{F} .*

Exemplo 2. *A função F_{Id} é implementada na biblioteca `Mat-Int`, pelo operador `&avalia`, de acordo com a precisão desejada:*

```
> with(mat_int):
> X:=1/3;
                                X := 1/3
> X1:= X &avalia 3;
                                X2 := [.333, .334]
> SOMA:= X &soma X;
```

SOMA := 2/3

Avaliação da soma com uma precisão igual a 3:

```
> &avalia(SOMA,3);
[.666, .667]
```

Resultado da soma utilizando o intervalo X1, cujos extremos foram previamente arredondados considerando uma precisão igual a 3:

```
> X1 &soma X1;
[.666, .668]
```

Definição 2. *Seja \mathbb{A} uma σ -álgebra e P a função probabilidade real. A função probabilidade intervalar P_v é definida por $P_v : \mathbb{A} \rightarrow \mathbb{IR}$, tal que $P_v(A) = (F \circ P)(A)$, onde F é a função de aproximação F_{Id} ou a extensão intervalar de uma função aritmética real.*

Definição 3. *Seja X uma variável aleatória discreta com uma função de probabilidade real $P(X = k)$, $k \in \mathbb{R}_X$, onde \mathbb{R}_X é o contra-domínio da variável aleatória. A função de probabilidade intervalar, $P_v(X = k)$, satisfaz às seguintes propriedades: (i) $P_v(X = k) \in \mathbb{IR}$ e (ii) $1 \in \sum_{k \in \mathbb{R}_X} P_v(X = k)$.*

3.2. Probabilidades Intervalares para Variáveis Aleatórias Discretas

Em [4] foram calculadas probabilidades intervalares para as variáveis aleatórias discretas Bernoulli, Binomial, Poisson, Poisson Truncada, Geométrica, Hipergeométrica e Pascal. A biblioteca intervalar **Prob-Int** [1] implementa operadores para estas variáveis aleatórias. Introduzem-se aqui alguns desses operadores.

Definição 4. *Seja $P(A) = p$ e X uma variável aleatória que conta o número de ocorrências de A nas n repetições independentes do experimento. A função de probabilidade intervalar binomial é definida por*

$$P_v(X = k) = \binom{n}{k} p^k q_v^{n-k}, k = 0, \dots, n.$$

Exemplo 3. *Implementação da Binomial intervalar na biblioteca Prob-Int.*

```
> prob_int['pB']:= proc(k,n,p,prec)
>     local q, pB;
>     q:= 1 &sub p;
>     Pb:= (binomial(n,k)) &mult (p &pot k) &mult (q &pot(n-k));
>     &avalia(Pb,prec);
> end;
```

Supondo $n = 3$, $p = 1/3$ e um precisão igual a 3, tem-se os seguintes resultados; observa-se que o intervalo somatório é uma aproximação do valor 1.

4. Cadeias de Markov Intervalares

Nesta seção serão combinados os métodos da álgebra linear intervalar [11] com a probabilidade intervalar, para introduzir as cadeias de Markov intervalares.

Definição 6. *Uma cadeia de Markov intervalar é aquela em que se consideram probabilidades intervalares.*

Definição 7. *Seja uma cadeia de Markov intervalar. Seja $p_{i,j}$ a probabilidade intervalar de transição do estado i para o estado j . Então a matriz $N \times N$, dada por $P = [p_{i,j}]$, denomina-se matriz intervalar de transição da cadeia de Markov intervalar.*

Definição 8. *Seja $p_{i,j}^n$ a probabilidade intervalar de transição do estado i para o estado j em n passos. Então a matriz $P^n = [p_{i,j}^n]$, dada pela potência intervalar de ordem n da matriz P , é a matriz intervalar de transição de n passos da cadeia.*

Exemplo 5. *A matriz intervalar de transição de n passos é obtida utilizando o operador `&pot_matriz` da biblioteca `Mat-Int`:*

```
> mat_int['&pot_matriz'] := (P::matriz_int, n::posint) ->
>   if (coldim(P) = rowdim(P))
>     then if n = 1
>           then P
>           else &pot_matriz(P, n-1) &mult_matriz P
>         fi;
>     else ERROR("invalid operation");
>   fi;
```

5. Aplicações

Nesta seção introduz-se uma aplicação trivial das cadeias de Markov intervalares em problemas de Genética, adaptando-se um exemplo extraído de [14].

Certas características das plantas e dos animais são determinadas por um par de genes, cada um dos quais podendo ser de dois tipos, denotados por A e a . Existem três genótipos possíveis: AA , Aa e aa (os genótipos Aa e aA são idênticos). Em alguns casos esses três genótipos resultam em três características distintas e em outros o AA e o Aa exibem uma mesma forma observável. Nesta última situação, diz-se que o gene A domina o gene a . O indivíduo chama-se dominante se tem o genótipo AA , heterozigoto se tem genótipo Aa e recessivo se tem o genótipo aa . Por conveniência, denota-se um indivíduo AA por D , um Aa por H e um aa por R . No caso de cruzamento, o filho herda um gene de cada um dos pais. Admita-se que as probabilidades intervalares, com precisão 5, dos genótipos dos filhos de acordo com os dos pais sejam as dadas nas Tabelas 1, 2 e 3, a seguir.

As cadeias de Markov intervalares podem auxiliar em cálculos sobre hereditariedade, como descrito neste próximo exemplo.

Tabela 1: Probabilidades intervalares dos genótipos do filho de dois indivíduos H

D (AA)	H (Aa)	R (aa)
[0.24999, 0.25001]	[0.49999, 0.50001]	[0.24999, 0.25001]

Tabela 2: Probabilidades intervalares dos genótipos do filho de um indivíduo H com outro D

D (AA)	H (Aa)	R (aa)
[0.49999, 0.50001]	[0.49999, 0.50001]	[0, 0]

Exemplo 6. *Suponha que no tempo 0, um indivíduo é acasalado com outro, sendo este do tipo H . No tempo 1, o produto do acasalamento é novamente acasalado com um indivíduo H . O processo repete-se então da mesma maneira. Considera-se como estado do sistema no tempo n o genótipo do n -ésimo filho. Tem-se como resultado uma cadeia de Markov com três estados (D, H, R), cuja matriz intervalar de transição é definida utilizando-se o Maple e as bibliotecas `Mat-Int` e `Prob-Int`:*

```
> P:= array(1..3,1..3,[[[.49999, .50001], [.49999, .50001], [0, 0]],
    [[0.24999, 0.25001], [.49999, .50001], [0.24999, 0.25001]],
    [[0, 0], [.49999, .50001], [.49999, .50001]]]);
    [[.49999, .50001], [.49999, .50001], [0, 0]]
P := [[.24999, .25001], [.49999, .50001], [.24999, .25001]]
    [[0, 0], [.49999, .50001], [.49999, .50001]]
```

Algumas matrizes intervalares de transição foram calculadas:

```
> with(mat_int);
> P2:=&avalua_matriz(P &pot_matriz 2,5); %(duas fases, precisao 5)
    [[.37498, .37502], [.49998, .50002], [.12499, .12501]]
P2 := [[.24998, .25002], [.49997, .50003], [.24998, .25002]]
    [[.12499, .12501], [.49998, .50002], [.37498, .37502]]
> P50:=&avalua_matriz(P &pot_matriz 50,3); %(50 fases, precisao 3)
    [[.249, .251], [.499, .501], [.249, .251]]
P50 := [[.249, .251], [.499, .501], [.249, .251]]
    [[.249, .251], [.499, .501], [.249, .251]]
```

Pela observação da matriz intervalar de transição de dois passos P^2 , que apresenta todas as entradas positivas, conclui-se que esta matriz intervalar aproxima uma matriz real regular que tem uma distribuição de equilíbrio v aproximada pelo vetor probabilístico intervalar $V = [v_1, v_2, v_3]$, com $v_1, v_2, v_3 \in \mathbb{IR}$, tal que $VP \equiv P$. O

Tabela 3: Probabilidades intervalares dos genótipos do filho de um indivíduo H com outro R

D (AA)	H (Aa)	R (aa)
[0, 0]	[0.49999, 0.50001]	[0.49999, 0.50001]

sistema intervalar correspondente é:

$$\begin{aligned}
 [.5, .5] v_1 + [.25, .25] v_2 &\equiv v_1 \\
 [.5, .5] v_1 + [.5, .5] v_2 + [.5, .5] v_3 &\equiv v_2 \\
 [.25, .25] v_2 + [.5, .5] v_3 &\equiv v_3 \\
 1 \in v_1 + v_2 + v_3.
 \end{aligned}$$

A solução deste sistema resulta na menor aproximação intervalar V (com precisão igual a 3) da distribuição real de equilíbrio $v = [.25, .5, .25]$, isto é,

$$V = [[.249, .251] \quad [.499, .501] \quad [.249, .251]] \supseteq v = [.25, .5, .25].$$

6. Conclusões

Este trabalho apresentou uma biblioteca intervalar implementada no Maple para a probabilidade intervalar. Introduziu o conceito de cadeias de Markov intervalares e mostrou uma aplicação. As facilidades de programação no Maple possibilitaram a implementação do tipo intervalo e das operações sobre o tipo. A partir disso, uma biblioteca intervalar (**Mat-Int**) foi desenvolvida, seguindo-se a da probabilidade intervalar (**Prob-Int**). Esta última suportou o cálculo de probabilidades intervalares de transição para cadeias de Markov. A introdução das cadeias de Markov intervalares permitirá a análise e o tratamento da incerteza em fenômenos markovianos, assim como o controle automático e rigoroso dos resultados em computações envolvendo esses modelos.

Agradecimentos

Os autores agradecem aos referees anônimos por suas valiosas observações e importantes sugestões. Salienta-se também o apoio recebido do Grupo de Matemática e Fundamentos da Computação (UCPel/UFPE/UFRN).

Abstract. The aim of this work is to present an implementation of the calculus of interval probabilities using Maple. Two Maple libraries were developed – the library **Mat-Int** which contains procedures for the basic interval operators of interval mathematics and interval matrix algebra, and the library **Prob-Int**, for interval probabilities, containing procedures for the calculus of interval probabilities for discrete random variables. The paper also presents the notion of interval Markov chain and its Maple implementation, and describes an application using those concepts.

Referências

- [1] J.F.F. Araújo, “Probabilidades Intervalares com Aplicações no Maple”, ESIN/UCPel, Pelotas, RS, 2001. (disponível em <http://gmc.ucpel.tche.br/fmc>)
- [2] H. Bunke e T. Caelli (Eds), Hidden Markov models applied in computer vision, em “Machine Perception and Artificial Intelligence”, *World Scientific*, Vol. 45, 2001.
- [3] M.A. Campos, “Uma Extensão Intervalar para a Probabilidade Real”, Tese de Doutorado, Centro de Informática, UFPE, PE, 1997.
- [4] M.A. Campos, Interval probabilities, application to discrete random variables, em “Seleção do XXII CNMAC” (E.X.L. de Andrade, J.M. Balthazar, S.M. Gomes, G.N. Silva, A. Sri Ranga, eds.), *Tendências em Matemática Aplicada e Computacional*, Vol. 1, No. 2, pp. 333-344, SBMAC, 2000.
- [5] A.M. Dias e G.P. Dimuro, “Matemática Intervalar com Aplicações no Maple”, ESIN/UCPel, Pelotas, 2000. (disponível em <http://gmc.ucpel.tche.br/mat-int>)
- [6] D. Goldberg, What every computer scientist should know about floating-point arithmetic, *ACM Computing Surveys*, **23**, No. 1 (1991), 5-48.
- [7] U.W. Kulisch e W.L. Miranker, “Computer Arithmetic in Theory and Practice”, Academic Press, New York, 1981.
- [8] H.E. Kyburg Jr., “Interval-valued Probabilities”, (disponível em <http://www.ensmain.rug.ac.be/ipp>, acessado em Maio, 2002).
- [9] M.B. Monagan, K.O. Geddes, K.M. Heal, G. Labahn and S.M. Vorkoetter, “Maple V: Program. Guide”, Springer, New York, 1998.
- [10] R.E. Moore, “Methods and Applications of Interval Analysis”, SIAM, Philadelphia, 1979.
- [11] A. Neumaier, “Interval Methods for Systems of Equations”, Cambridge University Press, Cambridge, 1990.
- [12] B. Tessem, Interval probability propagation, *International Journal of Approximate Reasoning*, **7** (1992), 95-120.
- [13] K. Weichselberger, Axiomatic foundations of the theory of interval-probability, em “Symposia Gaussiana, Conference B: Statistical Sciences”, pp 47-64, Munich, 1993.
- [14] W. Yoseloff, “Finite Mathematics”, Worth Publishing, New York, 1975.