

Atividade com C/OpenMP usando paralelismo funcional com *task*

Um pesquisador possui um conjunto de amostras agrupadas em matrizes, nas quais cada coluna representa uma amostra. No intuito de inferir sobre os dados obtidos, este pesquisador precisa calcular em paralelo, para cada amostra (coluna) as seguintes métricas: **média aritmética**, **média harmônica**, **mediana**, **moda**, **variância**, **desvio padrão** e **coeficiente de variação**. Cada métrica pode ser definida da seguinte forma:

Média aritmética: Somatório de todos os elementos da amostra, divididos pelo tamanho da amostra (somas das linhas da coluna da matriz divididos pela quantidade de linhas);

$$\frac{1}{n} \sum_{i=1}^n x_i$$

Média harmônica: Razão entre o tamanho da amostra e o somatório do inverso das amostras:

$$\frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

Mediana: Elemento médio da amostra (elemento médio da **coluna ordenada**). Para um número par de elementos, a mediana é a média entre os elementos do meio $((n/2+n/2+1)/2)$.

$$\frac{n+1}{2} \text{ ou } \frac{n}{2} \text{ e } \frac{n}{2} + 1$$

Moda: Elemento mais frequente da amostra (elemento que mais aparece na coluna, se houver mais de um, considera-se somente o primeiro. Se não houver, retorna -1).

Variância: Soma dos quadrados das diferenças entre o elemento da amostra e a **média aritmética** calculada.

$$\sum_{i=1}^n (x_i - \bar{x})^2$$

Desvio padrão: Raiz quadrada da **variância**.

$$\sigma = \sqrt{\sum_{i=1}^N p_i (x_i - \mu)^2}$$

Coeficiente de variação: Razão entre o **desvio padrão** e a **média aritmética**.

$$c_v = \frac{\sigma}{\mu}.$$

Exemplo:

Dada a matriz A[6,4]:

```
9 8 4 5
4 12 20 40
8 8 4 4
8 12 4 21
33 44 20 1
10 18 17 10
```

As métricas calculadas são:

Média aritmética: 12.0, 17.0, 11.5, 13.5

Média harmônica: 8.1, 12.1, 6.6, 3.7

Mediana: 8.5, 12.0, 10.5, 7.5

Moda: 8.0, 8.0, 4.0, -1.0

Variância: 110.0, 188.4, 68.7, 217.9

Desvio Padrão: 10.5, 13.7, 8.3, 14.8

Coefficiente de variação: 0.9, 0.8, 0.7, 1.1

O objetivo deste exercício é construir uma solução paralela para o cálculo das métricas estatísticas solicitadas utilizando os conceitos de OpenMP relativos à diretiva *task*.

Considere como entrada um arquivo de texto contendo, na primeira linha, a quantidade de linhas e colunas da matriz, separadas por um único espaço. Nas linhas posteriores, estão os elementos da matriz do tipo **double**, onde as linhas são separadas por uma quebra de linha simples e as colunas por um único espaço. A matriz deve ser lida por meio do redirecionamento de fluxo de entrada (**stdin**), ou seja, não é necessário usar ponteiros para arquivos.

Conteúdo do arquivo de entrada, por exemplo, **entrada.txt**:

```
6 4
9 8 4 5
4 12 20 40
8 8 4 4
8 12 4 21
33 44 20 1
10 18 17 10
```

Para executar no **bash**, por exemplo, utilize este padrão:

```
./est_desc < entrada.txt <enter>
```

Obs: na linha de comando acima, considera-se que o programa foi inserido em **est_desc.c** e o executável chama-se **est_desc** e está no diretório atual.

A saída deve ser impressa, utilizando o *output (stdout)* padrão, apenas com os elementos correspondentes de cada métrica solicitada, separados por um espaço simples. É importante notar que haverá um espaço a mais no final da linha. Cada métrica é separada por uma quebra de linha simples. A impressão deve ser feita considerando uma única casa decimal.

Saída:

```
12.0 17.0 11.5 13.5  
8.1 12.1 6.6 3.7  
8.5 12.0 10.5 7.5  
8.0 8.0 4.0 -1.0  
110.0 188.4 68.7 217.9  
10.5 13.7 8.3 14.8  
0.9 0.8 0.7 1.1
```

Entregue os três arquivos abaixo compactados em um arquivo **.zip** apenas. Por favor, não troque o padrão de compactação (entregue no padrão **zip**).

- 1) Projeto do algoritmo paralelo usando a Metodologia PCAM;
- 2) Código fonte em C com boas práticas de programação; e
- 3) *Makefile*.

Observações:

- 1) Coloque nos documentos entregues os nomes dos integrantes do grupo que realmente participaram do desenvolvimento.
- 2) Desenvolva e teste o seu algoritmo no cluster do LaSDPC. O exercício será corrigido nele.