

Bài tập thực hành 04

Nhân ma trận dùng bộ nhớ chia sẻ

I. Đề bài:

Mở rộng bài toán nhân hai ma trận trong bài thực hành 4 với kỹ thuật sử dụng bộ nhớ chia sẻ (Tính ma trận tích $P = M \times N$ với M có kích thước $m \times n$ và N có kích thước $n \times p$).

Yêu cầu:

- Cài đặt thuật toán 1: nhân ma trận trên CPU.
- Cài đặt thuật toán 2: nhân ma trận trên GPU (không dùng bộ nhớ chia sẻ).
- Cài đặt thuật toán 3: nhân ma trận trên GPU (dùng bộ nhớ chia sẻ).
- Đo thời gian thực thi của thuật toán nhân ma trận trên GPU và CPU, và tính tỉ lệ:

$$\text{time_rate_1} = \text{thời gian của Thuật toán 1} / \text{thời gian của Thuật toán 2}$$

- So sánh độ chính xác của thuật toán nhân ma trận trên GPU (không dùng bộ nhớ chia sẻ) so với CPU. Độ lỗi được tính toán như sau:

Gọi P_GPU là ma trận kết quả được tính bởi GPU và P_CPU là ma trận kết quả được tính bởi CPU.

$$\text{err_1} = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{p-1} (P_GPU[i, j] - P_CPU[i, j])^2}{\sum_{i=0}^{m-1} \sum_{j=0}^{p-1} P_CPU[i, j]^2}$$

- Đo thời gian thực thi của thuật toán nhân ma trận trên GPU dùng và không dùng bộ nhớ chia sẻ

$$\text{time_rate_2} = \text{thời gian của Thuật toán 2} / \text{thời gian của Thuật toán 3}$$

- So sánh độ chính xác của thuật toán nhân ma trận trên GPU (dùng bộ nhớ chia sẻ) so với CPU bằng cách tính độ lỗi (giá trị err_2).
- Báo cáo: lần lượt thử nghiệm với các kích thước ma trận 16x16, 48x48, 144x144, 432x432, 1296x1296, 3888x3888, 11664x11664. Với CPU, GPU không dùng bộ nhớ chia sẻ, GPU dùng bộ nhớ chia sẻ, báo cáo lại thời gian thực thi & độ lỗi, vẽ đồ thị và cho nhận xét.

II. Cấu trúc tập tin và thư mục:

Tham số dòng lệnh:

Tên tập tin input và output được truyền dưới dạng tham số dòng lệnh như sau:

`<MSSV>.exe <tên_tập_tin_input> <tên_tập_tin_output>`

Ví dụ: Sinh viên có mã số 07K2099 có:

- Tập tin thực thi là **07K2099.exe**
- Tập tin input là **input.txt**
- Tập tin output là **output.txt**

Tham số dòng lệnh sẽ là:

`07K2099.exe input.txt output.txt`

Tập tin input:

Tập tin input được truyền vào trong tham số dòng lệnh thứ 2 (argv[1]):

Cấu trúc:

- Dòng 1: giá trị m và n
- m dòng tiếp theo: ma trận M:
- Dòng m+2: giá trị n và p
- n dòng tiếp theo: ma trận N

Ví dụ:

Cho hai ma trận

M =	1	2	5
	3	5	6
N =	4	6	
	9	8	
	2	7	

Nội dung tập tin input sẽ là

`input.txt`

Giải thích

2 3	Ma trận M có 2 dòng, 3 cột
1 2 5	Dòng 1 của ma trận M
3 5 6	Dòng 2 của ma trận M
3 2	Ma trận N có 3 dòng, 2 cột
4 6	Dòng 1 của ma trận N
9 8	Dòng 2 của ma trận N
2 7	Dòng 3 của ma trận N

Tập tin output:

Tập tin input được truyền vào trong tham số dòng lệnh thứ 3 (argv[2]):

Cấu trúc tập tin output như sau:

- Dòng 1: Thời gian chạy trên CPU
- Dòng 2: Thời gian chạy trên GPU (không dùng bộ nhớ chia sẻ)
- Dòng 3: Thời gian chạy trên GPU (dùng bộ nhớ chia sẻ)
- time_rate_1
- err_1
- time_rate_2
- err_2

Ví dụ: Ma trận tích $P = M \times N$ của trong ví dụ tập tin input ở trên sẽ được lưu trong tập tin output như sau:

Output	Giải thích
0.036	<i>Thời gian nhân ma trận trên CPU</i>
0.00024	<i>Thời gian nhân ma trận trên GPU (không dùng bộ nhớ chia sẻ)</i>
0.000005	<i>Thời gian nhân ma trận trên GPU (dùng bộ nhớ chia sẻ)</i>
100	<i>time_rate_1: Tỷ lệ thời gian time_rate CPU/GPU (không dùng bộ nhớ chia sẻ)</i>
0.012	<i>err_1: Độ lỗi GPU (không dùng bộ nhớ chia sẻ) so với CPU</i>
48	<i>time_rate_2: Tỷ lệ thời gian time_rate GPU không dùng / GPU dùng bộ nhớ chia sẻ</i>
0.012	<i>err_2: Độ lỗi GPU (dùng bộ nhớ chia sẻ) so với CPU</i>

Cấu trúc thư mục

Trong thư mục MSSV gồm có:

- Thư mục *Source*: chứa toàn bộ source code chương trình.
- File *Report.pdf*

III. Quy định Chấm Bài:

- Trong chương trình cuối cùng đem nộp *không được dùng bất cứ hàm input từ màn hình nào* (như *getch*, *scanf*, *scanf_s*, *cin* v.v...).
- Đối với những bài giống nhau sẽ bị 0.