

Multicore Computing Lecture09 - Matrix Multiplication & Gaussian Elimination



남 범 석 bnam@skku.edu



Some Applications of Matrix

- In Physics related scientific domains, matrices are used in Electrical circuits, Quantum mechanics, Optics, and many more.
- Stochasic matrics and Eigen vector solvers are used in the page rank algorithms
- Encryption of message codes (cryptocurrency, etc)
- 3D modeling in Computer Graphics and Vision
- In robotics and automation, matrices are the base elements for the robot movements



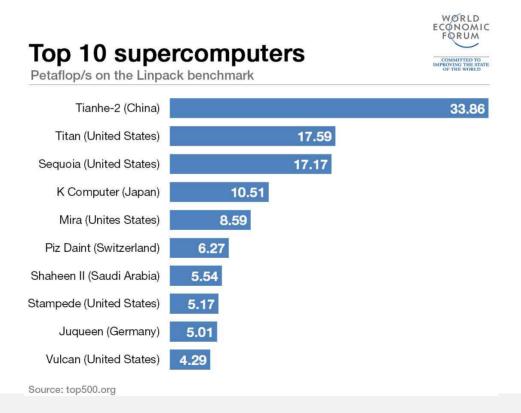
Linear Algebra Packages

- Basic Linear Algebra Subroutines (BLAS)
 - Level 1 (vector-vector): vectorization
 - Level 2 (matrix-vector): vectorization, parallelization
 - Level 3 (matrix-matrix): parallelization
- LINPACK (Fortran)
 - Linear equations and linear least-squares
- EISPACK (Fortran)
 - Eigenvalues and eigenvectors for matrix classes
- LAPACK (Fortran, C) (LINPACK + EISPACK)
 - Use BLAS internally
- ScaLAPACK (Fortran, C, MPI) (scalable LAPACK)



Supercomputers

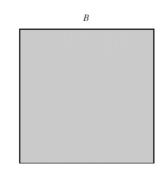
- Supercomputer: a computer with a high computing performance
- Top500.org
 - measures how fast a computer solves a dense N by N linear equations
 Ax = b, which is a common task in various domains.
 - HPL (High Performance LINPACK benchmark)





- $-A \times B = C$
- A[i,:] B[:,j] = C[i,j]
- Row partitioning
 - N tasks

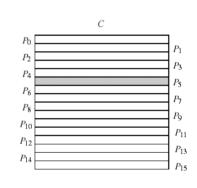
A X



(a)

(b)

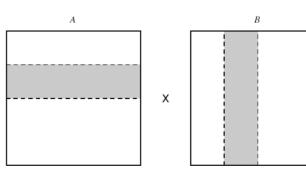
Χ



=

- Block partitioning
 - N*N/B tasks

Shading shows data sharing in B matrix



P_0	P_1	P_2	P ₃
P_4	P ₅	P_6	P ₇
P_8	P9	P ₁₀	P ₁₁
P ₁₂	P ₁₃	P ₁₄	P ₁₅



- The number of operations required = O(m x n x r)
- For simplicity, we analyze square matrices of order n. So, O(n3)



Sequential algorithm

Square Matrix Multiplication

```
for (i=0;i<n; i++)
{
    for (j=0; j<n; j++)
    {
        c[i][j] = 0;
        for (k=0; k<n; k++)
        {
            c[i][j] += a[i][k] * b[k][j];
        }
    }
}</pre>
```



Assumption:

- The number of processors available in parallel machine is p
- The processing nodes are homogeneous
 - Homogeneity make it possible to achieve load balancing

Parallelization

- Step 1) Partition the two matrices into p square blocks
- Step 2) Each square block of A and B are assigned to each process
 - Initial alignment is needed (shown in the next slide)
- Step 3) p processors process p blocks.
 - Each processor multiplies blocks and add the results to partial results in C.
- Step 4) The A blocks are rolled one step to the left and B blocks are rolled one step upward
- Repeat step 3 and 4 p^{1/2} times



A =
$$\begin{bmatrix} 2 & 1 & 5 & 3 & 7 & 3 ... \\ 0 & 7 & 1 & 6 & 2 & 1 ... \\ 9 & 2 & 4 & 4 & 3 & 2 ... \\ 3 & 6 & 7 & 2 & 5 & 9 ... \\ 1 & 3 & 2 & 0 & 3 & 1 ... \\ 4 & 2 & 0 & 1 & 2 & 0 ... \end{bmatrix}$$

$$B = \begin{bmatrix} 6 & 1 & 2 & 3 & 0 & 2 \dots \\ 4 & 5 & 6 & 5 & 2 & 1 \dots \\ 1 & 9 & 8 & -8 & 1 & 2 \dots \\ 4 & 0 & -8 & 5 & 0 & 8 \dots \\ 2 & 3 & 0 & 1 & 1 & 2 \dots \\ 0 & 1 & 2 & 3 & 4 & 0 \dots \\ \dots \end{bmatrix}$$

Initial alignment

0	1 7	5 1		2	1	
4 7	4 2	3 5	2 9	 9	2 6	←
3 2	1	 1 4	3 2	2 0	0	—

4 5	-8 5	4 0
		•••
1 9	0 1	0 2
4 0	2 3	2 1
	•••	
2 3	2 3	1 2
0 1	6 5	0 8
	†	†

8 -8

1 2

2x2 blocks

$$A = \begin{bmatrix} 2 & 1 & 5 & 3 \\ 0 & 7 & 1 & 6 \\ 9 & 2 & 4 & 4 \\ 3 & 6 & 7 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 6 & 1 & 2 & 3 \\ 4 & 5 & 6 & 5 \\ 1 & 9 & 8 & -8 \\ 4 & 0 & -8 & 5 \end{bmatrix}$$

 2 1
 5 3
 6 1
 8 -8

 0 7
 1 6
 4 5
 -8 5

 4
 4

 7
 2

 3
 6

Initial alignment



$$A = \begin{bmatrix} 2 & 1 & 5 & 3 \\ 0 & 7 & 1 & 6 \\ 9 & 2 & 4 & 4 \\ 3 & 6 & 7 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 6 & 1 & 2 & 3 \\ 4 & 5 & 6 & 5 \\ 1 & 9 & 8 & -8 \\ 4 & 0 & -8 & 5 \end{bmatrix}$$

B =
$$\begin{bmatrix} 6 & 1 & 2 & 3 \\ 4 & 5 & 6 & 5 \\ 1 & 9 & 8 & -8 \\ 4 & 0 & -8 & 5 \end{bmatrix}$$

P_{0,0}

P_{0,1}

P_{1,0}

P_{1,1}



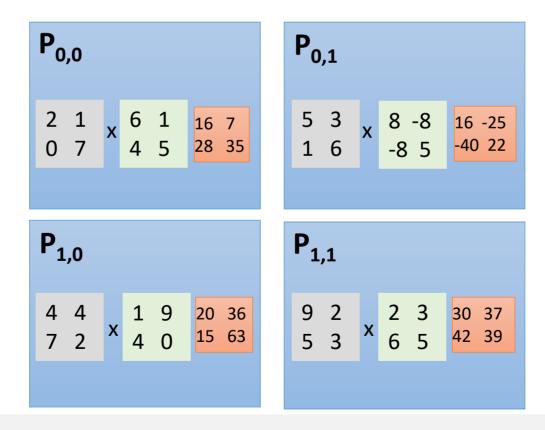
$$A = \begin{bmatrix} 2 & 1 & 5 & 3 \\ 0 & 7 & 1 & 6 \\ 9 & 2 & 4 & 4 \\ 3 & 6 & 7 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 6 & 1 & 2 & 3 \\ 4 & 5 & 6 & 5 \\ 1 & 9 & 8 & -8 \\ 4 & 0 & -8 & 5 \end{bmatrix}$$

P_{0,0} 2 1 x 6 1 4 5 16 7 28 35 P_{0,1} 5 3 x 8 -8 1 6 -8 5 16 -25 -40 22

P_{1,0} 1 9 4 4 20 36 7 2 × 4 0 15 63 P_{1,1} 9 2 x 2 3 6 5 30 37 42 39

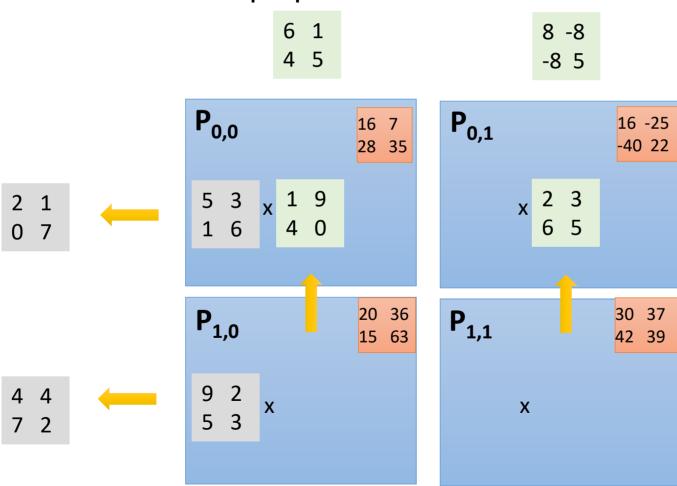


- Shift A one step to left
- Shift B one step up



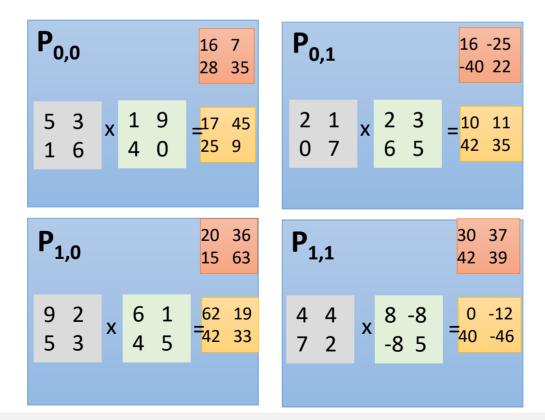


- Shift A one step to left
- Shift B one step up





- Shift A one step to left
- Shift B one step up





Done



HPL (High-Performance Linpack Benchmark)

- HPL is a parallel, blocked, LU decomposition solver.
- Suppose you want to solve these equations

$$5x + 3y - 2z = 23$$

 $7x + 9y + 3z = 102$
 $8x + 8y - 8z = 8$

You could put them into the form:

$$Ax = b$$

$$\begin{bmatrix} 5 & 3 & -2 \\ 7 & 9 & 3 \\ 8 & 8 & -8 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 23 \\ 102 \\ 8 \end{bmatrix}$$



Taking advantage of Properties of A

$$Ax = b$$

- A is a given n x n matrix
- b is a given n-vector
- x is unknown solution n-vector
- Properties of A can make this equation easier to solve.
 - If A = lower triangular matrix → forward substitution.
 - If A = upper triangular matrix → backward substitution.
- How to change the matrix to upper or lower triangular?



Gaussian Elimination for a System of Linear Equations

$\blacksquare Ax = b$

$$a_{0,0}x_0 + a_{0,1}x_1 + \dots + a_{0,n-1}x_{n-1} = b_0$$

 $a_{1,0}x_0 + a_{1,1}x_1 + \dots + a_{1,n-1}x_{n-1} = b_1$
...
$$A_{n-1,0}x_0 + a_{n-1,1}x_1 + \dots + a_{n-1,n-1}x_{n-1} = b_{n-1}$$

- Gaussian elimination (classic algorithm)
 - Forward elimination to Ux=y (U is upper triangular)
 - without or with partial pivoting
 - Back substitution to solve for x
 - Parallel algorithms based on partitioning of A



Gaussian Elimination for a System of Linear Equations

$$\left[egin{array}{ccc|c} 1 & 3 & 1 & 9 \ 1 & 1 & -1 & 1 \ 3 & 11 & 5 & 35 \end{array}
ight]
ightarrow \left[egin{array}{ccc|c} 1 & 3 & 1 & 9 \ 0 & -2 & -2 & -8 \ 0 & 2 & 2 & 8 \end{array}
ight]
ightarrow \left[egin{array}{ccc|c} 1 & 3 & 1 & 9 \ 0 & -2 & -2 & -8 \ 0 & 0 & 0 & 0 \end{array}
ight]
ightarrow \left[egin{array}{ccc|c} 1 & 0 & -2 & -3 \ 0 & 1 & 1 & 4 \ 0 & 0 & 0 & 0 \end{array}
ight]$$

- Use Elementary Row Operations
 - Swapping two rows,
 - Multiplying a row by a nonzero number,
 - Adding a multiple of one row to another row.
- Forward Elimination: Upper Triangular Matrix
 - row1 \rightarrow row2 \rightarrow row3
 - Then, Backward Substitution (x3 \rightarrow x2 \rightarrow x1)



Inverse Matrix using Gaussian Elimination



Sequential Gaussian Elimination

```
procedure GAUSSIAN ELIMINATION (A, b, y)
1.
2.
    begin
3.
       for k := 0 to n - 1 do /* Outer loop */
4.
      begin
5.
               for j := k + 1 to n - 1 do
6.
                       A[k, j] := A[k, j]/A[k, k];
7.
               y[k] := b[k]/A[k, k];
8.
               A[k, k] := 1;
9.
               for i := k + 1 to n - 1 do
10.
               begin
11.
                    for j := k + 1 to n - 1 do
12.
                       A[i, j] := A[i, j] - A[i, k] \times A[k, j];
13.
                   b[i] := b[i] - A[i, k] \times y[k];
14.
                   A[i, k] := 0;
15.
               endfor;
16.
       endfor;
17. end GAUSSIAN ELIMINATION
```

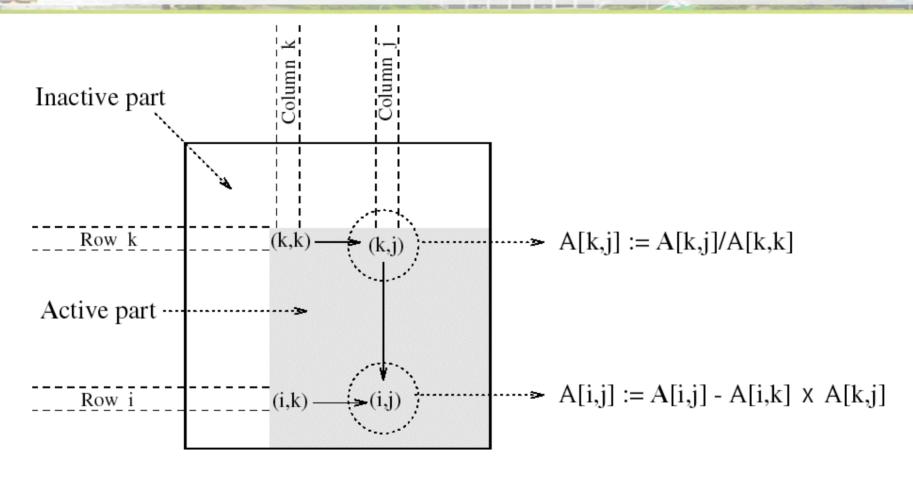


Gaussian Elimination

- Gaussian elimination : triple-nested loop
- for k

 for i $a_{ij} = a_{ij} (a_{ik}/a_{kk}) a_{kj}$ end
 end
 end

Computation Step in Gaussian Elimination



$$5x + 3y = 22$$

 $8x + 2y = 13$

$$x = (22 - 3y) / 5$$

 $8(22 - 3y)/5 + 2y = 13$



$$5x + 3y = 22$$

 $8x + 2y = 13$
 $x = (22 - 3y) / 5$
 $8(22 - 3y) / 5 + 2y = 13$
 $x = (22 - 3y) / 5$
 $y = (13 - 176/5) / (24/5 + 2)$



Row-wise Partitioning on Eight Processes

P_0	1	(0,1)	(0,2)	(0,3)	(0,4)	(0,5)	(0,6)	(0,7)
P ₁	0	1	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	(1,7)
P ₂	0	0	1	(2,3)	(2,4)	(2,5)	(2,6)	(2,7)
P ₃	0	0	0	(3,3)	(3,4)	(3,5)	(3,6)	(3,7)
P ₄	0	0	0	(4,3)	(4,4)	(4,5)	(4,6)	(4,7)
P ₅	0	0	0	(5,3)	(5,4)	(5,5)	(5,6)	(5,7)
P ₆	0	0	0	(6,3)	(6,4)	(6,5)	(6,6)	(6,7)
P ₇	0	0	0	(7,3)	(7,4)	(7,5)	(7,6)	(7,7)

P_0	1	(0,1)	(0,2)	(0,3) (0,4) (0,5) (0,6) (0,7)
P_1	0	1	(1,2)	(1,3) (1,4) (1,5) (1,6) (1,7)
P_2	0	0	1	(2,3) (2,4) (2,5) (2,6) (2,7)
P ₃	0	0	0	1 (3,4) (3,5) (3,6) (3,7)
P_4	0	0	0	(4,3) Y (4,4) Y (4,5) Y (4,6) Y (4,7)
P ₅	0	0	0	(5,3) Y (5,4) Y (5,5) Y (5,6) Y (5,7)
P ₆	0	0	0	(6,3) V (6,4) V (6,5) V (6,6) V (6,7)
P ₇	0	0	0	(7,3) \$\dagge(7,4)\$\dagge(7,5)\$\dagge(7,6)\$\dagge(7,7)\$

(a) Computation:

(i) A[k,j] := A[k,j]/A[k,k]

(ii) A[k,k] := 1

(b) Communication:

One-to-all broadcast of row A[k,*]



Row-wise Partitioning on Eight Processes

P_0	1	(0,1)	(0,2)	(0,3)	(0,4)	(0,5)	(0,6)	(0,7)
P ₁	0	1	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	(1,7)
P ₂	0	0	1	(2,3)	(2,4)	(2,5)	(2,6)	(2,7)
P ₃	0	0	0	1	(3,4)	(3,5)	(3,6)	(3,7)
P ₄	0	0	0	(4,3)	(4,4)	(4,5)	(4,6)	(4,7)
P ₅	0	0	0	(5,3)	(5,4)	(5,5)	(5,6)	(5,7)
P ₆	0	0	0	(6,3)	(6,4)	(6,5)	(6,6)	(6,7)
P ₇	0	0	0	(7,3)	(7,4)	(7,5)	(7,6)	(7,7)

(c) Computation:

- (i) $A[i,j] := A[i,j] A[i,k] \times A[k,j]$ for k < i < n and k < j < n
- (ii) A[i,k] := 0 for k < i < n

■ P4 ~ P7 : parallel computation



2D Partitioning on 64 Processes

1	(0,1)	(0,2)	(0,3)	(0,4)	(0,5)	(0,6)	(0,7)
0	1	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	(1,7)
0	0	1	(2,3)	(2,4)	(2,5)	(2,6)	(2,7)
0	0	0	(3,3)	(3,4)	(3,5)		(3,7)
0	0	0			(4,5)		
0	0	0			(5,5)		
0	0	0	(6,3)		(6,5)		
0	0	0	(7,3)	(7,4)	(7,5)	(7,6)	

(a) Rowwise broadcast of A[i,k] for (k - 1) < i < n

1	(0,1)	(0,2)	(0,3)	(0,4)	(0,5)	(0,6)	(0,7)
0	1	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	(1,7)
0	0	1	(2,3)	(2,4)	(2,5)	(2,6)	(2,7)
0	0	0	1	(3,4)	(3,5)	(3,6)	(3,7)
0	0	0	(4,3)	(4,4)	(4,5)	(4,6)	(4,7)
0	0	0	(5,3)	(5,4)	(5,5)	(5,6)	(5,7)
0	0	0	(6,3)	(6,4)	(6,5)	(6,6)	(6,7)
0	0	0	(7,3)	(7,4)	(7,5)	(7,6)	(7,7)

(c) Columnwise broadcast of A[k,j] for k < j < n

1	(0,1)	(0,2)	(0,3)	(0,4)	(0,5)	(0,6)	(0,7)
0	1	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	(1,7)
0	0	1	(2,3)	(2,4)	(2,5)	(2,6)	(2,7)
0	0	0	(3,3)	(3,4)	(3,5)	(3,6)	(3,7)
0	0	0	(4,3)	(4,4)	(4,5)	(4,6)	(4,7)
0	0	0	(5,3)	(5,4)	(5,5)	(5,6)	(5,7)
0	0	0	(6,3)	(6,4)	(6,5)	(6,6)	(6,7)
0	0	0	(7,3)	(7,4)	(7,5)	(7,6)	(7,7)

(b) A[k,j] := A[k,j]/A[k,k]for k < j < n

1	(0,1)	(0,2)	(0,3)	(0,4)	(0,5)	(0,6)	(0,7)
0	1	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	(1,7)
0	0	1	(2,3)	(2,4)	(2,5)	(2,6)	(2,7)
0	0	0	1	(3,4)	(3,5)	(3,6)	(3,7)
0	0	0	(4,3)	(4,4)	(4,5)	(4,6)	(4,7)
0	0	0	(5,3)	(5,4)	(5,5)	(5,6)	(5,7)
0	0	0	(6,3)	(6,4)	(6,5)	(6,6)	(6,7)
0	0	0	(7,3)	(7,4)	(7,5)	(7,6)	(7,7)

(d) $A[i,j] := A[i,j]-A[i,k] \times A[k,j]$ for k < i < n and k < j < n



Back Substitution to Find Solution

```
    procedure BACK SUBSTITUTION (U, x, y)
    begin
    for k := n - 1 downto 0 do /* Main loop */
    begin
    x[k] := y[k];
    for i := k - 1 downto 0 do
    y[i] := y[i] - x[k] * U[i, k];
    endfor;
    end BACK SUBSTITUTION
```



Drawback of Gaussian Elimination

- Drawback of Gaussian Elimination → lots of computations
 - n³/3 additions and multiplications
 - n²/2 divisions.
 - Equations, especially {b}, have to be changed in each step
 - What if we want to solve the equation for a different b?
 - Can we do better?

