Vv417 Lecture 4

Jing Liu

UM-SJTU Joint Institute

September 17, 2019

Definition

A square matrix in which all the entries below the main diagonal are 0 is called an upper triangular matrix. In case of all the entries above the main diagonal are 0,

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}, \qquad \mathbf{L} = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix}$$

it is called a lower triangular matrix. A square matrix in which all the entries off the main diagonal are zero is called a diagonal matrix, which is also triangular.

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix} = \operatorname{diag}(d_1, d_2, \dots, d_n)$$

Q: Is every diagonal matrix invertible?

Q: Why must the inverse of an invertible diagonal matrix also be diagonal?

Theorem

A diagonal matrix is invertible if and only if all of its diagonal entries are nonzero;

$$\mathbf{D} = \operatorname{diag}(d_1, d_2, \cdots, d_n) \quad \text{where} \quad d_i \neq 0 \quad \text{for all } i.$$

$$\mathbf{D}^{-1} = \operatorname{diag}(1/d_1, 1/d_2, \cdots, 1/d_n).$$

- Note powers and products of diagonal matrices are easy to compute

$$\mathbf{D}^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^k \end{bmatrix} \quad \text{and} \quad \mathbf{D}\mathbf{D}^* = \begin{bmatrix} d_1 d_1^* & 0 & \cdots & 0 \\ 0 & d_2 d_2^* & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n d_n^* \end{bmatrix}$$

where d_k^* is the kth diagonal element of the diagonal matrix $\mathbf{D}^*$.

Q: Is the product of triangular matrices also triangular?

Exercise

Find the inverse of $\mathbf{A} = \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & 4 \\ 0 & 0 & 5 \end{bmatrix}$.

Solution

$$\left[\begin{array}{ccc|ccc} 1 & 3 & -1 & 1 & 0 & 0 \\ 0 & 2 & 4 & 0 & 1 & 0 \\ 0 & 0 & 5 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 3 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0.5 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0.2 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1.5 & 1.4 \\ 0 & 1 & 0 & 0 & 0.5 & -0.4 \\ 0 & 0 & 1 & 0 & 0 & 0.2 \end{array}\right]$$

Theorem

1. The product of upper triangular matrices is upper triangular.
2. The product of lower triangular matrices is lower triangular.
3. A triangular matrix is invertible if and only if its diagonals are nonzero.
4. The inverse of an invertible upper triangular matrix is upper triangular.
5. The inverse of an invertible lower triangular matrix is lower triangular.

- To prove the invertibility of a lower triangular matrix $\mathbf{L}$ and the form of $\mathbf{L}^{-1}$, we need the following theorem, which involves the notion of transpose.

Theorem

If $\mathbf{A}$ is an invertible matrix, then $\mathbf{A}^{\mathrm{T}}$ is also invertible and

$$(\mathbf{A}^{\mathrm{T}})^{-1} = (\mathbf{A}^{-1})^{\mathrm{T}}$$

Definition

The transpose of a matrix $\mathbf{A}$, denoted by

$$\mathbf{A}^{\mathrm{T}}$$

is a matrix of which its columns are the rows of $\mathbf{A}$ or vice versa.

Properties of Transpose

Let $\mathbf{A}$ and $\mathbf{B}$ be matrices, and let α be a scalar.

1. $\left(\mathbf{A}^{\mathrm{T}}\right)^{\mathrm{T}} = \mathbf{A}$
2. $(\alpha\mathbf{A})^{\mathrm{T}} = \alpha\mathbf{A}^{\mathrm{T}}$
3. $(\mathbf{A}+\mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}}$
4. $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$

Q: Properties 1. $\sim$ 3. are easy to understand and prove, how can we prove 4.

- Now we ready to prove the last theorem.

Proof

- We need to show either $\mathbf{A}^{\mathrm{T}}(\mathbf{A}^{-1})^{\mathrm{T}} = \mathbf{I}$ or $(\mathbf{A}^{-1})^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} = \mathbf{I}$.

$$\mathbf{A}^{\mathrm{T}}(\mathbf{A}^{-1})^{\mathrm{T}} = \left(\left((\mathbf{A}^{-1})^{\mathrm{T}}\right)^{\mathrm{T}}\left(\mathbf{A}^{\mathrm{T}}\right)^{\mathrm{T}}\right)^{\mathrm{T}} = \left(\mathbf{A}^{-1}\mathbf{A}\right)^{\mathrm{T}} = (\mathbf{I})^{\mathrm{T}} = \mathbf{I}$$

Definition

A square matrix $\mathbf{A}$ is said to be symmetric if

$$\mathbf{A} = \mathbf{A}^{\mathrm{T}}$$

Theorem

If $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices of the same size, and α is a scalar, then

1. $\mathbf{A}^{\mathrm{T}}$ is symmetric.
2. $\alpha\mathbf{A}$ is symmetric.
3. $\mathbf{A}+\mathbf{B}$ is symmetric.

Q: Why the product of symmetric matrices is not necessarily symmetric?

Theorem

If $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices of the same size, then

$$\mathbf{AB}$$

is symmetric if and only if

$$\mathbf{AB} = \mathbf{BA}$$

Proof

- Since $\mathbf{A}$ and $\mathbf{B}$ are symmetric, and the property $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ implies

$$(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} = \mathbf{BA}$$

  from which, we can complete the proof

$$(\mathbf{AB})^{\mathrm{T}} = \mathbf{AB} \iff \mathbf{AB} = \mathbf{BA}$$

Q: In general, is a symmetric matrix always invertible?

Theorem

If $\mathbf{A}$ is an invertible symmetric matrix, then $\mathbf{A}^{-1}$ is symmetric.

Proof

- We need to show
$$\left(\mathbf{A}^{-1}\right)^{\mathrm{T}} = \mathbf{A}^{-1}$$
- Since $\mathbf{A}$ is invertible, by the theorem on page $\boxed{5}$, and $\mathbf{A}$ being symmetric,
$$\left(\mathbf{A}^{-1}\right)^{\mathrm{T}} = \left(\mathbf{A}^{\mathrm{T}}\right)^{-1} = (\mathbf{A})^{-1} = \mathbf{A}^{-1} \quad \square$$

Definition

A square matrix $\mathbf{A}$ is said to be skew-symmetric if
$$\mathbf{A} = -\mathbf{A}^{\mathrm{T}}$$

Theorem

Let $\mathbf{A}$ and $\mathbf{B}$ be square matrices of the same size, and α be a scalar.

1. If $\mathbf{A}$ is invertible and skew-symmetric, then $\mathbf{A}^{-1}$ is skew-symmetric.
2. If $\mathbf{A}$ and $\mathbf{B}$ are skew-symmetric, then $\mathbf{A}^{\mathrm{T}}$, $\alpha\mathbf{A}$, $\mathbf{A}+\mathbf{B}$ are skew-symmetric.
3. Every square matrix $\mathbf{A}$ can be uniquely decomposed into a sum of 2 matrices

$$\mathbf{A} = \operatorname{sym}(\mathbf{A}) + \operatorname{skew}(\mathbf{A})$$

where

$$\operatorname{sym}(\mathbf{A}) = \frac{1}{2}\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)$$

is a symmetric matrix and

$$\operatorname{skew}(\mathbf{A}) = \frac{1}{2}\left(\mathbf{A} - \mathbf{A}^{\mathrm{T}}\right)$$

$\operatorname{skew}(\mathbf{A})$ is a skew-symmetric matrix.

Proof

- Statements 1. and 2. are very similar to their symmetric counterparts.
- Now for statement 3., the decomposition is clearly valid

$$\mathbf{A} = \mathrm{sym}\left(\mathbf{A}\right) + \mathrm{skew}\left(\mathbf{A}\right) = \frac{1}{2}\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right) + \frac{1}{2}\left(\mathbf{A} - \mathbf{A}^{\mathrm{T}}\right)$$

so we simply need to show the two parts

$$\frac{1}{2}\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right) \qquad \text{and} \qquad \frac{1}{2}\left(\mathbf{A} - \mathbf{A}^{\mathrm{T}}\right)$$

are symmetric and skew-symmetric, respectively, and show uniqueness. Since

$$\left(\frac{1}{2}\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)\right)^{\mathrm{T}} = \frac{1}{2}\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)^{\mathrm{T}} = \frac{1}{2}\left(\mathbf{A}^{\mathrm{T}} + \mathbf{A}\right) = \frac{1}{2}\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)$$

$$\left(\frac{1}{2}\left(\mathbf{A} - \mathbf{A}^{\mathrm{T}}\right)\right)^{\mathrm{T}} = \frac{1}{2}\left(\mathbf{A} - \mathbf{A}^{\mathrm{T}}\right)^{\mathrm{T}} = \frac{1}{2}\left(\mathbf{A}^{\mathrm{T}} - \mathbf{A}\right) = -\frac{1}{2}\left(\mathbf{A} - \mathbf{A}^{\mathrm{T}}\right)$$

we conclude they are indeed symmetric and skew-symmetric, respectively.

Proof

- To prove uniqueness, we need to figure it out how in the first place to obtain

$$\operatorname{sym}(\mathbf{A}) = \frac{1}{2}\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)$$
$$\operatorname{skew}(\mathbf{A}) = \frac{1}{2}\left(\mathbf{A} - \mathbf{A}^{\mathrm{T}}\right)$$

- Suppose $\mathbf{A} = \mathbf{P} + \mathbf{Q}$, where $\mathbf{P}$ and $\mathbf{Q}$ are symmetric and skew-symmetric.
- Notice the following must be true if $\mathbf{P}$ is symmetric and $\mathbf{Q}$ is skew-symmetric

$$\mathbf{A}^{\mathrm{T}} = (\mathbf{P} + \mathbf{Q})^{\mathrm{T}} = \mathbf{P}^{\mathrm{T}} + \mathbf{Q}^{\mathrm{T}} = \mathbf{P} - \mathbf{Q}$$

- If we take the sum of and the difference of $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}$, we have

$$\mathbf{A} + \mathbf{A}^{\mathrm{T}} = 2\mathbf{P} \implies \mathbf{P} = \frac{1}{2}\left(\mathbf{A} + \mathbf{A}^{\mathrm{T}}\right)$$
$$\mathbf{A} - \mathbf{A}^{\mathrm{T}} = 2\mathbf{Q} \implies \mathbf{Q} = \frac{1}{2}\left(\mathbf{A} - \mathbf{A}^{\mathrm{T}}\right)$$

- To avoid a contradiction, we have to conclude that it is unique.

- Matrix products of the following form

$$\mathbf{A}\mathbf{A}^{\mathrm{T}} \quad \text{and} \quad \mathbf{A}^{\mathrm{T}}\mathbf{A}$$

arise in many applications, where $\mathbf{A}$ is not necessary a square matrix.

- Note both of the products $\mathbf{A}\mathbf{A}^{\mathrm{T}}$ and $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ are always square matrices.

Q: Are $\mathbf{A}\mathbf{A}^{\mathrm{T}}$ and $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ symmetric?

Theorem

If $\mathbf{A}$ is an invertible matrix, then $\mathbf{A}\mathbf{A}^{\mathrm{T}}$ and $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ are also invertible.

Proof

- If $\mathbf{A}$ is invertible, then $\mathbf{A}^{\mathrm{T}}$ is invertible by the theorem on page 5.
- Because

$$\mathbf{A}\mathbf{A}^{\mathrm{T}} \quad \text{and} \quad \mathbf{A}^{\mathrm{T}}\mathbf{A}$$

are both products of invertible matrices, thus both must be invertible. □

- A matrix can be partitioned into submatrices (also called blocks), e.g.

$$\left[\begin{array}{ccc:c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hdashline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right]$$

for which, if we denote

$$\mathbf{A}_{11} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}, \quad \mathbf{A}_{12} = \begin{bmatrix} a_{14} \\ a_{24} \end{bmatrix},$$
$$\mathbf{A}_{21} = \begin{bmatrix} a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad \mathbf{A}_{22} = \begin{bmatrix} a_{34} \end{bmatrix}$$

then we can view $\mathbf{A}$ as a 2×2 matrix whose entries are themselves matrices.

$$\left[\begin{array}{ccc:c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hdashline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

Definition

A sparse matrix is a matrix in which many or most of the elements are zero.

- Partitioning a matrix into blocks is often useful when the matrix is sparse.
- It is possible to multiply two large block matrices together by treating the blocks as if they are scalars, e.g. Consider two block matrices

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \end{bmatrix}$$

- The product can be computed in the following way

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} & \mathbf{A}_{11}\mathbf{B}_{13} + \mathbf{A}_{12}\mathbf{B}_{23} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} & \mathbf{A}_{21}\mathbf{B}_{13} + \mathbf{A}_{22}\mathbf{B}_{23} \end{bmatrix} \end{aligned}$$

provided all the matrix multiplications needed above are defined.

- This formula is called block matrix multiplication. Although a general proof is possible, we omit it here and focus on the applications of it.

Q: Do you notice that we have seen blocks and used block multiplication?

$$\mathbf{AB} = \left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ \hdashline a_{21} & a_{22} & a_{23} \\ \hdashline a_{31} & a_{32} & a_{33} \end{array}\right] \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \left[\begin{array}{c:c:c:} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{array}\right]$$

Exercise

Compute $\mathbf{AB}$ where $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 \end{bmatrix}$.

Solution

- Consider the following partitionings,

$$\mathbf{A} = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right] = \begin{bmatrix} \mathbf{C} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}$$

$$\mathbf{B} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 \end{array}\right] = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C} & \mathbf{C} \end{bmatrix}$$

- Using block multiplication, the product

$$\mathbf{AB} = \begin{bmatrix} \mathbf{C} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C} & \mathbf{C} \end{bmatrix} = \begin{bmatrix} 2\mathbf{C} & \mathbf{C} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} = \left[\begin{array}{cc|cc} 2 & 4 & 1 & 2 \\ 6 & 8 & 3 & 4 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$$

Definition

A square partitioned matrix $\mathbf{A}$ is said to be block diagonal if the submatrices on the main diagonal are square and all submatrices off the main diagonal are zero,

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_k \end{bmatrix} \quad \text{where} \quad \mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k \quad \text{are square.}$$

Q: Is it true that a block diagonal matrix $\mathbf{A}$ is invertible if and only if diagonals

$$\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$$

are invertible, and inverse is given by

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_1^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_k^{-1} \end{bmatrix}$$

Exercise

Is the following matrix invertible? If so, find $\mathbf{A}^{-1}$. If not, justify your answer.

$$\mathbf{A} = \begin{bmatrix} 8 & -7 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 5 & 2 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

Solution

- If we partition $\mathbf{A}$ as below, then there are three submatrices on the diagonal

$$\mathbf{A} = \left[\begin{array}{cc:cc:c} 8 & -7 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 5 & 2 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 4 \end{array}\right] \Longrightarrow \mathbf{A}^{-1} = \left[\begin{array}{cc:cc:c} 1 & -7 & 0 & 0 & 0 \\ 1 & -8 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & -5 & 3 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 1/4 \end{array}\right]$$

all of which are invertible.

Definition

A square matrix $\mathbf{A}$ is said to be block upper triangular if the submatrices on the main diagonal are square and all submatrices below the main diagonal are zero,

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1k} \\ \mathbf{0} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{kk} \end{bmatrix} \quad \text{where} \quad \mathbf{A}_{11}, \mathbf{A}_{22}, \ldots, \mathbf{A}_{kk} \quad \text{are square.}$$

The definition of a block lower triangular matrix is similar, instead of having zero submatrices below the main diagonal, all submatrices above are zero in this case.

- Many computer algorithm on large matrices exploit block structure to break the computation down into smaller pieces. For example, consider

$$\mathbf{T} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \qquad \text{where } \mathbf{A} \text{ and } \mathbf{C} \text{ are square.}$$

Q: Suppose $\mathbf{T}$ is very large, how can we take advantage of multi-core structure of modern CPUs to invert $\mathbf{T}$ by parallel processing?

- It is clear that the followings are row equivalent by the Gaussian elimination,

$$\left[\begin{array}{cc:cc} \mathbf{A} & \mathbf{B} & \mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} & \mathbf{I}_q \end{array}\right] \sim \left[\begin{array}{cc:cc} \mathbf{A} & \mathbf{B} & \mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_q & \mathbf{0} & \mathbf{C}^{-1} \end{array}\right]$$

Q: How to proceed the next elimination step? What is the matrix needed here?

$$\left[\begin{array}{cc:cc} \mathbf{A}_p & \mathbf{B}_{p\times q} & \mathbf{I}_p & \mathbf{0}_{p\times q} \\ \mathbf{0}_{q\times p} & \mathbf{I}_q & \mathbf{0}_{q\times p} & \mathbf{C}_q^{-1} \end{array}\right]$$

- Now consider the following following matrix multiplications,

$$\begin{aligned} \left[\begin{array}{cc:cc} \mathbf{A} & \mathbf{B} & \mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} & \mathbf{I}_q \end{array}\right] &\sim \left[\begin{array}{cc} \mathbf{I}_p & -\mathbf{B}_{p\times q} \\ \mathbf{0}_{q\times p} & \mathbf{I}_q \end{array}\right] \left[\begin{array}{cc:cc} \mathbf{A}_p & \mathbf{B}_{p\times q} & \mathbf{I}_p & \mathbf{0}_{p\times q} \\ \mathbf{0}_{q\times p} & \mathbf{I}_q & \mathbf{0}_{q\times p} & \mathbf{C}_q^{-1} \end{array}\right] \\ &\sim \left[\begin{array}{cc} \mathbf{A}_p^{-1} & \mathbf{0}_{p\times q} \\ \mathbf{0}_{q\times p} & \mathbf{I}_q \end{array}\right] \left[\begin{array}{cc:cc} \mathbf{A}_p & \mathbf{0}_{p\times q} & \mathbf{I}_p & -\mathbf{B}_{p\times q}\mathbf{C}_q^{-1} \\ \mathbf{0}_{q\times p} & \mathbf{I}_q & \mathbf{0}_{q\times p} & \mathbf{C}_q^{-1} \end{array}\right] \\ &\sim \left[\begin{array}{cc:cc} \mathbf{I}_p & \mathbf{0}_{p\times q} & \mathbf{A}_p^{-1} & -\mathbf{A}_p^{-1}\mathbf{B}_{p\times q}\mathbf{C}_q^{-1} \\ \mathbf{0}_{q\times p} & \mathbf{I}_q & \mathbf{0}_{q\times p} & \mathbf{C}_q^{-1} \end{array}\right] \end{aligned}$$

Q: The matrices in blue are not elementary. Why the eliminations are valid?

- Therefore, if

$$\mathbf{T} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \qquad \text{where } \mathbf{A} \text{ and } \mathbf{C} \text{ are invertible.}$$

then

$$\mathbf{T}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{C}^{-1} \\ \mathbf{0} & \mathbf{C}^{-1} \end{bmatrix}$$

- In general, if

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{D} & \mathbf{C} \end{bmatrix} \qquad \text{where } \mathbf{A} \text{ and } \mathbf{C} \text{ are invertible.}$$

then

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{C} - \mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{D}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{C} - \mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -(\mathbf{C} - \mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{D}\mathbf{A}^{-1} & (\mathbf{C} - \mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix}$$

provided $\mathbf{C} - \mathbf{D}\mathbf{A}^{-1}\mathbf{B}$ is invertible.