Chapter 2. Matrix Algebra (2/2)

Sim, Min Kyu, Ph.D., mksim@seoultech.ac.kr

2.5. Matrix Factorizations

Matrix Factorizations

- A *factorization* of a matrix A is an equation that expresses A as a product of two or more matrices.
- Whereas matrix multiplication involves a *synthesis* of data (combining the effects of two or more linear transformations into a single matrix), matrix factorization is an *analysis* of data.

The LU Factorization

- The LU factorization, described on the next few slides, is motivated by the fairly common industrial and business problem of solving a sequence of equations, all with the same coefficient matrix:

$$A\mathbf{x} = b_1, \quad A\mathbf{x} = b_2, \quad ..., A\mathbf{x} = \mathbf{b}_p \tag{5}$$

- When A is invertible, one could compute A^{-1} and then compute $A^{-1}b_1$, $A^{-1}b_2$, and so on.
- However, it is more efficient to solve the first equation in the sequence (5) by row reduction and obtain the LU factorization of A at the same time. Thereafter, the remaining equations in sequence (5) are solved with the LU factorization.

The LU Factorization

- At first, assume that A is an $m \times n$ matrix that can be row reduced to echelon form, *without row interchanges*.
- Then A can be written in the form $A = LU$, were L is an $m \times n$ lower triangular matrix with 1's on the diagonal and U is an $m \times n$ echelon form of A.
- For instance, see Fig. 1 below. Such a factorization is called an **LU factorization** of A. The matrix L is invertible and is called a unit lower triangular matrix.

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{U}$$

The LU Factorization

- Before studying how to construct L and U, we should look at why they are so useful. When $A = LU$, the equation $A\mathbf{x} = \mathbf{b}$ can be written as $L(U\mathbf{x}) = \mathbf{b}$.
- Writing $\mathbf{y}$ for $U\mathbf{x}$, we can find $\mathbf{x}$ by solving the pair of equations

$$Ly = \mathbf{b}$$
$$U\mathbf{x} = y$$

- First solve $Ly = \mathbf{b}$ for y, and then solve $U\mathbf{x} = y$ for $\mathbf{x}$. See Fig. 2 on the next slide. Each equation is easy to solve because L and U are triangular.

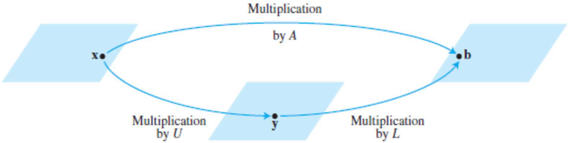


FIGURE 2 Factorization of the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

The LU Factorization

- **Example 1** It can be verified that

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = LU$$

- Use this factorization of A to solve $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$

- **Solution**
  - The solution of $L\mathbf{y} = \mathbf{b}$ needs only 6 multiplications and 6 additions, because the arithmetic takes place only in column 5.

$$[L \mid \mathbf{b}] = \left[\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{array}\right] \sim \left[\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right] = [I \mid \mathbf{y}]$$

  - Then, for $U\mathbf{x} = \mathbf{y}$, the "backward" phase of row reduction requires 4 divisions, 6 multiplications, and 6 additions.
  - For instance, creating the zeros in column 4 of $[U \mid \mathbf{y}]$ requires 1 division in row 4 and 3 multiplication - addition pairs to add multiples of row 4 to the rows above.

$$[U|\mathbf{y}] = \begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$$

  - To find $\mathbf{x}$ requires 28 arithmetic operations, or "*flops*' ' (floating point operations), excluding the cost of finding L and U. In contrast, row reduction of $[A|\mathbf{b}]$ to $[I|\mathbf{x}]$ takes 62 operations.

An LU Factorization Algorithm

- Suppose A can be reduced to an echelon form U using only row replacements that add a multiple of one row to another below it.
- In this case, there exist unit lower triangular elementary matrices $E_1, ..., E_p$ such that $E_p ... E_1 A = U$. Then,

$$A = (E_p ... E_1)^{-1} U = LU \qquad (3)$$

where

$$L = (E_p ... E_1)^{-1} \qquad (4)$$

- It can be shown that products and inverses of unit lower triangular matrices are also unit lower triangular. Thus L is unit lower triangular.
- Note that row operations in equation (3), which reduce A to U, also reduce the L in equation (4) to I, because $E_p ... E_1 L = (E_p ... E_1)(E_p ... E_1)^{-1} = I$. This observation is the key to *constructing L*.

An LU Factorization Algorithm

Algorithm for an LU Factorization

- Step 1) Reduce A to an echelon form U by a sequence of row replacement operations, if possible. e.g.) $(R2 \leftarrow R2 - 3R1)$
- Step 2) Place entries in L such that the same sequence of row operations reduces L to I.
- Step 1 is not always possible, but when it is, the argument above shows that an LU factorization exists.
- Example 2 on the followings slides will show how to implement Step 2. By construction, L will satisfy

$$(E_p ... E_1)L = I$$

using the same $E_p ... E_1$ as in equation (3). Thus L will be invertible, by the Invertible Matrix Theorem, with $(E_p ... E_1) = L^{-1}$. From (3), $L^{-1}A = U$, and $A = LU$. So Step 2 will produce an acceptable L.

An LU Factorization Algorithm

- **Example 2** Find an LU factorization of

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

- **Solution:**
  - Since A has four rows, L should be 4×4. The first column of L is the first column of A divided by the top pivot entry:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & & 1 & 0 \\ -3 & & & 1 \end{bmatrix}$$

- **(Solution continued:)**
  - Compare the first columns of A and L. The row operations that create zeros in the first column of A will also create zeros in the first column of L.
  - To make this same correspondence of row operations on A hold for the rest of L, watch a row reduction of A to an echelon form U. That is, *highlight the entries* in each matrix that are used to determine the sequence of row operations that transform A onto U.

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} = A_1$$

$$\sim A_2 = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = U$$

- **(Solution continued:)**
  - The highlighted entries above determine the row reduction of A to U. At each pivot column, divide the highlighted entries by the pivot and place the result onto L:
  - **An easy calculation verifies that this L and U satisfy $LU = A$.**

$$\begin{bmatrix} 2 \\ -4 \\ 2 \\ -6 \end{bmatrix} \begin{bmatrix} 3 \\ -9 \\ 12 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 5 \end{bmatrix}$$

$$\div 2 \quad \div 3 \quad \div 2 \quad \div 5$$

$$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow$$

$$\begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ 1 & -3 & 1 & \\ -3 & 4 & 2 & 1 \end{bmatrix}, \quad \text{and} \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix}$$

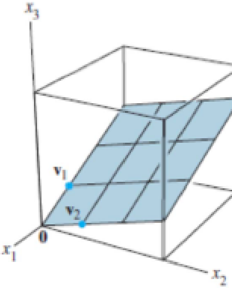
Suggested Exercises

- 2.5.7
- 2.5.9
- 2.5.11

2.8. Subspaces of $\mathbb{R}^n$

Subspaces of $\mathbb{R}^n$

- **Definition:** A **subspace** of $\mathbb{R}^n$ is any set H in $\mathbb{R}^n$ that has three properties:
  a) The zero vector is in H.
  b) For each $\mathbf{u}$ and $\mathbf{v}$ in H, the sum $\mathbf{u} + \mathbf{v}$ is in H. (*closed under vector addition*)
  c) For each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H. (*closed under scalar multiplication*)
- A plane through the origin is the standard way to visualize the subspace in Example 1 on the next slide. See Fig. 1 below:



Subspaces of $\mathbb{R}^n$

- **Example 1** If v_1 and v_2 are in $\mathbb{R}^n$ and $H = Span\{v_1, v_2\}$, prove that H is a subspace of $\mathbb{R}^n$.

- **Proof**
  1. To verify this statement, note that the zero vector is in H (because $v_1 + v_2$ is a linear combination of v_1 and v_2).
  2. Now take two arbitrary vectors in H, say

     $$u = s_1 v_1 + s_2 v_2 \quad and \quad v = t_1 v_1 + t_2 v_2$$

     Then,

     $$u + v = (s_1 + t_1)v_1 + (s_2 + t_2)v_2,$$

     which shows that $u + v$ is a linear combination of v_1 and v_2 and hence is in H.
  3. Also, for any scalar c, the vector cu is in H, because

     $$cu = c(s_1 v_1 + s_2 v_2) = cs_1(v_1) + cs_2(v_2)$$

Column space and Null space of a matrix

- **Definition:** The **column space** of a matrix A is the set $Col\ A$ of all linear combinations of the columns of A.
- If $A = [\mathbf{a}_1, ..., \mathbf{a}_n]$ with the columns of $\mathbb{R}^m$, then $Col\ A$ is the same as $Span\{a_1...a_n\}$.
- Example 4 shows that the column space of an $m \times n$ matrix is a subspace of $\mathbb{R}^m$.

- **Example 4** Determine whether $\mathbf{b}$ is in the column space of A, where

$$A = \begin{bmatrix} 1 & -3 & -4 \\ -4 & 6 & -2 \\ -3 & 7 & 6 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 3 \\ 3 \\ -4 \end{bmatrix}.$$

- **Solution:**
  - The vector $\mathbf{b}$ is a linear combination of the columns of A if and only if $\mathbf{b}$ can be written as $A\mathbf{x}$ for some $\mathbf{x}$. That is, if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
  - Row reducing the augmented matrix $[A \mid \mathbf{b}]$,

$$\left[\begin{array}{ccc|c} 1 & -3 & -4 & 3 \\ -4 & 6 & -2 & 3 \\ -3 & 7 & 6 & -4 \end{array}\right] \sim \begin{bmatrix} 1 & -3 & -4 & 3 \\ 0 & 6 & -18 & 15 \\ 0 & -2 & -6 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & -4 & 3 \\ 0 & 6 & -18 & 15 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

  - We conclude that
    - $A\mathbf{x} = \mathbf{b}$ is **consistent**
    - $\mathbf{b}$ is in $Col\ A$.

Column space and Null space of a matrix

- **Definition:** The **null space** of a matrix A is the set $Nul\ A$ of all solutions of the homogeneous equation $A\mathbf{x} = 0$.

- **Theorem 12:** The null space of an $m \times n$ matrix A is a subspace of $\mathbb{R}^n$. Equivalently, the set of all solutions of a system $A\mathbf{x} = 0$ of m homogenous linear equations in n unknowns is a subspace of $\mathbb{R}^n$.

- **Proof:**
  1. The zero vector is in $NulA$ (because $A0 = 0$).
  2. To show that $NulA$ satisfies that other two properties required for a subspace, take any $\mathbf{u}$ and $\mathbf{v}$ in $Nul\ A$. That is, suppose $A\mathbf{u} = 0$ and $A\mathbf{v} = 0$. Then, by a property of matrix multiplication, $A(u + v) = Au + Av = 0 + 0 = 0$. Thus, $\mathbf{u} + \mathbf{v}$ satisfies $A = 0$, and so $\mathbf{u} + \mathbf{v}$ is in $NulA$.
  3. Also, if $\mathbf{u} \in Nul\ A$, then for any scalar c, $A(c\mathbf{u}) = c(A\mathbf{u}) = c(0) = 0$, which shows that $c\mathbf{u}$ is in $Nul\ A$.

Basis for a subspace

- **Definition:** A **basis** for a subspace H of $\mathbb{R}^n$ is a linearly independent set in H that spans H.
- **Example 5**
  - The columns of an invertible $n \times n$ matrix form a basis because they are linearly independent and span $\mathbb{R}^n$, by the Invertible Matrix Theorem.
  - One such matrix is the $n \times n$ identity matrix. Its columns are denoted by $e_1, ..., e_n$:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix}, e_1 = \begin{bmatrix} 0 \\ 1 \\ . \\ . \\ . \\ 0 \end{bmatrix}, \ e_n = \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ . \\ 1 \end{bmatrix}$$

  - The set $\{e_1, ..., e_n\}$ is called the **standard basis** for $\mathbb{R}^n$. See Fig. 3 on the next slide.

Basis for a subspace

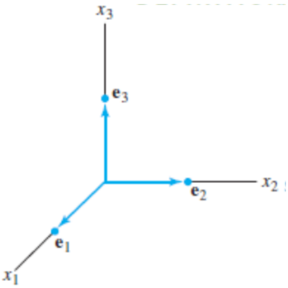


FIGURE 3
The standard basis for $\mathbb{R}^3$.

- **Theorem 13:** The pivot columns of a matrix A form a basis for the column space of A.

Suggested Exercises

- 2.8.11
- 2.8.12

2.9. Dimension and rank

The dimension of a subspace

- **Definition:** The **dimension** of a nonzero subspace H, denoted by $dim\ H$, is the number of vectors in any basis for H. The dimension of the zero subspace $\{0\}$ is defined to be zero.

- **Definition:** The **rank** of a matrix A, denoted by $rank\ A$, is the dimension of the column space of A.

The dimension of a subspace

- **Example 3** Determine the rank of the matrix

$$A \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & -6 & 4 & 14 & -20 \\ 0 & -9 & 6 & 5 & -6 \end{bmatrix}$$

- **Solution:**
  - Reduce A to echelon form:

$$\mathrm{A} \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & -6 & 4 & 14 & -20 \\ 0 & -9 & 6 & 5 & -6 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & 0 & 0 & 4 & -6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Pivot columns

- The matrix A has 3 pivot columns, so $rank\ A = 3$.

The dimension of a subspace

- **Theorem 14** If a matrix A has n columns, then $rank\, A + dim\, Nul\, A = n$.
- **Theorem 15** Let H be a p-dimensional subspace of $\mathbb{R}^n$. Any linearly independent set of exactly p elements in H is automatically a basis for H. Also, any set of p elements of H that spans H is automatically a basis for H.

Recap - The invertible matrix theorem

- **Theorem 8:** Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.
  - **a)** A is an invertible matrix.
  - **b)** A is row equivalent to the $n \times n$ identity matrix.
  - **c)** A has n pivot positions.
  - **d)** The equation $A\mathbf{x} = 0$ has only the trivial solution.
  - **e)** The columns of A form a linearly independent set.
  - **f)** The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
  - **g)** The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.
  - **h)** The columns of A span $\mathbb{R}^n$.
  - **i)** The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$
  - **j)** There is an $n \times n$ matrix C such that $CA = I$.
  - **k)** There is an $n \times n$ matrix D such that $AD = I$.
  - **l)** A^T is an invertible matrix.

Addendum on the Invertible Theorem

- **The Invertible Theorem (continued)**

  m) The columns of A form a basis of $\mathbb{R}^n$.
  n) $Col\ A = \mathbb{R}^n$
  o) $dim\, Col\ A = n$
  p) $rank\ A = n$
  q) $Nul\ A = \{0\}$
  r) $dim\ Nul\ A = 0$

The proof for the added statements

- Statement (m) is logically equivalent to statements (e) and (h) regarding linear independence and spanning.
- The other five statements are linked to the earlier ones of the theorem by the following chain of almost trivial implications:
  $(g) \Rightarrow (n) \Rightarrow (o) \Rightarrow (p) \Rightarrow (r) \Rightarrow (q) \Rightarrow (d)$
- Statement (g), which says that the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$, implies statement (n), because $Col\ A$ is precisely the set of all $\mathbf{b}$ such that the equation $A\mathbf{x} = \mathbf{b}$ is consistent.
- The implications $(n) \Rightarrow (o) \Rightarrow (p)$ follow from the definitions of $dimension$ and $rank$.
- If the rank of A is n, the number of columns of A, then $dim\ Nul\ A = 0$, by the Rank Theorem, and so $Nul\ A = \{0\}$. Thus $(p) \Rightarrow (r) \Rightarrow (q)$
- Also, statement (q) implies that the equation $A\mathbf{x} = 0$ has only the trivial solution, which is statement (d).
- Since statements (d) and (g) are already known to be equivalent to the statement that A is invertible, the proof is complete.

Suggested Exercises

- Supplementary Exercises
  - (At the end of the chapter, p.178-179)
  - 2.1 (all subproblems)

Acknowledgement

- This lecture note is based on the instructor's lecture notes (originally formatted as ppt files) provided by the publisher (Pearson Education) and the textbook authors (David Lay and others)
- The pdf conversion project for this chapter was possible thanks to the hard work by Taehyeon Kwon (ITM 17'). Mr. Kwon was a captain for ITM's soccer club "ITM United" from Feb 2021 to Feb 2022. In Professor Sim's lab, he developed reinforcement learning applications for optimal use strategies of energy storage systems (ESS).