Introduction to Linear Algebra

Midterm 1 Review Class

Wang Tianyu[1]

Fall, 2019

Outline

1. A system of linear equations is **consistent** if it has at least one solution;
2. Every system of linear equations has either **zero, one, or infinitely** many solutions;
3. Every **homogeneous** system of linear equations is consistent;
4. **Back substitution**;
5. **Elementary row operations**:
    - Type I: Interchange two rows ($E_{i,j}$)
    - Type II: Multiply a row by a nonzero constant ($E_{(\alpha)i}$)
    - Type III: Add a constant times one row to another ($E_{(\alpha)i,j} \neq E_{i,(\alpha)j}$)
6. Gaussian Elimination: reduce the matrix to **ref** and apply back substitution.

Question1

Construct a system of equations with the general solution

$$\mathbf{x} = c1 \begin{bmatrix} 2 \\ -3 \\ 1 \\ 0 \end{bmatrix} + c2 \begin{bmatrix} -2 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

Question2

For the following equation

$$B = \begin{bmatrix} 1 & 1 & 1+\lambda & \lambda \\ 1 & 1+\lambda & 1 & 3 \\ 1+\lambda & 1 & 1 & 0 \end{bmatrix}$$

for what λ will the equation have

(a) unique solution?

(b) no solution?

(c) infinite many solutions?

1. Reduced row echelon form (rref):
   - If a row does not consist entirely of zeros, then the first nonzero entry of this row is **1**, which is known as a **leading 1**;
   - If the matrix has any rows that consist entirely of zeros, they occur at the **bottom** of the matrix;
   - In any two successive rows that do not consist entirely of zeros, the leading 1 in the **lower** row occurs further to the **right** than the leading 1 above;
   - Each column that contains a leading 1 has zeros everywhere else.
2. Row echelon form (ref): satisfy the first three properties.
3. A linear system is consistent **if and only if** there **exists** a ref of the corresponding augmented matrix that has no row of $[0 \ \cdots \ 0 \ a]$ where $a \neq 0$.
4. Gauss-Jordan Elimination: result in **rref**;
5. The rref of a matrix is **unique**;
6. The number of pivots is the **rank** of the matrix;
7. A **homogeneous** system of linear equations with more unknowns than equations has infinitely many solutions (not hold for nonhomogeneous system).

pivot column: a pivot column is a column of matrix that contains a pivot position, *i.e.*, the leftmost nonzero column.

$$
\left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 2 & 3 & 5 & 8 \\ 4 & 0 & 5 & 2 \end{array}\right]
\xrightarrow{R_2-2R_1}
\left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & 1 & 3 & -2 \\ 4 & 0 & 5 & 2 \end{array}\right]
$$

$$
\xrightarrow{R_3-4R_1}
\left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & 1 & 3 & -2 \\ 0 & -4 & 1 & -18 \end{array}\right]
$$

$$
\xrightarrow{R_3+4R_2}
\left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 13 & -26 \end{array}\right]
$$

$$
\xrightarrow{\frac{1}{13}R_3}
\left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 1 & -2 \end{array}\right]
$$

$$
\xrightarrow{R_2-3R_3}
\left[\begin{array}{ccc|c} 1 & 1 & 1 & 5 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -2 \end{array}\right]
$$

$$
\xrightarrow{R_1-R_3}
\left[\begin{array}{ccc|c} 1 & 1 & 0 & 7 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -2 \end{array}\right]
$$

$$
\xrightarrow{R_1-R_2}
\left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -2 \end{array}\right]
$$

Figure: Working Example of Gaussian-Jordan Elimination

1. $A(BC) = (AB)C$;
2. A matrix can be understood from three perspectives: **row (intersection)**, **column (combination)** and **matrix (inverse image)**;
3. Matrix multiplication by row and by column (left matrix always involved with rows and right matrix always involved with columns);
4. Elementary matrix differs from the identity matrix by **one single** elementary row operation.

Validity with **square matrix**:

1. Mutual relationship of the inverse: $AB = BA = I$ (**unique**);
2. $(AB \cdots Z)^{-1} = Z^{-1} \cdots B^{-1}A^{-1}$;
3. $(A^n)^{-1} = A^{-n} = (A^{-1})^n$;
4. Equivalence theorem;
5. Inversion algorithm (augment the matrix with identity matrix and eliminate the original matrix into identity matrix);
6. Either rref(A) has a row of zeros or rref(A)=I;
7. If AB is invertible, then A and B must also be invertible;

Equivalent statements of a square matrix A being invertible:

- $A\mathbf{x} = 0$ has only the trivial solution;
- The rref of A is identity matrix;
- A is expressible as a product of elementary matrices;
- $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b} \in \mathbb{R}^n$;
- $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $\mathbf{b} \in \mathbb{R}^n$;
- $\det(A) \neq 0$.

Inverse of triangular matrix:

$$\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{a} & -\frac{b}{ad} & \frac{be-cd}{afd} \\ 0 & \frac{1}{d} & -\frac{e}{fd} \\ 0 & 0 & \frac{1}{f} \end{bmatrix}$$

Question1: Suppose A, B and C are matrices where the following matrix multiplications are defined, and A is invertible. Show

$$C(A-B)=A^{-1}B$$

given

$$(A-B)C=BA^{-1}$$

Solution:

$$I+(A-B)C=I+BA^{-1}$$

$$AA^{-1}+(A-B)C-BA^{-1}=I$$

$$(A-B)(C+A^{-1})=I,(C+A^{-1})(A-B)=I$$

$$C(A-B)=A^{-1}B$$

1. Upper triangular matrix, lower triangular matrix, diagonal matrix, symmetric matrix, skew-symmetric matrix, sparse matrix;
2. A diagonal matrix (triangular matrix) is invertible **if and only if** all of its diagonal entries are nonzero;
3. The product and inverse of triangular matrix do not change its property;
4. Symmetric ($A = A^{\mathrm{T}}$) and skew-symmetric ($A = -A^{\mathrm{T}}$);
5. Every square matrix A can be uniquely decomposed into a sum of symmetric and skew-symmetric matrices: $A = \frac{1}{2}(A + A^{\mathrm{T}}) + \frac{1}{2}(A - A^{\mathrm{T}})$;
6. If A is an invertible matrix, then AA^{T} and $A^{\mathrm{T}}A$ are also invertible;
7. Block-form matrix.

Permutation:
Any arrangement of a set $\mathcal{S} = \{1, 2, ..., n\}$ in a specific order, for example,

$$\sigma_{\text{no}} = (1, 2, ..., n) \quad \text{or} \quad \sigma = (k_1, k_2, ..., k_i, ..., k_j, ..., k_n)$$

is called a permutation of $\mathcal{S}$, where σ_{no} above is defined to be in the **nature order**.
Out of the nature order:
A pair of elements (k_i, k_j) in σ satisfies $k_i > k_j$ where $i < j$.

- Even permutation: even number of pairs out of the nature order;
- Odd permutation: odd number of pairs out of the nature order.

Levi-Civita symbol:

$$\epsilon_\sigma = \epsilon_{k_1 \cdots k_n} = \begin{cases} +1 & \text{if } \sigma \text{ is an even permutation} \\ -1 & \text{if } \sigma \text{ is an odd permutation} \\ 0 & \text{otherwise} \end{cases}$$

1. $\det(A) = \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} \epsilon_{k_1 \cdots k_n} a_{1k_1} a_{2k_2} a_{3k_3} \cdots a_{nk_n}$ (this explains why determinant is not defined for non-square matrix);
2. $\det(E_{i,j}A)$=-$\det(A)$;
3. $\det(E_{(\alpha)i}A) = \alpha\det(A)$;
4. $\det(E_{(\alpha)i,j}A) = \det(A)$;
5. $\det(A^{\mathrm{T}}) = \det(A)$;
6. $\det(AB) = \det(A)\det(B)$;
7. $\mathrm{adj}(A) = C^{\mathrm{T}}$;
8. If $\det(A) \neq 0$, then $A^{-1} = \frac{1}{\det(A)}\mathrm{adj}(A)$;
9. The determinant of diagonal and triangular matrices is the product of diagonal entries.

Question1:
Find the determinant of the matrix

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

(using some properties to quickly find the answer)

Reasons to introduce LU decomposition

1. *Previous ways of calculating determinant can sometimes be very tedious when the size becomes large;*
2. *The calculation of a triangular matrixs determinant is simple.(product of diagonal entries);*
3. *Transforming a matrix into triangular form is simple (Gauss Elimination);*
4. *Estimating the influence of trasformation on the determinant is simple(there are only elementary row operations).*

Solving $A\mathbf{x} = \mathbf{b}$ with LU decomposition: first solve $L\mathbf{c} = \mathbf{b}$, then solve $U\mathbf{x} = \mathbf{c}$.

Interesting facts

1. *When a row of A* **starts with** *zeros, so does that row of L; When a column of A* **starts with** *zeros, so does that column of U;*
2. *U has the pivots on its diagonal; L has all the 1's on its diagonal;*
3. *The multipliers $l_{i,j}$ are below the diagonal of L;*
4. *If A is symmetric, it can be shown that $A = LDL^{\mathrm{T}}$.*

Consider $A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & 2 & 1 \\ 6 & -3 & 4 \end{bmatrix}$, we can perform the LU decomposition to obtain $A = E_{(2)1,2}E_{(3)1,3}E_{3,2}U = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -6 & -5 \\ 0 & 0 & -5 \end{bmatrix}$, which seems to contradict with our observation. What goes wrong?

Note: If only **Type III** operations are used when applying Gaussian elimination to reduce A to U, then $A = LU$.

Better balance for this decomposition ...

$A = LU$ is "unsymmetric" because U has the pivots on its diagonals where L has 1's. Therefore, we may prefer more the decomposition $A = LD\hat{U}$, where D is the diagonal matrix with pivots. Each row of $\hat{U}$ is that of the original U divided by its pivot.

If A is **invertible** and $A = \hat{L}D\hat{U}$ **exists** where $\hat{L}$ and $\hat{U}$ are upper and lower triangular with unit diagonal and D is a diagonal matrix, then this decomposition is unique.

Suppose A is an invertible $n \times n$ matrix. The leading principle submatrices A_k of A are invertible for $k = 1, ..., n-1$ **if and only if** A has the decomposition $A = \hat{L}U$.

Back to our question on the last slide, how to adapt LU decomposition to matrix for which row interchanges are necessary during Gaussian elimination.

1. A permutation matrix, often denoted by P, is a product of elementary matrices corresponding to row interchanges;
2. If P is a permutation matrix, then $P^{-1} = P^{\mathrm{T}}$;
3. Every invertible matrix A has a decomposition of the form $A = P^{\mathrm{T}}\hat{L}U$;

In common practice, row exchanges are done in advance, which results in $PA = \hat{L}U$. Here row exchanges mean bringing the pivot into its usual place.

Jacobi iteration (slow convergence for large linear systems):

$$\mathbf{x}^{(k+1)} = D^{-1}(\mathbf{b} + (D - A)\mathbf{x}^{(k)})$$

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right), \text{for } i = 1, 2, ..., n$$

Gauss-Seidel iteration:

$$\mathbf{x}^{(k+1)} = (D + L)^{-1}(\mathbf{b} - U\mathbf{x}^{(k)})$$

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right), \text{for } i = 1, 2, ..., n$$

Suppose we are given the following linear system:

$$
\begin{aligned}
10x_1 - x_2 + 2x_3 &= 6, \\
-x_1 + 11x_2 - x_3 + 3x_4 &= 25, \\
2x_1 - x_2 + 10x_3 - x_4 &= -11, \\
3x_2 - x_3 + 8x_4 &= 15.
\end{aligned}
$$

If we choose $(0, 0, 0, 0)$ as the initial approximation, then the first approximate solution is given by

$$
\begin{aligned}
x_1 &= (6 + 0 - (2 * 0))/10 = 0.6, \\
x_2 &= (25 + 0 + 0 - (3 * 0))/11 = 25/11 = 2.2727, \\
x_3 &= (-11 - (2 * 0) + 0 + 0)/10 = -1.1, \\
x_4 &= (15 - (3 * 0) + 0)/8 = 1.875.
\end{aligned}
$$

Using the approximations obtained, the iterative procedure is repeated until the desired accuracy has been reached. The following are the approximated solutions after five iterations.

x_1	x_2	x_3	x_4
0.6	2.27272	-1.1	1.875
1.04727	1.7159	-0.80522	0.88522
0.93263	2.05330	-1.0493	1.13088
1.01519	1.95369	-0.9681	0.97384
0.98899	2.0114	-1.0102	1.02135

The exact solution of the system is $(1, 2, -1, 1)$.

Figure: Working Example of Jacobi iteration

Convergence of the iteration:

1. A **square** matrix is **strictly diagonally dominant** if $|a_{ii}| > \sum_{j=1}^{i-1} |a_{ij}| + \sum_{j=i+1}^{n} |a_{ij}|$ for all $i = 1, 2, ..., n$.
2. If A is strictly diagonally dominant, then $A\mathbf{x} = \mathbf{b}$ has a unique solution (**invertible**) and for any choice of the initial guess $\mathbf{x}^{(0)}$, the sequence $\{\mathbf{x}^{(k)}\}$ produced by Jacobi or Gauss-Seidel iteration **converge** to the exact solution.

Question: Why is this lecture called "sparse"?
Save storage requirements to large extent.

Ex.1: Suppose $Q^{\mathrm{T}} = Q^{-1}$ (Q is then called **orthogonal matrix**).

(a) Show that the columns $q_1, ..., q_n$ are unit vectors;

(b) Show that every two columns of Q are perpendicular;

(c) Find a 2 by 2 example with first entry $q_{11} = \cos\theta$.

Ex.2: Show that $I + BA$ and $I + AB$ are both invertible or both singular. (*hint*: consider the identity $A(I + BA) = (I + AB)A$)

Ex.3: Find the inverse of $\begin{bmatrix} 1 & -a & 0 & 0 \\ 0 & 1 & -b & 0 \\ 0 & 0 & 1 & -c \\ 0 & 0 & 0 & 1 \end{bmatrix}$ with **proper** method.

Thank you!

Good luck for Midterm 1!