CS3430 S24: Scientific Computing

Lecture 02: Gauss-Jordan Reduction, Determinants, Leibnitz' Formula, Cramer's Rule

Vladimir Kulyukin
Department of Computer Science
Utah State University

Review of Lecture 01: Linear Systems and Matrices

Recall that in Lecture 01, we established that there is a meaningful connection between linear systems and matrices in that a linear system can be represented as a matricx. Let's consider this linear system of 3 equation in 3 unknowns:

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= 39 \\ 2x_1 + 3x_2 + x_3 &= 34 \\ x_1 + 2x_2 + 3x_3 &= 26 \end{aligned}$$

This linear system can be represented with two matrices.

$$\begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 2 & 3 \end{bmatrix} \quad \begin{bmatrix} 39 \\ 34 \\ 26 \end{bmatrix}$$

The left matrix contains the variable coefficients; the right matrix (that's a *column matrix* or *column vector*, by the way) contains the right hand side values of each equation. These two matrices preserve all essential information about the linear system, except the variable names (but those are not important).

Review of Lecture 01: Augmented Matrix

An augmented matrix is a matrix in which rows or columns of another matrix of the appropriate order are appended to the original matrix (typically, to the right of the original matrix). If $\mathbf{A}$ is augmented on the right with $\mathbf{B}$, the resultant matrix is denoted as $(\mathbf{A}|\mathbf{B})$ or $[\mathbf{A}|\mathbf{B}]$. For example, if

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 5 & 6 \end{bmatrix} \ \mathbf{B} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \ \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

then $(\mathbf{A}|\mathbf{B}) = \begin{bmatrix} 1 & 4 & | & 3 \\ 5 & 6 & | & 1 \end{bmatrix}$ and $(\mathbf{A}|\mathbf{I}) = \begin{bmatrix} 1 & 4 & | & 1 & 0 \\ 5 & 6 & | & 0 & 1 \end{bmatrix}$.

Review of Lecture 01: Matrices and Linear Systems

Let's consider the linear system we reviewed on the first slide:

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= 39 \\ 2x_1 + 3x_2 + x_3 &= 34 \\ x_1 + 2x_2 + 3x_3 &= 26 \end{aligned}$$

The coefficient matrix and the column vector of the right-hand side values can be combined into one augmented matrix.

$$\begin{bmatrix} 3 & 2 & 1 & | & 39 \\ 2 & 3 & 1 & | & 34 \\ 1 & 2 & 3 & | & 26 \end{bmatrix}$$ or, if we drop the vertical bars, which are convenient but unnecessary, $$\begin{bmatrix} 3 & 2 & 1 & 39 \\ 2 & 3 & 1 & 34 \\ 1 & 2 & 3 & 26 \end{bmatrix}$$

Review of Lecture 01: Elementary Row Operations

There are 3 *elementary row operations* **R1**, **R2**, and **R3** defined below that we briefly discussed in Lecture 01 on Jan. 8, 2024.

- **R1** (Row interchange): Interchange any two rows in a matrix.
- **R2** (Row scaling): Multiply any row in the matrix by a nonzero scalar (a real number).
- **R3** (Row addition): Replace any row in the matrix with the sum of that row and another row in the matrix.

This brings to the end our brief review of Lecture 01.

Row Equivalence

If $\mathbf{A}$ and $\mathbf{B}$ are two matrices such that $\mathbf{B}$ can be obtained from $\mathbf{A}$ by a finite chain of the elementary row operations (we don't have to use all three types, e.g., we can use only row interchages or row additions or a combination of the two or all three), in that case we say that $\mathbf{A}$ is *row equivalent* to $\mathbf{B}$, or formally,

$$\mathbf{A} \sim \mathbf{B}.$$

Row Echelon Form and Pivots

A matrix is in **row echelon form** if:

- All rows containing only zeros appear below the rows containing nonzero entries.
- The first nonzero entry in any row appears in a column to the right of the column of the first nonzero entry in any preceding row.

The first nonzero entry in a row of a row echelon form matrix is called a **pivot**.

A Fundamental Theorem of Linear Algebra

Before we proceed, let's review an important theorem of linear algebra, because our method of solving linear systems with matrices is based on it.

If $[\mathbf{A}|\mathbf{b}] \sim [\mathbf{H}|\mathbf{c}]$, then the corresponding linear systems $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{Hx} = \mathbf{c}$ have the same solution set.

Back Substitution: Solving $\mathbf{Hx} = \mathbf{c}$

Why do we need the row echelon form? Because it makes it easier to solve linear systems. Let's go though an example to see how easy it is to determine all solutions of the system if $\mathbf{H}$ is in row echelon form. Let's solve $\mathbf{Hx} = \mathbf{c}$, where

$$[\mathbf{H}|\mathbf{c}] = \begin{bmatrix} -5 & -1 & 3 & | & 3 \\ 0 & 3 & 5 & | & 8 \\ 0 & 0 & 2 & | & -4 \end{bmatrix}$$

Note that the coefficient matrix $\mathbf{H}$ is in the row echelon form. The equations corresponding to $[\mathbf{H}|\mathbf{c}]$ are

1. $-5x_1 - x_2 + 3x_3 = 3.$
2. $0x_1 + 3x_2 + 5x_3 = 8.$
3. $0x_1 + 0x_2 + 2x_3 = -4.$

We solve for $x_3 = -2$ in equation 3, substitute x_3 in equation 2 to solve for $x_2 = 6$, and then substitute x_2 and x_3 into equation 1 to solve for $x_1 = -3$. This method is called **back substitution**.

An Algorithm for Solving a Linear System

Given a linear system $\mathbf{Ax} = \mathbf{b}$, obtain $[\mathbf{A}|\mathbf{b}]$, row-reduce it to $[\mathbf{H}|\mathbf{c}]$, where $\mathbf{H}$ is in row echelon form and use back substitution to find a solution.

This method does not always return a unique solution (i.e., when $\mathbf{H}$ has a pivot in every column). For example, the original system $\mathbf{Ax} = \mathbf{b}$ may be inconsistent (i.e., may have no solutions, when $\mathbf{H}$ has a row of 0's but the corresponding entry in the $\mathbf{c}$ column is not 0) or may have infinitely many solutions (if some columns of $\mathbf{H}$ have no pivots).

Gauss-Jordan Reduction of $\mathbf{Ax} = \mathbf{b}$ to $\mathbf{Hx} = \mathbf{c}$

A common method of solving linear systems, which we'll study in this course, is known as *Gauss-Jordan Elimination* or *Gauss-Jordan Reduction*. Here is how it works.

1. Push all rows containing only 0's to the bottom of $\mathbf{A}$.
2. If the first column of $\mathbf{A}$ contains only 0's, disregard it and go on to the next column. Keep skipping zero columns until the current column has a non-zero entry or the matrix is exhausted.
3. Use the row exchange as necessary to obtain the pivot entry in the top row of the first non-zero column.
4. Use the scalar multiplication to set the value of the pivot entry in the top row of the first non-zero column to 1.
5. Use row addition and scalar multiplication to create 0's below the pivot in the pivot's column.
6. Go back to step 1 on the smaller matrix that starts from the next row and the next column.

Gauss-Jordan Elimination Formulated More Rigorously

1. Push the zero rows down to the bottom row of the matrix. If all rows are zero rows, then you're done.

2. Let the current row with non-zero entries be r. For example, initially, $r = 1$ (or $r = 0$ if indexing is zero-based). Move right skipping the zero columns. When you find a column c with a non-zero entry, use row swaps to move the entry in the position. (r, c) That's a **pivot**. You may want to scale the row (i.e., use the scalar multiplication) to set the pivot to 1.

3. Once you have a pivot at (r, c), use row addition and row scaling to zero out all entries below (r, c) in column c. You can also zero out all entries above the pivot but it's not necessary for the back substitution to work.

Gauss-Jordan Elimination Formulated More Rigorously

Steps 1 – 3 are on the previous slide.

4. Set the current row to $r+1$ and the current column to $c+1$ and go back to step 1 with the matrix whose top left corner is at $(r+1, c+1)$ (unless you've run out of rows or columns).

5. If you've run out of rows or columns, check if there's a pivot in every column. If yes, the system has a unique solution; if no, the system has inifinitely many solutions (because the columns without pivots correspond to **free** variables; if you have a row where all entries are 0's except for the last one, the system is inconsistent (i.e., it has no solutions).

Gauss-Jordan Reduction Example

Let's apply Gauss-Jordan Reduction to the following matrix.

$$\begin{bmatrix} 0 & 1 & -3 & -5 \\ 2 & 3 & -1 & 7 \\ 4 & 5 & -2 & 10 \end{bmatrix}.$$

1. We interchange row 1 and row 2 to obtain

$$\begin{bmatrix} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 4 & 5 & -2 & 10 \end{bmatrix}.$$

2. We add -2 $\times$ row 1 to row 3 (i.e., replace row 3 with the sum of -2 $\times$ row 1 and itself) to obtain

$$\begin{bmatrix} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & -1 & 0 & -4 \end{bmatrix}.$$

Gauss-Jordan Reduction Example

3. We add row 2 to row 3 (i.e., replace row 3 with the sum of row 2 and itself) to obtain

$$\begin{bmatrix} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & -3 & -9 \end{bmatrix}.$$

4. We multiply row 3 by $-\frac{1}{3}$ to obtain

$$\begin{bmatrix} 2 & 3 & -1 & 7 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

5. We add row 3 to row 1 to obtain

$$\begin{bmatrix} 2 & 3 & 0 & 10 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

We can use backsubmitution after step 5, because the above coefficient matrix os in row echelon form. However, let's keep going to reduce all pivots to 1 and turn all entries above and below each pivot in each column to 0. That'll make our life much easier in the end.

Gauss-Jordan Reduction Example

6. We add 3 $\times$ row 3 to row 2 to obtain

$$\begin{bmatrix} 2 & 3 & 0 & 10 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

7. We add -3 $\times$ row 2 to row 1 to obtain

$$\begin{bmatrix} 2 & 0 & 0 & -2 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

8. We scale row 1 by $\frac{1}{2}$ to obtain

$$\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

Now we can read the solution from the above matrix directly without any backsubstitution, i.e., the solution is $x_1 = -1$, $x_2 = 4$, $x_3 = 3$.

Determinants

If we consider each row in an n x n matrix as a vector in the n-dimensional space, then the determinant of the matrix, if one exists, is a real number whose absolute value can be interpreted as the volume of the hyper box in the n dimension space.

Thus, in 1D, the determinant of a matrix with exactly 1 point, i.e., $[a_{11}]$, is the value of that point, i.e., a_{11}. In 2D, the determinant of a 2 x 2 matrix is the area of the parallelogram determined by the two row vectors. In 3D, the determinant of a 3 x 3 matrix is the volume of the box determined by the three row vectors, etc.

First- and Second-Order Determinants

The determinant of a 1×1 matrix is its single value. This determinant is called the **first-order determinant**. Let us define a generic 2×2 matrix

$$A = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix},$$

where a_1, a_2, b_1, b_2 are real numbers. The following equation defines the **second-order determinant**.

$$det(A) = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1.$$

The absolute value of the second-order determinant is the area of the parallelogram defined by the vectors (a_1, a_2) and (b_1, b_2).

Third-Order Determinants

Let us define the determinant of a 3×3 matrix, i.e., the **third-order determinant**. Here is a generic 3×3 matrix

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}.$$

The third-order determinant is computed as follows:

$$det(A) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \cdot \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \cdot \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \cdot \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.$$

The absolute value of the third-order determinate is the volume of the box in 3D defined by the vectors (a_1, a_2, a_3), (b_1, b_2, b_3), and (c_1, c_2, c_3).

Minor Matrices

In order to state a generic formula for computing determinants of an n x n matrix, $n > 0$, we need to define the concept of a minor matrix. Once we do that, then we will see how to compute the second-order determinant in terms of the first-order determinants and the third-order determinant in terms of the second-order determinants, the fourth-order determinant in terms of the third-order determinants, the fifth-order in terms of the fourth-order ones, etc.

Let's define the **minor matrix** A_{ij} of an $n \times n$ matrix A is an $(n-1) \times (n-1)$ matrix obtained from A by removing row i and column j. For example, let

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

The minor matrix A_{11} is obtained from A by removing row 1 and column 1 from A; the minor matrix A_{12} is obtained from A by removing row 1 and column 2 from A, etc.

Cofactors

If A is an $n \times n$ matrix, then $det(A)$ can be written as $|A|$. If A_{ij} is a minor of A, then $det(A_{ij}) = |A_{ij}|$. If A is a 3×3 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

then

$$det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}|A_{11}| - a_{12}|A_{12}| + a_{13}|A_{13}|.$$

The numbers $|A_{11}|$, $-|A_{12}|$, and $|A_{13}|$ are called **cofactors**.

Leibnitz's Determinant Formula

The general formula for the cofactor c_{ij} of the entry a_{ij} in an $n \times n$ matrix A is defined as

$$c_{ij} = (-1)^{i+j} det(A_{ij})$$

Let A_{ij} be the minor of A. Then the determinant of an $n \times n$ matrix A is

$$det(A) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{11}c_{11} + a_{12}c_{12} + \dots + a_{1n}c_{1n}.$$

General Expansion by Minors

You can use the Leibnitz formula along any row or column of a matrix of an $n \times n$ matrix A. The value of the determinant will remain the same. This is known as the **general expansion by minors**.

More Efficient Determinant Computation

1. Reduce A to row echelon form with Gauss elimination.
2. Keep track of the number of row interchanges (this is the only elementary row operation that affects the value of the determinant).
3. After row reduction is done, $det(A) = (-1)^n \times$(product of pivots), where n is the number of row interchanges.

Special Matrices

The following few slides discuss special kinds of matrices (diagonal, identity, inverse, and transpose). Make sure that you understand what these matrices are.

Diagonal Matrix

A **diagonal** matrix is a square matrix (i.e., the number of rows is equal to the number of columns) whose off-diagonal elements (i.e., $a_{ij}, i \neq j$) are equal to 0.

$$\mathbf{A} = [a_{ij}] = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

Identity Matrix

An **identity** matrix is a diagonal matrix where all diagonal elements (main diagonal's elements) are equal to 1 and the other elements are equal to 0. It is typically denoted as $\mathbf{I}_m$.

$$\mathbf{I_2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{I_4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Invertible Matrix

An $n \times n$ matrix $\mathbf{A}$ is invertible if there exists an $n \times n$ matrix $\mathbf{C}$ such that $\mathbf{AC} = \mathbf{CA} = \mathbf{I}$.

If $\mathbf{A}$ is not invertible, it is **singular**.

The matrix $\mathbf{C}$ is called the inverse of $\mathbf{A}$ and is unique in that if $\mathbf{AC} = \mathbf{AD} = \mathbf{I}$, then $\mathbf{C} = \mathbf{D}$.

If $\mathbf{A}$ is invertible, its inverse matrix is denoted as $\mathbf{A}^{-1}$.

Computation of $\mathbf{A}^{-1}$

To find $\mathbf{A}^{-1}$, if it exists, proceed as follows:

1. Form the augmented matrix $[\mathbf{A}|\mathbf{I}]$.
2. Use the Gauss-Jordan reduction to attempt to reduce $[\mathbf{A}|\mathbf{I}]$ to $[\mathbf{I}|\mathbf{C}]$. If the reduction can be carried out, then $\mathbf{C} = \mathbf{A}^{-1}$. Otherwise, $\mathbf{A}$ is singular.

Matrix Transpose

The **transpose** of a matrix $\mathbf{A}$, denoted as $\mathbf{A}^T$ is a reordering of $\mathbf{A}$ where the rows are interchanged with columns, in order. Row 1 of $\mathbf{A}$ becomes column 1 of $\mathbf{A}^T$, row 2 of $\mathbf{A}$ becomes column 2 of $\mathbf{A}^T$, etc. In other words,

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix} \qquad \mathbf{A}^T = \begin{bmatrix} a_{1,1} & a_{2,1} & \dots & a_{m,1} \\ a_{1,2} & a_{2,2} & \dots & a_{m,2} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ a_{1,n} & a_{2,n} & \dots & a_{m,n} \end{bmatrix}$$

Cramer's Rule

Gabriel Cramer, a Swiss mathematician, discovered a method to compute the solution to the linear system $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is an invertible $n \times n$ matrix, $\mathbf{x}$ is an $n \times 1$ matrix (i.e., a vector) of variables, and $\mathbf{b}$ is an $n \times 1$ vector of the right-hand sides of the linear equations whose coefficients are in $\mathbf{A}$.

Cramer's rule states that the solution $\mathbf{x}$ of $\mathbf{Ax} = \mathbf{b}$ is a vector $(x_1, x_2, ..., x_n)$, where $x_k = det(\mathbf{B_k})/det(\mathbf{A})$, $i \leq k \leq n$, and $\mathbf{B_k}$ is a matrix obtained from $\mathbf{A}$ by replacing column k in $\mathbf{A}$ with $\mathbf{b}$.

Example

Let's apply Cramer's rule to solve $\mathbf{Ax} = \mathbf{b}$ such that

$$\begin{bmatrix} 5 & -2 & 1 \\ 3 & 2 & 0 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}.$$

We have

$$\mathbf{B_1} = \begin{bmatrix} 1 & -2 & 1 \\ 3 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix}; \mathbf{B_2} = \begin{bmatrix} 5 & 1 & 1 \\ 3 & 3 & 0 \\ 1 & 0 & -1 \end{bmatrix}; \mathbf{B_3} = \begin{bmatrix} 5 & -2 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 0 \end{bmatrix}.$$

Now we compute $det(\mathbf{A}) = -15$; $det(\mathbf{B_1}) = -5$; $det(\mathbf{B_2}) = -15$; $det(\mathbf{B_3}) = -20$. Then $x_1 = det(\mathbf{B_1})/det(\mathbf{A}) = \frac{1}{3}$; $x_2 = det(\mathbf{B_2})/det(\mathbf{A}) = 1$; $x_3 = det(\mathbf{B_3})/det(\mathbf{A}) = \frac{4}{3}$.

Numpy Tip: How to Replace Columns in Numpy Matrices

Here's how we can replace columns in 2D numpy matrices. We create the matrix of the Cramer's rule example and replace each column vector with **b**.

```
A = np.array([[5, -2, 1],
              [3, 2, 0],
              [1, 1, -1]])
b = np.array([1, 3, 0])

B1 = A.copy()
B1[:,0] = b ### column 0 is now b
B2 = A.copy()
B2[:,1] = b ### column 1 is now b
B3 = A.copy()
B3[:,2] = b ### column 2 is now b
```

References

1. The file np_col_replacement_ops.py contains examples of how to replace colums with vectors.
2. J. Fraleigh, R. Beauregard. *Linear Algebra*, Ch. 01.