Chapter 2: Solving Linear Equations $A\vec{x} = \vec{b}$

Sarah Helmbrecht

1 Systems of Equations

Definition 1.1. *A **linear system** in the variables $x_1, \ldots, x_n$ is a list of equations of the form*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

*where $a_1, \ldots, a_n, b$ are constants. An assignment of numbers to the variables $x_1, \ldots, x_n$ is a **solution** if the assignment satisfies each of the equations. The **solution set** is the collection of all solutions. **Solving** the system means finding the solution set.*

Proposition 1.2. *A linear system of the form*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

can be expressed in matrix notation as $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \ \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \ \text{and } \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

2 Elimination

Definition 2.1. *We call the following operations **elementary row operations (EROs)**:*

1. *Multiply all entries in a row by a nonzero number.*
2. *Add a scalar multiple of one row to another row.*
3. *Swap two rows.*

Theorem 2.2. *EROs preserve the set of solutions to a linear system.*

Definition 2.3. *Two matrices are called **row equivalent** if one can be obtained from the other through EROs.*

Definition 2.4. *A system of equations is called **inconsistent** if it has no solution. It is **consistent** otherwise.*

Algorithm 2.5. ***(Solving Linear Systems)*** *Suppose we are given a system of m equations in n unknowns:*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}.$$

This system can be written in matrix form as:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

*This system can be written in **augmented form** as:*

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

Next, we can use elimination (via EROs) to solve the system.

3 Existence and Uniqueness of Solutions

Theorem 3.1. *If $A\vec{x} = \vec{b}$ is a system of n equations for n unknowns, then $A\vec{x} = \vec{b}$ can have exactly one solution $\vec{x}$, no solutions, or infinitely many solutions.*

1. *There is exactly one solution when all the columns of A are independent. In this case, the only solution to $A\vec{x} = \vec{0}$ is $\vec{x} = \vec{0}$, and A has an inverse matrix A^{-1}.*
2. *There is no solution (inconsistent) when $\vec{B}$ is not a linear combination of the columns of A. In other words, $\vec{b}$ is not in the column space of A.*
3. *There are infinitely many solutions to $A\vec{x} = \vec{0}$ when the columns of A are not all independent.*

4 Matrix Operations

Definition 4.1. ***(Matrix Addition)*** *If A and B are $n \times m$ matrices, then*

$$A + B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}.$$

Definition 4.2. ***(Scalar Multiplication)*** *If A is an $n \times m$ matrix and c is a scalar, then*

$$cA = c\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}.$$

Definition 4.3. ***(Matrix Multiplication)*** *If A is an $n \times m$ matrix and B is an $m \times p$ matrix, then the ijth entry of AB is the ith row of A times the jth column of B:*

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

Proposition 4.4. *Suppose A is $m \times n$, and B and C are of appropriate sizes to make matrix multiplication possible. Then:*

1. $A(BC) = (AB)C$.
2. $(B + C)A = BA + CA$.
3. $c(AB) = A(cB)$.
4. $AI_n = A$.
5. $A(B + C) = AB + AC$.
6. $c(AB) = (cA)B$.
7. $I_mA = A$.

Note that matrix multiplication is not commutative.

5 Elimination and Permutation Matrices

Definition 5.1. *An **elimination matrix** E_{ij} adds a multiple l_{ij} of equation j to equation i for any matrix it is multiplied by.*

Algorithm 5.2. ***(Elimination Matrices)*** *To construct an elimination matrix E_{ij} that adds a multiple l_{ij} of equation j to equation i, take the identity matrix and replace the zero in position ij with l_{ij}.*

Definition 5.3. *A **permutation matrix** P_{ij} swaps rows i and j for any matrix it is multiplied by.*

Algorithm 5.4. ***(Permutation Matrices)*** *To construct a permutation matrix P_{ij} that swaps rows i and j, take the identity matrix and swap rows i and j.*

Proposition 5.5. ***(Properties of Permutation Matrices)***

1. *A permutation matrix P has a one in every row and a one in every column, and all other entries are zero.*
2. *Let P be an $n \times n$ permutation matrix. Then the n ones appear in n different rows and n different columns of P.*
3. *The product of two permutation matrices is a permutation matrix.*
4. *The inverse of a permutation matrix is also a permutation matrix.*
5. *If A is an invertible $n \times n$ matrix, there is a permutation matrix P to order its rows in advance so that elimination on PA results in no zeros in the pivot positions.*

6 Inverse Matrices

Definition 6.1. *Suppose A is an $n \times n$ matrix. Then A is **invertible** if there exists an **inverse** matrix A^{-1} such that*

$$AA^{-1} = A^{-1}A = I.$$

Proposition 6.2. ***(Properties of Inverses)*** *Let A be an $n \times n$ matrix.*

1. *The inverse exists if and only if elimination produces n pivots (allowing row exchanges). Elimination solves $A\vec{x} = \vec{b}$ without explicitly using A^{-1}.*
2. *The inverse of a matrix A is unique. If $BA = I$ and $AC = I$, then by the associative law, $B(AC) = (BA)C \Rightarrow BI = IC \Rightarrow B = C$.*
3. *If A is invertible, then the one and only solution to $A\vec{x} = \vec{b}$ is $\vec{x} = A^{-1}\vec{b}$. To see this, take $A\vec{x} = \vec{b} \Rightarrow A^{-1}A\vec{x} = A^{-1}\vec{b} \Rightarrow I\vec{x} = A^{-1}\vec{b} \Rightarrow \vec{x} = A^{-1}\vec{b}$.*
4. *Suppose there is a nonzero vector $\vec{x}$ such that $A\vec{x} = \vec{0}$. Then A has dependent columns, so A is not invertible.*
5. *If A is invertible, then $A\vec{x} = \vec{0}$ has only the zero solution $\vec{x} = A^{-1}\vec{0} = \vec{0}$.*
6. *A square matrix is invertible if and only if its columns are independent.*
7. *A 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if the number $ad - bc \neq 0$. In this case, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. The number $ad - bc$ is the **determinant** of the matrix. A matrix A is invertible if $\det(A) \neq 0$.*
8. *An upper triangular matrix has an inverse provided no diagonal entries d_i are zero. If $A = \begin{bmatrix} d_1 & * & \cdots & * \\ 0 & d_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$, then*

$$A^{-1} = \begin{bmatrix} \frac{1}{d_1} & * & \cdots & * \\ 0 & \frac{1}{d_2} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & \frac{1}{d_n} \end{bmatrix}.$$

Algorithm 6.3. ***(Inverses)*** *To find the inverse A^{-1} of an invertible $n \times n$ matrix A, first augment A with the $n \times n$ identity matrix to obtain $\begin{bmatrix} A \mid I \end{bmatrix}$. Next, use elimination (via EROs) until the left-hand side is the identity matrix. Then we have $\begin{bmatrix} I \mid A^{-1} \end{bmatrix}$.*

Theorem 6.4. *If A and B are both invertible $n \times n$ matrices, then the inverse of AB is $(AB)^{-1} = B^{-1}A^{-1}$.*

Proposition 6.5. *Even if A and B are both invertible $n \times n$ matrices, we cannot guarantee that the matrix $(A + B)$ is invertible.*

Algorithm 6.6. ***(Inverses of Elimination Matrices)*** *To find the inverse E_{ij}^{-1} of an elimination matrix E_{ij} that has l_{ij} in position ij, we simply replace the entry l_{ij} with $-l_{ij}$.*

Theorem 6.7. ***(Inverses of Permutation Matrices)*** *The inverse of a permutation matrix is its transpose.*

7 The Transpose of a Matrix

Definition 7.1. *The* ***transpose*** *of an $m \times n$ matrix A is the $n \times m$ matrix A^T whose rows are the columns of A. In other words, the ij-th entry of A^T is a_{ji}.*

Proposition 7.2. ***(Properties of the Transpose)***

1. $(A + B)^T = A^T + B^T$.
2. $(AB)^T = B^T A^T$.
3. $(A^T)^{-1} = (A^{-1})^T$.

Proposition 7.3. *Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. We can write their dot product as $\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y}$.*

8 LU Factorization

Definition 8.1. *The* ***LU factorization*** *of an $n \times n$ matrix A is its decomposition into the product $A = LU$, where L is an $n \times n$ lower triangular matrix and U is an $n \times n$ upper triangular matrix.*

Algorithm 8.2. ***(LU Factorization)*** *To find the LU factorization of an $n \times n$ matrix A,*

1. *Use elimination (via EROs) on A to obtain an upper triangular matrix U. If U has nonzero pivots, continue. Otherwise, see Algorithm 8.7 to find $PA = LU$ instead.*
2. *Take note of the elimination matrices that performed each step of the elimination process.*
3. *Find the inverse of each elimination matrix.*
4. *Let L equal the product of the inverses of the elimination matrices.*
5. *Finally, $A = LU$.*

Algorithm 8.3. ***(Solving Systems using LU Factorization)*** *To solve the system $A\vec{x} = \vec{b}$ given LU factorization $A = LU$,*

1. *Solve the lower triangular system $L\vec{y} = \vec{b}$ for $\vec{y}$ using forward substitution.*
2. *Solve the upper triangular system $U\vec{x} = \vec{y}$ for $\vec{x}$ using back substitution.*

Definition 8.4. *The* ***LDU factorization*** *of an $n \times n$ matrix A is its decomposition into the product $A = LDU$, where L is an $n \times n$ lower triangular matrix, D is an $n \times n$ diagonal matrix, and U is an $n \times n$ upper triangular matrix.*

Algorithm 8.5. ***(LDU Factorization)*** *To find the LDU factorization of an $n \times n$ matrix A,*

1. *Find the LU factorization $A = LU$ using Algorithm 8.2.*
2. *Let D be the diagonal matrix that has the diagonal entries of U on its diagonal, and zeros everywhere else.*
3. *Divide each row of U by its diagonal entry to obtain an upper triangular matrix U'.*
4. *Finally, $A = LDU'$.*

Proposition 8.6. *If Algorithm 8.2 fails to produce an upper triangular matrix U that has nonzero pivots, we cannot find $A = LU$, but we can find $PA = LU$.*

Algorithm 8.7. ***(PA=LU Factorization)***

1. *Apply a permutation matrix P to A such that elimination on A results in nonzero pivots.*

2. *Use elimination (via EROs) on PA to obtain an upper triangular matrix U that has nonzero pivots.*
3. *Take note of the elimination matrices that performed each step of the elimination process.*
4. *Find the inverse of each elimination matrix.*
5. *Let L equal the product of the inverses of the elimination matrices.*
6. *Finally, $PA = LU$.*

9 Symmetric Matrices

Definition 9.1. *A matrix S is* ***symmetric*** *if $S^T = S$.*

Proposition 9.2. ***(Properties of Symmetric Matrices)***

1. *If S is a symmetric matrix, then S^{-1} is also symmetric.*
2. *If A is an $n \times m$ matrix, then the matrices AA^T and A^TA are symmetric.*
3. *If A is a symmetric $n \times n$ matrix, then the LDU factorization of A is $A = LDL^T$.*