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

2.1 Elimination and Back Substitution

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.

- 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}.
- No solution when $\vec{b}$ is not a LC of the columns of A. In other words, $\vec{b}$ is not in the column space of A.
- Infinitely many solutions to $A\vec{x}=\vec{0}$ when the columns of A are not all independent.

Elimination is a procedure to simplify A into a matrix U without changing any solution $\vec{x}$ to $A\vec{x}=b$. We do the same (reversible) operations to both sides of the equation. Elimination keeps all solutions $\vec{x}$ and creates no new ones. U is an upper triangular matrix.

Since U is upper triangular, we can use back substitution to solve $U\vec{x}=\vec{c}$, which gives us the solutions to $A\vec{x}=\vec{b}$.

An $n\times n$ matrix A has independent columns IFF U has n nonzero pivots (after possible row exchanges).

Every square matrix A with independent columns (full rank) can be reduced to a triangular matrix U with nonzero pivots.

U is the product of elimination matrices and A. The elimination matrices are whatever it takes to get A into upper triangular form U.

If zero appears in a pivot position, try multiplying by a permutation matrix P in order to exchange rows.

If U doesn't have a pivot in every column, then A doesn't have full rank, so A has dependent columns. Therefore $A\vec{x}=\vec{0}$ has infinitely many solutions.

A triangular matrix U has full rank exactly when its main diagonal has no zeros.

Summary: an elimination matrix E will act on $A\vec{x}=\vec{b}$. If zero appears in a pivot position, use a permutation matrix P. This gives us an upper triangular matrix U and a new right hand side c. Then $U\vec{x}=\vec{c}$ is solved by back substitution.

To make sure that the operations of E and P on the matrix A are also executed on $\vec{b}$, we can apply E and P to the augmented matrix $[A \mid \vec{b}]$.

The overall equation for elimination on A is $PA=LU$, where $L=E^{-1}$.

There are $n!$ permutation matrices P that permute the rows of $n\times n$ matrices, including $P=I$ for no row exchanges.

2.2 Elimination Matrices and Inverse Matrices

The basic elimination step subtracts a multiple ℓ_{ij} of equation j from equation i.

The matrix A is invertible if $\exists A^{-1}$ s.t. $A^{-1}A=AA^{-1}=I$.

Properties of Inverses:

- The inverse exists IFF elimination produces n pivots (allowing row exchanges). Elimination solves $A\vec{x}=\vec{b}$ without explicitly using A^{-1}.
- The inverse of a matrix A is unique. If $BA=I$ and $AC=I$, then by the associative law, $B(AC)=(BA)C \Rightarrow BI=IC \Rightarrow B=C$.
- 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 $A\vec{x}=\vec{b} \Rightarrow A^{-1}A\vec{x}=A^{-1}\vec{b} \Rightarrow I\vec{x}=A^{-1}\vec{b} \Rightarrow \vec{x}=A^{-1}\vec{b}$.
- Suppose there is a nonzero vector $\vec{x}$ s.t. $A\vec{x}=\vec{0}$. Then A has dependent columns, so A cannot have an inverse.
- If A is invertible, then $A\vec{x}=\vec{0}$ has only the zero solution $\vec{x}=A^{-1}\vec{0}=\vec{0}$.
- A square matrix is invertible IFF its columns are independent.
- A 2×2 matrix is invertible IFF the number $ad-bc\neq 0$: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}=\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. The number $ad-bc$ is the determinant of A. A matrix is invertible if $\det(A)\neq 0$.
- A triangular matrix has an inverse provided no diagonal entries d_i are zero:

  If $A=\begin{bmatrix} d_1 & * & * & * \\ 0 & \ddots & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & d_n \end{bmatrix}$, then $A^{-1}=\begin{bmatrix} 1/d_1 & * & * & * \\ 0 & \ddots & * & * \\ 0 & 0 & \ddots & * \\ 0 & 0 & 0 & 1/d_n \end{bmatrix}$.

If A and B (same size) are invertible, then the inverse of AB is $(AB)^{-1}=B^{-1}A^{-1}$.

For square matrices, an inverse on one side (i.e., $AB=I$) is automatically an inverse on the other side (i.e., $BA=I$).

E is the product of all the elimination matrices E_{ij}, taking A into its upper triangular form $EA=U$. Assume for now that no row exchanges are involved (i.e., $P=I$). Multiplying all the separate elimination steps to get E is messy. Let's look at the $n=3$ case:

$$E=E_{32}E_{31}E_{21}=\begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & -\ell_{32} & 1 \end{bmatrix}\begin{bmatrix} 1 & & \\ 0 & 1 & \\ -\ell_{31} & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & & \\ -\ell_{21} & 1 & \\ 0 & 0 & 1 \end{bmatrix}=\begin{bmatrix} 1 & & \\ -\ell_{21} & 1 & \\ (\ell_{32}\ell_{21}-\ell_{31}) & -\ell_{32} & 1 \end{bmatrix}.$$

$$\text{But } E^{-1}=\begin{bmatrix} 1 & & \\ \ell_{21} & 1 & \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & & \\ 0 & 1 & \\ \ell_{31} & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & \ell_{32} & 1 \end{bmatrix}=\begin{bmatrix} 1 & & \\ \ell_{21} & 1 & \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix}=L, \text{ which is much easier.}$$

2.3 Matrix Computations and $A=LU$

To find the inverse of an $n\times n$ matrix A, we augment it by I and use Gauss-Jordan elimination to go from $[A \mid I]$ to $[I \mid A^{-1}]$.

Reducing A to U requires about $\frac{1}{3}n^3$ multiplications and $\frac{1}{3}n^3$ subtractions.

Going from $\vec{b}$ to $\vec{c}$ requires n^2 multiplications and n^2 subtractions.

Key reason why $A=LU$: the pivot rows that are subtracted from lower rows aren't always the original rows of A because elimination probably changed them. But these pivot rows are rows of U, because pivot rows never change again. In the 3×3 case, row 3 of U is (Row 3 of A) $-\ell_{31}$(Row 1 of U) $-\ell_{32}$(Row 2 of U). Rewriting this, we can see that the row $[\ell_{31}\ \ell_{32}\ 1]$ is multiplying the matrix U: (Row 3 of A) $=\ell_{31}$(Row 1 of U) $+\ell_{32}$(Row 2 of U) $+$ (Row 3 of U). This is exactly row 3 of $A=LU$. That row of L holds $\ell_{31}, \ell_{32}, 1$. All rows look like this, whatever the size of A. With no row exchanges, we have $A=LU$.

See textbook page 60 for a second proof that $A=LU$.

$A=LU$ is possible with no row exchanges ($P=I$) and no zeros in the pivots if for $k=1,\dots,n$, all upper left submatrices of A are invertible.

2.4 Permutations and Transposes

Permutation matrices P swap the rows of matrices. Permutation matrices have a 1 in every row and a 1 in every column, and all other entries are zero.

Half of the $n!$ permutations of size n are **even**, and half are **odd**. An even permutation needs an even # of simple row exchanges to reach I.

Properties of Permutation Matrices:

- The n ones appear in n different rows and n different columns of P.
- The columns of P are orthogonal: dot products between columns are all zero.
- The product P_1P_2 of permutations is a permutation. The inverse of a permutation is also a permutation.
- If A is invertible, there is a permutation P to order its rows in advance, so that elimination on PA meets no zeros in the pivot positions. Then $PA = LU$.

Elimination can often succeed, even when a zero appears in the pivot position. We just multiply A by an appropriate permutation matrix P to swap the rows.

If $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ and $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, then $PA = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \end{bmatrix}$ only has rows swapped. But column permutation $Q = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ gives $PAQ = \begin{bmatrix} a_{23} & a_{22} & a_{21} \\ a_{33} & a_{32} & a_{31} \\ a_{13} & a_{12} & a_{11} \end{bmatrix}$, with columns swapped.

The **transpose** A^T is the matrix whose columns are the rows of A, $(A^T)_{ij} = A_{ji}$.

Properties of the Transpose:

- The transpose of $A+B$ is $(A+B)^T = A^T + B^T$.
- The transpose of AB is $(AB)^T = B^TA^T$.
- The transpose of A^{-1} is $(A^{-1})^T = (A^T)^{-1}$.

$A\vec{x}$ combines the columns of A, while $\vec{x}^TA^T$ combines the rows of A^T.

Transpose of inverse: $A^{-1}A = I$ is transposed to $A^T(A^{-1})^T = I$.

The **dot product/inner product** of $\vec{x}$ and $\vec{y}$ is $\vec{x}^T\vec{y}$, which is 1×1 and therefore a scalar.

The **rank one product/outer product** of $\vec{x}$ and $\vec{y}$ is $\vec{x}\vec{y}^T$, which is an $n\times n$ matrix.

A^T is the matrix such that $(A\vec{x})^T\vec{y} = \vec{x}^T(A^T\vec{y})$.

A matrix S is **symmetric** if $S^T = S$. So every $s_{ji} = s_{ij}$.

The inverse of a symmetric matrix is a symmetric matrix. $(S^{-1})^T = (S^T)^{-1} = S^{-1}$.

For any matrix A, the product $S = A^TA$ is a square symmetric matrix: the transpose of A^TA is $A^T(A^T)^T = A^TA$. The matrix AA^T is also symmetric, but $AA^T \neq A^TA$.

Symmetric matrices make elimination twice as fast. The symmetry is in the triple product $S = LDL^T$. The diagonal matrix D of pivots can be divided out, to leave $U = L^T$.

For a rectangular A, the **saddle-point matrix** $S = \begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} = S^T$ is symmetric. It has block factorization $S = LDL^T = \begin{bmatrix} I & 0 \\ A^T & I \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & -A^TA \end{bmatrix}\begin{bmatrix} I & A \\ 0 & I \end{bmatrix}$.

S is invertible $\Leftrightarrow$ A^TA is invertible $\Leftrightarrow$ $A\vec{x} \neq \vec{0}$ whenever $\vec{x} \neq 0$.