16 Gaussian Elimination

Last week we studied the finite difference method for finding approximate solutions to partial differential equations, which led to the construction of linear systems of equations that must be solved. For the next few weeks, we'll study how to do that. We begin with an overview of the most classic method, but our main focus will be on methods for special matrices.

16.1 Gaussian Elimination

Suppose we want to solve $A\mathbf{x} = \mathbf{b}$, where $A \in \mathbb{R}^{m\times m}$ is nonsingular and $\mathbf{x}, \mathbf{b} \in \mathbb{R}^m$. The classical method that we all learned at school is called *Gaussian Elimination*: preform a series of elementary row operations to reduce the system to a lower triangular one, which is then easily solve with back-substitution.

Suppose the matrix is

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

The first step is to eliminate the term a_{21}. This is done by replacing A with

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} + \mu_{21}a_{12} & a_{23} + \mu_{21}a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = A + \mu_{21}\begin{pmatrix} 0 & 0 & 0 \\ a_{11} & a_{12} & a_{13} \\ 0 & 0 & 0 \end{pmatrix}$$

where $\mu_{21} = -a_{21}/a_{11}$. Because

$$\begin{pmatrix} 0 & 0 & 0 \\ a_{11} & a_{12} & a_{13} \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}, \end{pmatrix}$$

the row operation can be written as $\big(I + \mu_{21}E^{(21)}\big)A$, where $E^{(pq)} \in \mathbb{R}^{m\times m}$ has all zeros, except for $e_{pq} = 1$.

Next we eliminate the term a_{31} and then a_{32} so that, when we are done, A has been reduced to an *upper triangular matrix.*

In general each of the row operations in Gaussian Elimination can be written as

$$\big(I + \mu_{pq}E^{(pq)}\big)A \quad \text{where } 1 \leqslant q < p \leqslant m, \qquad (16.1)$$

and $\big(I + \mu_{pq}E^{(pq)}\big)$ is an example of a *Unit Lower Triangular Matrix.*

16.2 Triangular Matrices

Recall from Section 2.4 that $L \in \mathbb{R}^{m\times m}$ is a *lower triangular (LT) matrix* if the only non-zero entries are on or below the main diagonal, i.e., if $l_{ij} = 0$ for $1 \leqslant i \leqslant j \leqslant m$. It is a *unit lower triangular (ULT) Matrix* is $l_{ii} = 1$.

The analogous definitions of Upper Traiangular and Unit Upper Triangular matrices is obvious.

In Example 2.3 we saw that the product of two lower (upper) triangular matrices is lower (upper) triangular. Most importantly for us, as seen in Q4 of Problem Set 1:

(a) the product of two ULT matrices is a ULT matrix.

(b) the inverse of a ULT matrix always exists, and is a ULT matrix.

16.3 Factorising A

Each elementary row operation in Gaussian Elimination (GE) involves replacing A with $(I + \mu_{rs}E^{(rs)})A$. But $(I + \mu_{rs}E^{(rs)})$ is a unit lower triangular matrix. Also, when we are finished we have an upper triangular matrix. So we can write the whole process as

$$L_kL_{k-1}L_{k-2}\dots L_2L_1A = U, \qquad (16.2)$$

where each of the L_i is a ULT matrix. Since we know that the product of ULT matrices is itself a ULT matrix, we can write the whole process as

$$\tilde{L}A = U.$$

We also know that the inverse of a ULT matrix exists and is a ULT matrix. So we can write

$$A = LU,$$

where L is unit lower triangular and U is upper triangular. This is called the LU-*factorization* of the matrix.

16.4 Exercises

Exercise 16.1. Suppose that $B \in \mathbb{R}^{m\times m}$ is a matrix of the form

$$B = I + \mathbf{v}\mathbf{e}_j^\star.$$

where $\mathbf{v} \in \mathbb{R}^m$ with $v_j = 0$. Show that $B^{-1} = I - \mathbf{v}\mathbf{e}_j^\star$.

Exercise 16.2. We'll now use the exercise above to show that the inverses of the matrices $L_1, L_2, \dots, L_k$ on the left of (16.2) have a convenient form. If we let $l_{i,j} = x_{ij}/x_{jj}$, then we can write

$$L_j = I + \mathbf{l}_j\mathbf{e}_j,$$

where $\mathbf{l}_j^\star = (0, 0, \dots, 0, -l_{j+1,j}, -l_{j+2,j}, \dots, -l_{m,j})$.

Show that if $A = LU$, then

$$L = \prod_{j=1}^{k} L_j^{-1},$$

can be expressed as

$$L = I + \sum_{j=1}^{k} \mathbf{l}_j\mathbf{e}_j^\star.$$