Gaussian Elimination

Angjoo Kanazawa

October 27, 2011

1 Gaussian Elimination and LU Factorization

Gaussian Elimination is a way of solving systems of linear equations. When you apply the algorithm in a matrix form, it lets us write a matrix as the product of a unit lower-triangular matrix and an upper-triangular matrix.
Problem Statement: Solve $Ax = b$, where A is a N by M non-singular matrix.
Goal: Compute L, a unit lower-triangular matrix and U an upper-triangular matrix (with a P permutation matrix) s.t. $A = LU$. Then, we can solve $Ax = b$ by

$$\begin{aligned}
Ax &= b \\
LUx &= b \\
L\hat{b} &= b \text{ where } \hat{b} = L^{-1}b \\
&\Rightarrow Ux = \hat{b}
\end{aligned}$$

Where since L is a lower triangular matrix, we can solve for $\hat{b}$ using forward-substitution in $\mathcal{O}(n^2)$, and similarly, U is an upper triangular matrix so using backward-substitution we can solve for x in $\mathcal{O}(n^2)$
Doing the LU factorization costs $\mathcal{O}(\frac{2}{3}n^3) = \mathcal{O}(n^3)$.

2 How to do LU Factorization

1. Add multiple of first row of A and first entry of b to the second to last rows to produce zeros on the first column. i.e. Multpily A with L_1 to get $A^{(1)}$ ($A^{(i)}$ is our A after ith many steps).

2. Repeat, then we will get our upper triangular matrix $U = L_{n-1} \cdots L_1 A$, and our $L = L_1^{-1} \cdots L_{n-1}^{-1}$.

   (Why?) because $A = L_1^{-1} L_1 A = L_1^{-1} U_1$ where $U_1 = L_1 A$ we get:

$$\begin{aligned}
A &= (L_1^{-1} \cdots L_{n-1}^{-1})(L_{n-1} \cdots L_1 A) \\
&= LU
\end{aligned}$$

3 Step-by-Step example with 3 by 3 case

Let's walk through the entire thing with a 3 by 3 A:

3.1 LU Factorization

$$Ax = b \Rightarrow \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Where x is unknown.

1. Formally, our L_1 is $\left(\begin{smallmatrix} 1 & 0 & 0 \\ -a_{21}/a_{11} & 1 & 0 \\ -a_{31}/a_{11} & 0 & 1 \end{smallmatrix}\right)$, so

$$\begin{aligned} L_1 A &= \begin{pmatrix} 1 & 0 & 0 \\ -a_{21}/a_{11} & 1 & 0 \\ -a_{31}/a_{11} & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ \frac{-a_{21}}{a_{11}} a_{11} + a_{21} & \frac{-a_{21}}{a_{11}} a_{12} + a_{22} & \frac{-a_{21}}{a_{11}} a_{13} + a_{23} \\ \frac{-a_{31}}{a_{11}} a_{11} + a_{31} & \frac{-a_{31}}{a_{11}} a_{12} + a_{32} & \frac{-a_{31}}{a_{11}} a_{13} + a_{33} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} \end{pmatrix} \end{aligned}$$

2. $L_2 L_1 A = L_2 U^{(1)} =$

$$\begin{aligned} &\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -a_{32}^{(1)}/a_{22}^{(1)} & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & \frac{-a_{32}^{(1)}}{a_{22}^{(1)}} a_{22}^{(1)} + a_{32}^{(1)} & \frac{-a_{32}^{(1)}}{a_{22}^{(1)}} a_{23}^{(1)} + a_{33}^{(1)} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & 0 & a_{33}^{(2)} \end{pmatrix} = U \end{aligned}$$

Now What's L? Note that for all of L_i's we use, $L_i^{-1} = -L_i$ (check it). So our $L = L_1^{-1} L_2^{-1} =$

$$\begin{pmatrix} 1 & 0 & 0 \\ a_{21}/a_{11} & 1 & 0 \\ a_{31}/a_{11} & a_{32}^{(1)}/a_{22}^{(1)} & 1 \end{pmatrix}$$

Or just all of the nonzero elements of $-L_i$'s.

(Remember that inverse/multiplication of a lower triangular matrix is also a lower triangular matrix)

Finally we have:

$$
\begin{aligned}
A &= LU \\
&= \begin{pmatrix} 1 & 0 & 0 \\ a_{21}/a_{11} & 1 & 0 \\ a_{31}/a_{11} & a_{32}^{(1)}/a_{22}^{(1)} & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & 0 & a_{33}^{(2)} \end{pmatrix}
\end{aligned}
$$

3.2 Forward-substitution

Going back to our objective of solving $Ax = b$, now we do $\hat{b} = L^{-1}b$, or often written as $Ly = b$ ($y = \hat{b}$).

Looking at

$$
\begin{pmatrix} 1 & 0 & 0 \\ a_{21}/a_{11} & 1 & 0 \\ a_{31}/a_{11} & a_{32}^{(1)}/a_{22}^{(1)} & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}
$$

The first two equations are $y_1 = b_1$ and $\frac{a_{21}}{a_{11}} y_1 + y_2 = b_2$. But since we know y_1 from the first line, we can subsitute b_1 for y_1 and yeild $y_2 = b_2 - b_1 \frac{a_{21}}{a_{11}}$. We can keep on substituting earlier results to obtian all y_1 and this is *forward-substitution*.

Using this, our $\hat{b}$ is:

$$
\begin{aligned}
&y_1 = b_1 \\
&b_2 = y_2 + y_1 \frac{a_{22}}{a_{11}} \\
&\quad \rightarrow y_2 = b_2 - b_1 \frac{a_{22}}{a_{11}} \text{ substitute } y_1 \\
&b_3 = y_3 + y_1 \frac{a_{31}}{a_{11}} + y_2 \frac{a_{32}^{(1)}}{a_{22}^{(1)}} \\
&\quad \rightarrow y_3 = b_3 - (y_1 \frac{a_{31}}{a_{11}} + y_2 \frac{a_{32}^{(1)}}{a_{22}^{(1)}})
\end{aligned}
$$

Formally:

$$
y_m = b_m - \sum_{i=1}^{m-1} l_{mi} y_i
$$

This is long and details don't matter so just denote $\hat{b} = (\begin{smallmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \end{smallmatrix})$

3.3 Backward-substitution

Similarly, now that we have our $\hat{b}$, we do $U\hat{b} = x$. Since U is an upper triangular, it's the same thing as forward substitution but we progress from x_n ($x_n = \hat{b}_1$) For us the x is:

$$\begin{aligned}
x_3 &= \frac{\hat{b}_3}{u_{33}} \\
\hat{b}_2 &= u_{22}x_2 + u_{23}x_3 \\
&\rightarrow x_2 = \frac{\hat{b}_2 - u_{23}x_3}{u_{22}} \\
&\rightarrow x_2 = \frac{\hat{b}_2 - u_{23}\frac{\hat{b}_3}{u_{33}}}{u_{22}} \text{ substitute } x_3 \\
\hat{b}_1 &= u_{11}x_1 + u_{12}x_2 + u_{13}x_3 \\
&\rightarrow x_1 = \frac{\hat{b}_1 - (u_{12}x_2 + u_{13}x_3)}{u_{11}}
\end{aligned}$$

Formally:

$$x_m = (\hat{b}_m - \sum_{i=m+1}^{1} u_{mi}x_i)/u_{mm}$$

And finally we get our x!!

4 Pivoting

If $a_{11} = 1$, everything above falls apart. In such a case we need to permute the first row of A with another row of A with $max_j|a_{j1}|$. This is the same as multiplying $Ax = b$ with a *permutation matrix*, P_1, where all the diagnoal values are 1 but a_{11}, a_{jj} and rest all 0 but a_{1j} and a_{j1}

$$P_1 = \begin{pmatrix}
\mathbf{0} & \cdots & \cdots & \cdots & \mathbf{1} = a_{1j} & \cdots & 0 \\
0 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\
0 & 0 & 1 & \cdots & \cdots & \cdots & 0 \\
\vdots & & & & & & \\
\mathbf{1}_{(a_{j1})} & 0 & 0 & \mathbf{0}_{(a_{jj})} & \cdots & \cdots & 0 \\
\vdots & & & & & & \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & 1
\end{pmatrix}$$

So in the end, we do the same as above and compute L, U, and $P = P_1P_2\cdots P_k$ where $k \leq n-1$ and get $PA = LU$. Note that there is no mutliplication/divsion in doing PA.

5 Cost Analysis/Complexity

P is costless in terms of $flop$s (in the context of number of multiplication and divisions). Components that contribute to the cost are:

1. **LU decomposition**: to make L_1 requires $n-1$ divisions, and doing L_1A requires $(n-1)(n-1)$ multiplications. So total of $(n-1)^2+n-1 = n^2-n$ or $n(n-1)$ operations for step 1 (n diagonal elements to multpily with elements that required $(n-1)$ flops to make).

   So to do step 2 is $(n-1)(n-2)$, and so forth until $n-1$ steps are done, which is:

   $$\sum_{i=1}^{n-1} i(i+1) = sum_{i=1}^{n-1} i^2 + \sum_{i=1}^{n-1} i$$

   Recalling calc 2 arithmetic series and special sums, $\sum_{i=1}^{n} i = n(n+1)/2$ and $\sum_{i=1}^{n} i^2 = n(n+1)(2n+1)/6$, so the total sum is $\mathcal{O}n^3/3 + (n^2 + n)/2 = \mathcal{O}(2n^3/3)$

2. **forward/backward substitution**: The first(last for back) operation is 1 multiplication, the second equation requires 2 multiplication, and in the end it requires n multiplication. Disregarding the additions, this is just $\sum_{i=1}^{n} i = n(n+1)/2 = \mathcal{O}(n^2)$. This analysis holds for both, so total $\mathcal{O}(2n^2)$ here.

So in total, we have

$$\mathcal{O}(2n^3/3) + \mathcal{O}(2n^2) = \mathcal{O}(n^3)$$

for large n.