2013

The complexity of linear algebra

LeAnn Kay Christensen

The Complexity of Linear Algebra

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

in

Mathematics

by

LeeAnn Kay Christensen

June 2013

The Complexity of Linear Algebra

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

---

by

LeeAnn Kay Christensen

June 2013

Approved by:

[redacted]

Chris Freiling, Committee Chair

5/29/2013

Date

[redacted]

Gary Griffing, Committee Member

[redacted]

Laura Wallace, Committee Member

[redacted]

Peter Williams, Chair,
Department of Mathematics

[redacted]

Charles Stanton
Graduate Coordinator,
Department of Mathematics

Abstract

This paper examines the complexity of linear algebra. Complexity means how much work, or the number of calculations or time it takes, to perform a task. The hypothesis is that tasks performed using linear algebra are not more complicated than a factor of the work of matrix multiplication. If the amount of work could be reduced to multiply matrices, than other tasks could also become more efficient.

The paper first looks at reducing the amount of work it takes to perform matrix multiplication by studying Strassen's and Laderman's algorithms. Strassen's algorithm reduces the amount of work it takes to multiply 2 x 2 matrices. This idea is expanded to include multiplying any square matrices that have rows/columns as a power of two. Included are examples. Then we show how much work it takes to use Strassen's idea in a recursive formula. We also look at Laderman's algorithm and the work he did to reduce the amount of work it takes to multiply 3 x 3 matrices. We demonstrate his algorithm with an example. Finally, we compare the work of multiplying matrices by traditional methods and by Strassen's and Laderman's algorithms using examples and a table.

Next we study the amount of work it takes to perform basic linear algebra tasks that include inverses, solving a system of equations, and determinants. The paper researches sources proposing that these operations are no more complex than matrix multiplication.

We then focus on specific operations used in linear algebra–finding an inverse, solving a system of equations, and finding determinants. We look at each of these and how the work is not more than a factor of the work involved in multiplying matrices of the same size. We look at finding the inverse of a triangular matrix. And develop a formula that shows that the work of finding the inverse of an n x n matrix is less than a constant factor of multiplying two n x n matrices.

Further, we demonstrate how to take a square matrix and turn it into the product of a lower triangular, upper triangular, and permutation matrix. This is called LUP decomposition. Now any square, invertible matrix may be turned into the product of lower triangular, upper triangular and permutation matrices, and we can find the inverse. We show the LUP decomposition algorithm, using an explanation and an example. We look at the work needed to find the LUP decomposition of a square n x n matrix. The work is shown to be less than a constant factor of the work to multiply two n x n matrices. This is important because the work of solving a system of equations and finding

a determinant is based on using LUP decomposition.

Finally, we look at solving systems of equations using LUP decomposition. By example we show that this method involves matrix multiplication and that the work of solving them with it is not greater than a factor of the work needed to multiply matrices. We also show that finding a determinant can be done by performing the LUP decomposition. Therefore, the amount of work for finding a determinant is roughly the same as for finding an LUP decomposition.

This study shows that the linear operations of finding an inverse, solving a system of equations and finding a determinant are not more complex than a constant factor of multiplying two matrices of the original size. It shows ways to reduce the amount of work that it takes to multiply matrices. An even faster method, proposed by Coppersmith, not examined in this paper, merits a study similar to this one. Studying the reduction of the work involved in matrix multiplication continues.

ACKNOWLEDGEMENTS

I would like to thank Dr. Chris Freiling for all his help, support, patience and encouragement during the development of this paper. I would also like to thank Dr. Gary Griffing and Dr. Laura Wallace for their helpful comments and friendly assistance. It has been a privilege to work with people dedicated to furthering education in mathematics.

I would like to thank William for his help with the diagrams. And I would also like to thank Pennie and Ryan for their assistance. I am grateful to my family for their encouragement and support.

Table of Contents

Chapter 1

Introduction

In our ever changing world, technology has become a major factor in our daily lives. It seems that there is a constant search to have electronic devices perform faster and more conveniently. The idea of making things operate faster has been pursued in the study of linear algebra as well. If the mathematics of linear algebra could be more efficient, then computers could run faster when performing linear algebra calculations.

As linear algebra is used more and more in different fields, it becomes useful to study ways of reducing the amount of work required to complete basic procedures. This paper looks at the complexity of procedures in linear algebra. Here, "complexity" suggests the amount of work needed to complete a task. The term "work" refers to the number of calculations that are needed to perform the procedure. If the number of calculations can be reduced, then the amount of work and time required to perform the procedure is also reduced. The procedures of linear algebra studied here include finding inverses, solving systems of equations, and finding determinants. It has been proposed that these procedures can be performed through matrix multiplication and that the procedures are no more complex than matrix multiplication. Therefore, reducing the amount of work necessary to perform matrix multiplication would reduce the work needed to perform these basic procedures.

The reduction of work needed for matrix multiplication is presented. If operations are to be performed using matrix multiplication and the amount of work needed to perform that multiplication is reduced, then the amount of work needed to perform linear algebra procedures is also reduced. In 1969, Volker Strassen was the first to present an algorithm that reduced the amount of work necessary to multiply matrices[Wei99]. Strassen proposed an algorithm that reduces the number of calculations needed to perform the multiplication of two 2 x 2 matrices. Similarly, in 1975, Julian D. Laderman found an algorithm that reduces the number of calculations needed to multiply two 3 x 3 matrices [Lad79]. Their algorithms are presented with examples. Also, the procedure of finding inverses through matrix multiplication is presented.

To study other procedures of linear algebra requires working with matrices in a specific form. It is possible to rewrite a matrix as the product of simpler matrices, namely lower triangular (L), upper triangular (U), and permutation (P) matrices. This process of factoring is called LUP decomposition. Any nonsingular, square matrix may be factored by following the LUP decomposition algorithm. Then the matrix may be expressed in terms of three matrices that will be easier to work with separately. The procedure for finding an LUP decomposition is presented along with examples. An examination of this procedure is necessary before we can study the use of matrix multiplication for solving a system of linear equations and finding determinants. Finally, solving systems of linear equations and finding determinants by matrix multiplication are shown.

Chapter 2

Strassen's Algorithm

2.1 Presentation of Strassen's Algorithm

Remark 1. *Many of the equations typed will contain matrices with emphasis on the row and column number for the entry. For notation, we will use a subscript with the first digit designating the row number and the second digit representing the column number. For example a_{23} would represent the entry of the matrix in row 2, column 3. While this may not be the ideal notation for many cases where there is discussion of matrix entries, there should not be confusion here because all matrices that may have a row or column number larger than a single digit are represented by a single-letter variable. There are several lengthy equations where this notation allows brevity in typing and ease in reading.*

To study Strassen's Algorithm, consider the traditional methods of matrix multiplication. Matrix multiplication is not commutative. Not all matrices may be multiplied. The ability to multiply matrices depends on their size. The number of columns in the first matrix must be the same as the number of rows in the second matrix. Each entry in the product is found by aligning the elements of the row in the first matrix with the entries of the column in the second matrix. These terms are multiplied and then the entry in the product matrix is found by adding those products. The resulting matrix will have the same number of rows as the first matrix and the same number of colums as the second matrix. Therefore, the difficulty in finding the product's entries is based on the number of columns in the first matrix. If the first matrix has many columns, there will be many multiplication steps for each entry in the product matrix.

For example, consider the product of an m x n matrix times an n x p matrix

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{n1} & \dots & b_{nj} & \dots & b_{np} \end{pmatrix} =$$

$$\begin{pmatrix} c_{11} & \dots & c_{1j} & \dots & c_{1p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{i1} & \dots & c_{ij} & \dots & c_{ip} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{m1} & \dots & c_{mj} & \dots & c_{mp} \end{pmatrix}.$$

The entry in the i^{th} row and j^{th} column of the product matrix is

$$c_{ij} = a_{i1}b_{1j} + \ldots + a_{ik}b_{kj} + \ldots + a_{in}b_{nj}.$$

There are n number of multiplications to find each entry of the product matrix. There are $n-1$ number of additions to find each entry. There are mp number of entries in the product matrix, each requiring n multiplications and $n-1$ additions. Thus, to find the product matrix requires mpn multiplications and $mp(n-1)$ addition steps.

Strassen considered the possiblility of reducing the amount of work necessary to multiply matrices. Strassen's Algorithm is presented in the article, "Geometry and the Complexity of Matrix Multiplication" by J. M. Landsberg [Lan08]. It is also presented in the book "The Design and Analysis of Computer Algorithms" by Aho, Hopcroft, and Ullman [AHU74]. There are differences in the way the two algorithms are shown and not all equations are written the same. Here Landsberg's equations are used for discussion of Strassen's Algorithm.

To study Strassen's Algorithm, consider traditional matrix multiplication methods for mutliplying two 2 x 2 matrices. That is,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.$$

The above product matrix contains 8 multiplication steps and 4 addition steps. Strassen's goal was to find a method that would reduce the number of multiplication steps when multiplying matrices. His algorithm for multiplying two 2 x 2 matrices uses only seven multiplications. While this reduction may not seem significant for a small matrix, the amount of work would be greatly reduced for much larger matrices.

The following is a presentation of Strassen's Algorithm [Lan08].

Algorithm 1. *Suppose,*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}.$$

The following muliplications are made using the entries of the two matrices being multiplied.

$$\begin{aligned}
\mathrm{I} &= (a_{11} + a_{22}) \cdot (b_{11} + b_{22}) \\
\mathrm{II} &= (a_{21} + a_{22}) \cdot\ b_{11} \\
\mathrm{III} &= a_{11}\,(b_{12} - b_{22}) \\
\mathrm{IV} &= a_{22}\,(-b_{11} + b_{21}) \\
\mathrm{V} &= (a_{11} + a_{12})\, b_{22} \\
\mathrm{VI} &= (-a_{11} + a_{21})\,(b_{11} + b_{12}) \\
\mathrm{VII} &= (a_{12} - a_{22})\,(b_{21} + b_{22})
\end{aligned}$$

The entries of matrix C can now be found by using the above multiplications in the following equations:

$$\begin{aligned}
c_{11} &= \mathrm{I} + \mathrm{IV} - \mathrm{V} + \mathrm{VII}, \\
c_{12} &= \mathrm{III} + \mathrm{V}, \\
c_{21} &= \mathrm{II} + \mathrm{IV}, \\
c_{22} &= \mathrm{I} + \mathrm{III} - \mathrm{II} + \mathrm{VI}.
\end{aligned}$$

Consider each entry in the two matrices to have real number values, then the alogrithm uses 7 multiplication and 18 addition steps.

An example of using Strassen's Algorithm follows.

$$\begin{pmatrix} 2 & -1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -1 & -2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

Using Strassen's method gives the following information:

$$\begin{aligned}
\mathrm{I} &= (2+4)(-1+1) &&= 0, \\
\mathrm{II} &= (3+4)-1 &&= -7, \\
\mathrm{III} &= 2(-2-1) &&= -6, \\
\mathrm{IV} &= 4(-(-1)+3) &&= 16, \\
\mathrm{V} &= (2+-1)\,1 &&= 1, \\
\mathrm{VI} &= (-(2)+3)(-1+-2) &&= -3, \\
\mathrm{VII} &= (-1-4)(3+1) &&= -20, \\
c_{11} &= 0+16-1+(-20) &&= -5, \\
c_{21} &= -7+16 &&= 9, \\
c_{12} &= -6+1 &&= -5, \\
c_{22} &= 0+(-6)-(-7)+(-3) &&= -2.
\end{aligned}$$

The algorithm yields

$$\begin{pmatrix} 2 & -1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -1 & -2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} -5 & -5 \\ 9 & -2 \end{pmatrix}.$$

2.2 Verification of Strassen's Algorithm

The following is a verification of Strassen's Algorithm by comparing the output of traditional matrix multiplication with the output from the algorithm. Let

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}.$$

As mentioned earlier, traditional methods yield

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.$$

Now, evaluate each entry in the product. In the simplification for the following equations, underbraces show which terms cancel. The matching numbers underneath the braces indicate that those two terms add to zero and are no longer necessary to the equation. For the first entry, we know $c_{11} = a_{11}b_{11} + a_{12}b_{21}$. The algorithm produces

$$\begin{aligned} c_{11} &= \mathrm{I} + \mathrm{IV} - \mathrm{V} + \mathrm{VII} \\ &= (a_{11} + a_{22})(b_{11} + b_{22}) + a_{22}(\neg b_{11} + b_{21}) - (a_{11} + a_{12})b_{22} + (a_{12} - a_{22})(b_{21} + b_{22}) \\ &= a_{11}b_{11} + \underbrace{a_{11}b_{22}}_{1} + \underbrace{a_{22}b_{11}}_{2} + \underbrace{a_{22}b_{22}}_{3} - \underbrace{a_{22}b_{11}}_{2} + \underbrace{a_{22}b_{21}}_{4} - \underbrace{a_{11}b_{22}}_{1} - \underbrace{a_{12}b_{22}}_{5} + \\ &\quad + a_{12}b_{21} + \underbrace{a_{12}b_{22}}_{5} - \underbrace{a_{22}b_{21}}_{4} - \underbrace{a_{22}b_{22}}_{3}, \text{ and so} \\ &= a_{11}b_{11} + a_{12}b_{21}. \end{aligned}$$

The next two entries are not as lengthy to verify. From traditional multiplication methods, $c_{12} = a_{11}b_{12} + a_{12}b_{22}$ and $c_{21} = a_{21}b_{11} + a_{22}b_{21}$.

$$\begin{aligned} c_{12} &= \mathrm{III} + \mathrm{V} \\ c_{12} &= a_{11}(b_{12} - b_{22}) + (a_{11} + a_{12})b_{22} \\ c_{12} &= a_{11}b_{12} - a_{11}b_{22} + a_{11}b_{22} + a_{12}b_{22} \\ c_{12} &= a_{11}b_{12} + a_{12}b_{22} \end{aligned}$$

Similarily

$$\begin{aligned} c_{21} &= \mathrm{II} + \mathrm{IV} \\ c_{21} &= (a_{21} + a_{22})b_{11} + a_{22}(-b_{11} + b_{21}) \\ c_{21} &= a_{21}b_{11} + a_{22}b_{11} - a_{22}b_{11} + a_{22}b_{21} \\ c_{21} &= a_{21}b_{11} + a_{22}b_{21}. \end{aligned}$$

Finally, the last entry in the matrix simplifies as

$$
\begin{aligned}
c_{22} &= \mathrm{I} + \mathrm{III} - \mathrm{II} + \mathrm{VI} \\
c_{22} &= (a_{11}+a_{22})(b_{11}+b_{22}) + a_{11}(b_{12}-b_{22}) - (a_{21}+a_{22})b_{11} + (-a_{11}+a_{21})(b_{11}+b_{12}) \\
&= \underbrace{a_{11}b_{11}}_{1} + \underbrace{a_{11}b_{22}}_{2} + \underbrace{a_{22}b_{11}}_{3} + a_{22}b_{22} + \underbrace{a_{11}b_{12}}_{4} - \underbrace{a_{11}b_{22}}_{2} - \underbrace{a_{21}b_{11}}_{5} - \\
&\quad - \underbrace{a_{22}b_{11}}_{3} - \underbrace{a_{11}b_{11}}_{1} - \underbrace{a_{11}b_{12}}_{4} + \underbrace{a_{21}b_{11}}_{5} + a_{21}b_{12} \\
c_{22} &= a_{21}b_{12} + a_{22}b_{22}.
\end{aligned}
$$

Therefore, using Strassen's Algorithm finds the entries of the product matrix to be the same as if they were found by traditional matrix multiplication methods.

2.3 Expanding Strassen's Algorithm

The intent of the algorithm is to reduce the amount of work necessary to multiply matrices. Consider in more detail the amount of work it takes to multiply two 2 x 2 matrices using Strassen's Algorithm compared to the traditional method of multiplying matrices. In Strassen's Algorithm, it takes 7 multiplications and 18 additions for a total of 25 arithmetic operations. Using traditional methods, we have 8 multiplications and 4 additions for a total of 12 arithmetic operations.

Now, consider if the matrices being multiplied were larger. Expand Strassen's method to a product of two 4 x 4 matrices. Each 4 x 4 matrix is divided into four 2 x 2 matrices. The expectation is that the amount of work will be significantly reduced by using Strassen's Algorithm more than once. Consider

$$
\left(\begin{array}{cc|cc}
a_{11} & a_{12} & a_{13} & a_{14} \\
a_{21} & a_{22} & a_{23} & a_{24} \\
\hline
a_{31} & a_{32} & a_{33} & a_{34} \\
a_{41} & a_{42} & a_{43} & a_{44}
\end{array}\right) \cdot
\left(\begin{array}{cc|cc}
b_{11} & b_{12} & b_{13} & b_{14} \\
b_{21} & b_{22} & b_{23} & b_{24} \\
\hline
b_{31} & b_{32} & b_{33} & b_{34} \\
b_{41} & b_{42} & b_{43} & b_{44}
\end{array}\right).
$$

Let

$$
A_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad A_{12} = \begin{pmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{pmatrix}
$$

$$A_{21} = \begin{pmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{pmatrix} \qquad A_{22} = \begin{pmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{pmatrix}$$

$$B_{11} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \qquad B_{12} = \begin{pmatrix} b_{13} & b_{14} \\ b_{23} & b_{24} \end{pmatrix}$$

$$B_{21} = \begin{pmatrix} b_{31} & b_{32} \\ b_{41} & b_{42} \end{pmatrix} \qquad B_{22} = \begin{pmatrix} b_{33} & b_{34} \\ b_{43} & b_{44} \end{pmatrix}.$$

The multiplication becomes

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$

Strassen's method reduces the operation to the following seven 2 x 2 multiplications:

$$\begin{aligned} \mathrm{I} &= (A_{11} + A_{22}) \cdot (B_{11} + B_{22}), \\ \mathrm{II} &= (A_{21} + A_{22}) \cdot B_{11}, \\ \mathrm{III} &= A_{11}(B_{12} - B_{22}), \\ \mathrm{IV} &= A_{22}(-B_{11} + B_{21}), \\ \mathrm{V} &= (A_{11} + A_{12})B_{22}, \\ \mathrm{VI} &= (-A_{11} + A_{21})(B_{11} + B_{12}), \\ \mathrm{VII} &= (A_{12} - A_{22})(B_{21} + B_{22}). \end{aligned}$$

Each 2 x 2 multiplication takes seven multiplications and 18 additions. Therefore a 4 x 4 matrix multiplication is reduced to 49 multiplications and 198 additions instead of 64 multiplications and 48 additions as with traditional methods. This may not seem like a reduction of work because the overall number of operations went from 12 to 25. However, the number of multiplications was reduced. Think of multiplication of whole numbers as just an advanced form of addition and addition as just an advanced form of counting. Say you want to add $56+82$. An electronic device counts to 56 and then counts 82 more. But if you want to multiply (56)(82), then the device must count 56 a total of 82 times. The counting cycle is increased from two cycles to 82 cycles. Therefore, the amount of work involved for multiplication can be much more significant as the numbers increase.

The following is an example of multiplication with two 4 x 4 matrices having entries from the set of real numbers. Consider

$$\left(\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ -1 & 3 & 1 & 4 \\ \hline 2 & 2 & 0 & 1 \\ 3 & 2 & 1 & 1 \end{array}\right) \cdot \left(\begin{array}{cc|cc} 3 & 2 & 1 & 1 \\ 1 & 0 & -2 & -1 \\ \hline 2 & 1 & 2 & 0 \\ 3 & 2 & 2 & 3 \end{array}\right).$$

Then

$$\mathrm{I} = \left[\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 2 & 3 \end{pmatrix}\right],$$

and so

$$\mathrm{I} = \begin{pmatrix} 1 & 3 \\ 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 5 & 2 \\ 3 & 3 \end{pmatrix}.$$

Now Strassen's Algorithm is applied to multiply the two 2 x 2 matrices. Using lower case Roman numerals for Strassen's Algorithm, the following values are found.

$$\mathrm{i} = 40, \quad \mathrm{ii} = 20, \quad \mathrm{iii} = -1, \quad \mathrm{iv} = -8, \quad \mathrm{v} = 12, \quad \mathrm{vi} = -7, \quad \mathrm{vii} = -6.$$

Thus

$$\mathrm{I} = \begin{pmatrix} 14 & 11 \\ 12 & 12 \end{pmatrix}.$$

The other matrices are found for II, III, IV, V, VI, and VII:

$$\mathrm{II} = \left[\begin{pmatrix} 2 & 2 \\ 3 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\right] \cdot \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix}$$

$$\mathrm{II} = \begin{pmatrix} 2 & 3 \\ 4 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix}$$

$$\mathrm{i} = 15, \quad \mathrm{ii} = 21, \quad \mathrm{iii} = 4, \quad \mathrm{iv} = -6, \quad \mathrm{v} = 0, \quad \mathrm{vi} = 10, \quad \mathrm{vii} = 0;$$

$$\mathrm{II} = \begin{pmatrix} 9 & 4 \\ 15 & 8 \end{pmatrix}$$

$$\begin{aligned}
\mathrm{III} &= \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \cdot \left[\begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 2 & 3 \end{pmatrix} \right] \\
\mathrm{III} &= \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} -1 & 1 \\ -4 & -4 \end{pmatrix}
\end{aligned}$$

$\mathrm{i} = -20, \quad \mathrm{ii} = -2, \quad \mathrm{iii} = 5, \quad \mathrm{iv} = -9, \quad \mathrm{v} = -12, \quad \mathrm{vi} = 0, \quad \mathrm{vii} = 8;$

$$\mathrm{III} = \begin{pmatrix} -9 & -7 \\ -11 & -13 \end{pmatrix}$$

$$\begin{aligned}
\mathrm{IV} &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \left[- \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} \right] \\
\mathrm{IV} &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & -1 \\ 2 & 2 \end{pmatrix}
\end{aligned}$$

$\mathrm{i} = 1, \quad \mathrm{ii} = -2, \quad \mathrm{iii} = 0, \quad \mathrm{iv} = 3, \quad \mathrm{v} = 2, \quad \mathrm{vi} = -2, \quad \mathrm{vii} = 0;$

$$\mathrm{IV} = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}$$

$$\begin{aligned}
\mathrm{V} &= \left[\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 4 \end{pmatrix} \right] \cdot \begin{pmatrix} 2 & 0 \\ 2 & 3 \end{pmatrix} \\
\mathrm{V} &= \begin{pmatrix} 1 & 3 \\ 0 & 7 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 2 & 3 \end{pmatrix}
\end{aligned}$$

$\mathrm{i} = 40, \quad \mathrm{ii} = 14, \quad \mathrm{iii} = -3, \quad \mathrm{iv} = 0, \quad \mathrm{v} = 12, \quad \mathrm{vi} = -2, \quad \mathrm{vii} = -20;$

$$\mathrm{V} = \begin{pmatrix} 8 & 9 \\ 14 & 21 \end{pmatrix}$$

$$
\begin{aligned}
\mathrm{VI} &= \left[-\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 3 & 2 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix}\right] \\
\mathrm{VI} &= \begin{pmatrix} 1 & 0 \\ 4 & -1 \end{pmatrix} \cdot \begin{pmatrix} 4 & 3 \\ -1 & -1 \end{pmatrix}
\end{aligned}
$$

$$
\mathrm{i} = 0, \quad \mathrm{ii} = 12, \quad \mathrm{iii} = 4, \quad \mathrm{iv} = 5, \quad \mathrm{v} = -1, \quad \mathrm{vi} = 21, \quad \mathrm{vii} = -2;
$$

$$
\mathrm{VI} = \begin{pmatrix} 4 & 3 \\ 17 & 13 \end{pmatrix}
$$

$$
\begin{aligned}
\mathrm{VII} &= \left[\begin{pmatrix} 0 & 1 \\ 1 & 4 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 2 & 3 \end{pmatrix}\right] \\
\mathrm{VII} &= \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 4 & 1 \\ 5 & 5 \end{pmatrix}
\end{aligned}
$$

$$
\mathrm{i} = 27, \quad \mathrm{ii} = 12, \quad \mathrm{iii} = 0, \quad \mathrm{iv} = 3, \quad \mathrm{v} = 0, \quad \mathrm{vi} = 0, \quad \mathrm{vii} = -30;
$$

$$
\mathrm{VII} = \begin{pmatrix} 0 & 0 \\ 15 & 15 \end{pmatrix}
$$

After calculating the necessary multiplcations, the following additions are needed to finish the algorithm:

$$
\begin{aligned}
C_{11} &= \mathrm{I} + \mathrm{IV} - \mathrm{V} + \mathrm{VII} \\
&= \begin{pmatrix} 14 & 11 \\ 12 & 12 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} - \begin{pmatrix} 8 & 9 \\ 14 & 21 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 15 & 15 \end{pmatrix} \\
&= \begin{pmatrix} 8 & 4 \\ 14 & 7 \end{pmatrix};
\end{aligned}
$$

$$
\begin{aligned}
C_{12} &= \mathrm{III} + \mathrm{V} \\
&= \begin{pmatrix} -9 & -7 \\ -11 & -13 \end{pmatrix} + \begin{pmatrix} 8 & 9 \\ 14 & 21 \end{pmatrix} \\
&= \begin{pmatrix} -1 & 2 \\ 3 & 8 \end{pmatrix};
\end{aligned}
$$

$$
\begin{aligned}
C_{21} &= \mathrm{II} + \mathrm{IV} \\
&= \begin{pmatrix} 9 & 4 \\ 15 & 8 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 11 & 6 \\ 16 & 9 \end{pmatrix};
\end{aligned}
$$

$$
\begin{aligned}
C_{22} &= \mathrm{I} + \mathrm{III} - \mathrm{II} + \mathrm{VI} \\
&= \begin{pmatrix} 14 & 11 \\ 12 & 12 \end{pmatrix} + \begin{pmatrix} -9 & -7 \\ -11 & -13 \end{pmatrix} - \begin{pmatrix} 9 & 4 \\ 15 & 8 \end{pmatrix} \begin{pmatrix} 4 & 3 \\ 17 & 13 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 3 \\ 3 & 4 \end{pmatrix}.
\end{aligned}
$$

The resulting product is found by placing each 2 x 2 matrix in the appropriate location of the product matrix.

$$
\left(\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ -1 & 3 & 1 & 4 \\ \hline 2 & 2 & 0 & 1 \\ 3 & 2 & 1 & 1 \end{array}\right) \cdot \left(\begin{array}{cc|cc} 3 & 2 & 1 & 1 \\ 1 & 0 & -2 & -1 \\ \hline 2 & 1 & 2 & 0 \\ 3 & 2 & 2 & 3 \end{array}\right) = \left(\begin{array}{cc|cc} 8 & 4 & -1 & 2 \\ 14 & 7 & 3 & 8 \\ \hline 11 & 6 & 0 & 3 \\ 16 & 9 & 3 & 4 \end{array}\right)
$$

The product is complete for this example of multiplying two 4 x 4 matrices.

2.4 Work Needed to Complete Strassen's Algorithm

Now look at the amount of work for finding the product of two n x n matrices such that $n = 2^k$ using Strassen's Algorithm compared to traditional methods. As the matrices get larger, there should be a greater reduction in work by using the algorithm.

Let $T(n)$ represent the number of operations needed to complete the multiplication of two n x n matrices.

Consider the following square matrix where n represents the number of rows and columns and $n = 2^k$ with k a natural number:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Partitioning the matrix into four $\frac{n}{2}$ x $\frac{n}{2}$ submatrices, we get

$$\left(\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1\frac{n}{2}} & a_{1(\frac{n}{2}+1)} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{\frac{n}{2}1} & \cdots & a_{\frac{n}{2}\frac{n}{2}} & a_{\frac{n}{2}(\frac{n}{2}+1)} & \cdots & a_{\frac{n}{2}n} \\ \hline a_{(\frac{n}{2}+1)1} & \cdots & a_{(\frac{n}{2}+1)\frac{n}{2}} & a_{(\frac{n}{2}+1)(\frac{n}{2}+1)} & \cdots & a_{(\frac{n}{2}+1)n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{n\frac{n}{2}} & a_{n(\frac{n}{2}+1)} & \cdots & a_{nn} \end{array}\right).$$

Now consider multiplying this matrix by a similar n x n matrix that has been partitioned in the same manner. Looking at each submatrix as an entry, there are two 2 x 2 matrices being multiplied. To accomplish this requires seven multiplication steps and 18 addition steps. Each multiplication will require the amount of work needed to multiply two $\frac{n}{2}$ x $\frac{n}{2}$ matrices. This work can be represented as $7T\left(\frac{n}{2}\right)$. An addition step involves adding two $\frac{n}{2}$ x $\frac{n}{2}$ matrices meaning $\left(\frac{n}{2}\right)^2$ entry additions. Since there are 18 addition steps when applying the algorithm, there are a total of $18\left(\frac{n}{2}\right)^2$ individual addition steps. Note that some of the work represented by $7T\left(\frac{n}{2}\right)$ includes addition steps also. The total number of operations for multiplying the two n x n matrices is represented by the following formula

$$T(n) = 7T\left(\frac{n}{2}\right) + 18\left(\frac{n}{2}\right)^2.$$

Looking at the traditional method of multiplying matrices, the amount of work needed would be n multiplications and $(n-1)$ additions for each entry. There would be n^2 entries. The total amount of work, $T(n)$, for multiplying two n x n matrices is

$$T(n) = n^3 + n^2\,(n-1)\,.$$

Table 2.1: Comparative Work for Multiplying Matrices

Traditional Method				Strassen's Algorithm	
k	2^kx2^k	mult	add	mult	add steps
1	2 x 2	8	4	7	18
2	4 x 4	64	48	49	198
3	8 x 8	512	448	343	1674
4	16 x 16	4096	3840	2401	12,870
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	2^kx2^k	$2^{3k} = 8^k$	$2^{3k} - 2^{2k}$	7^k	$7T\left(2^{k-1}\right) + 18\left(2^{2k-2}\right) - 7^k$

Table 2.1 compares the amount of work for matrix multiplication using traditional methods with the amount of work using Strassen's Algorithm.

Looking at the chart and the number of calculations needed for the two methods, it shows that while the number of multiplication steps decreases, the number of addition steps greatly increases. This may not seem like a reduction in work. However, when using large amounts of data in linear systems of higher dimensions, the reduction of multiplication steps may be of significant benefit over the increase of addition steps.

Strassen's Algorithm can be used to multiply 2^k x 2^k matrices using 7^k multiplications. Since not all matrices are of order 2^k x 2^k, additional rows and/or columns of zeros may be inserted such that dimensions are n x n where $n = 2^k$. This means $k = \log_2 n$. Therefore, any two n x n matrices may be multiplied using $7^{\lceil \log_2 n \rceil}$ multiplications. Through principles of logarithms, we see that $7^{\log_2 n} = n^{\log_2 7}$ and $\log_2 7 \approx 2.81$. Thus, $n^{\log_2 7} < n^3$.

Chapter 3

Laderman's Algorithm

3.1 Presentation of Laderman's Algorithm

Laderman looked at finding an algorithm for multiplying 3 x 3 matrices that would be more efficient than traditional methods. Using the findings from Section 2.4, multiplying two 3 x 3 matrices would require at most $7^{\log_2 3} \approx 21.8$ multiplications. To improve on Strassen's algorithm, he would need to find one that required 21 or fewer multiplications. Gastinel found an algorithm requiring 25 multiplications and Hopcraft and Kerr found one requiring 24 multiplications. However, Laderman was able to find one using 23 multiplications. His algorithm is presented as follows [Lad76].

Algorithm 2. *Consider the product*

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}.$$

Let m_n denote the different multiplication steps specified by the following equations

$$m_1 = (a_{11} + a_{12} + a_{13} - a_{21} - a_{22} - a_{32} - a_{33})\,, \cdot b_{22}$$
$$m_2 = (a_{11} - a_{21}) \cdot (-b_{12} + b_{22})\,,$$
$$m_3 = a_{22}\,(-b_{11} + b_{12} + b_{21} - b_{22} - b_{23} - b_{31} + b_{33})\,,$$

$$m_4 = (-a_{11} + a_{21} + a_{22}) \cdot (b_{11} - b_{12} + b_{22}),$$

$$m_5 = (a_{21} + a_{22}) \cdot (-b_{11} + b_{12}),$$

$$m_6 = a_{11}b_{11},$$

$$m_7 = (-a_{11} + a_{31} + a_{32}) \cdot (b_{11} - b_{13} + b_{23}),$$

$$m_8 = (-a_{11} + a_{31}) \cdot (b_{13} - b_{23}),$$

$$m_9 = (a_{31} + a_{32}) \cdot (-b_{11} + b_{13}),$$

$$m_{10} = (a_{11} + a_{12} + a_{13} - a_{22} - a_{23} - a_{31} - a_{32}) \cdot b_{23},$$

$$m_{11} = a_{32}(-b_{11} + b_{13} + b_{21} - b_{22} - b_{23} - b_{31} + b_{32}),$$

$$m_{12} = (-a_{13} + a_{32} + a_{33}) \cdot (b_{22} + b_{31} - b_{32}),$$

$$m_{13} = (a_{13} - a_{33}) \cdot (b_{22} - b_{32}),$$

$$m_{14} = a_{13}b_{31},$$

$$m_{15} = (a_{32} + a_{33}) \cdot (-b_{31} + b_{32}),$$

$$m_{16} = (-a_{13} + a_{22} + a_{23}) \cdot (b_{23} + b_{31} - b_{33}),$$

$$m_{17} = (a_{13} - a_{23}) \cdot (b_{23} - b_{33}),$$

$$m_{18} = (a_{22} + a_{23}) \cdot (-b_{31} + b_{33}),$$

$$m_{19} = a_{12}b_{21},$$

$$m_{20} = a_{23}b_{32},$$

$$m_{21} = a_{21}b_{13},$$

$$m_{22} = a_{31}b_{12}, \text{ and}$$

$$m_{23} = a_{33}b_{33}.$$

After the above calculations, the entries of the desired matrix can be found with the following additions:

$$c_{11} = m_6 + m_{14} + m_{19},$$

$$c_{12} = m_1 + m_4 + m_5 + m_6 + m_{12} + m_{14} + m_{15},$$

$$c_{13} = m_6 + m_7 + m_9 + m_{10} + m_{14} + m_{16} + m_{18},$$

$$c_{21} = m_2 + m_3 + m_4 + m_6 + m_{14} + m_{16} + m_{17},$$

$$c_{22} = m_2 + m_4 + m_5 + m_6 + m_{20},$$

$$c_{23} = m_{14} + m_{16} + m_{17} + m_{18} + m_{21},$$

$$c_{31} = m_6 + m_7 + m_8 + m_{11} + m_{12} + m_{13} + m_{14},$$

$$c_{32} = m_{12} + m_{13} + m_{14} + m_{15} + m_{22}, \text{ and}$$

$$c_{33} = m_6 + m_7 + m_8 + m_9 + m_{23}.$$

This algorithm contains 23 multiplication steps and 98 addition steps and is specific to multiplying two 3 x 3 matrices.

3.2 Verification of Laderman's Algorithm

Laderman's Algorithm is shown to be correct by comparing the results of traditional matrix multiplication with those of Ladermans. Let

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}.$$

Each entry of the matrix is shown using Laderman's Algorithm. After simplification, with underbraces designating terms that canel, we see that each entry is the same as using traditional multiplication methods.

$$\begin{aligned} c_{11} &= m_6 + m_{14} + m_{19} \\ &= a_{11}b_{11} + a_{13}b_{31} + a_{12}b_{21}, \\ c_{12} &= m_1 + m_4 + m_5 + m_6 + m_{12} + m_{14} + m_{15} \\ &= (a_{11} + a_{12} + a_{13} - a_{21} - a_{22} - a_{32} - a_{33})\, b_{22} + \\ &\quad (-a_{11} + a_{21} + a_{22})\,(b_{11} - b_{12} + b_{22}) + \\ &\quad (a_{21} + a_{22})\,(-b_{11} + b_{12}) + a_{11}b_{11} + \\ &\quad (-a_{13} + a_{32} + a_{33})\,(b_{22} + b_{31} - b_{32}) + \\ &\quad a_{13}b_{31} + (a_{32} + a_{33})\,(-b_{31} + b_{32}) \end{aligned}$$

$$= \underbrace{a_{11}b_{22}}_{1} + a_{12}b_{22} + \underbrace{a_{13}b_{22}}_{2} - \underbrace{a_{21}b_{22}}_{3} - \underbrace{a_{22}b_{22}}_{4} - \underbrace{a_{32}b_{22}}_{5} - \underbrace{a_{33}b_{22}}_{6} -$$

$$- \underbrace{a_{11}b_{11}}_{7} + a_{11}b_{12} - \underbrace{a_{11}b_{22}}_{1} + \underbrace{a_{21}b_{11}}_{8} - \underbrace{a_{21}b_{12}}_{9} + \underbrace{a_{21}b_{22}}_{3} + \underbrace{a_{22}b_{11}}_{10} -$$

$$- \underbrace{a_{22}b_{12}}_{11} + \underbrace{a_{22}b_{22}}_{4} - \underbrace{a_{21}b_{11}}_{8} + \underbrace{a_{21}b_{12}}_{9} - \underbrace{a_{22}b_{11}}_{10} + \underbrace{a_{22}b_{12}}_{11} + \underbrace{a_{11}b_{11}}_{7} -$$

$$- \underbrace{a_{13}b_{22}}_{2} - \underbrace{a_{13}b_{31}}_{12} + a_{13}b_{32} + \underbrace{a_{32}b_{22}}_{5} + \underbrace{a_{32}b_{31}}_{13} - \underbrace{a_{32}b_{32}}_{14} + \underbrace{a_{33}b_{22}}_{6} +$$

$$+ \underbrace{a_{33}b_{31}}_{15} - \underbrace{a_{33}b_{32}}_{16} + \underbrace{a_{13}b_{31}}_{12} - \underbrace{a_{32}b_{31}}_{13} + \underbrace{a_{32}b_{32}}_{14} - \underbrace{a_{33}b_{31}}_{15} + \underbrace{a_{33}b_{32}}_{16}$$

$$= a_{12}b_{22} + a_{11}b_{12} + a_{13}b_{32},$$

$$c_{13} = m_6 + m_7 + m_9 + m_{10} + m_{14} + m_{16} + m_{18}$$

$$= a_{11}b_{11} + (-a_{11} + a_{31} + a_{32})(b_{11} - b_{13} + b_{23}) +$$

$$+ (a_{31} + a_{32})(-b_{11} + b_{13}) +$$

$$+ (a_{11} + a_{12} + a_{13} - a_{22} - a_{23} - a_{31} - a_{32})\, b_{23} + a_{13}b_{31} +$$

$$+ (-a_{13} + a_{22} + a_{23})(b_{23} + b_{31} - b_{33}) + (a_{22} + a_{23})(-b_{31} + b_{33})$$

$$= \underbrace{a_{11}b_{11}}_{1} - \underbrace{a_{11}b_{11}}_{1} + a_{11}b_{13} - \underbrace{a_{11}b_{23}}_{2} + \underbrace{a_{31}b_{11}}_{3} - \underbrace{a_{31}b_{13}}_{4} + \underbrace{a_{31}b_{23}}_{5} +$$

$$+ \underbrace{a_{32}b_{11}}_{6} - \underbrace{a_{32}b_{13}}_{7} + \underbrace{a_{32}b_{23}}_{8} - \underbrace{a_{31}b_{11}}_{3} + \underbrace{a_{31}b_{13}}_{4} - \underbrace{a_{32}b_{11}}_{6} + \underbrace{a_{32}b_{13}}_{7} +$$

$$+ \underbrace{a_{11}b_{23}}_{2} + a_{12}b_{23} + \underbrace{a_{13}b_{23}}_{9} - \underbrace{a_{22}b_{23}}_{10} - \underbrace{a_{23}b_{23}}_{11} - \underbrace{a_{31}b_{23}}_{5} - \underbrace{a_{32}b_{23}}_{8} +$$

$$+ \underbrace{a_{13}b_{31}}_{12} - \underbrace{a_{13}b_{23}}_{9} - \underbrace{a_{13}b_{31}}_{12} + a_{13}b_{33} + \underbrace{a_{22}b_{23}}_{10} + \underbrace{a_{22}b_{31}}_{13} - \underbrace{a_{22}b_{33}}_{14} +$$

$$+ \underbrace{a_{23}b_{23}}_{11} + \underbrace{a_{23}b_{31}}_{15} - \underbrace{a_{23}b_{33}}_{16} - \underbrace{a_{22}b_{31}}_{13} + \underbrace{a_{22}b_{33}}_{14} - \underbrace{a_{23}b_{31}}_{15} + \underbrace{a_{23}b_{33}}_{16}$$

$$= a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33},$$

$$c_{21} = m_2 + m_3 + m_4 + m_6 + m_{14} + m_{16} + m_{17}$$

$$= (a_{11} - a_{21})(-b_{12} + b_{22}) + a_{22}(-b_{11} + b_{12} + b_{21} - b_{22} - b_{23} - -b_{31} + b_{33}) +$$

$$+ (-a_{11} + a_{21} + a_{22})(b_{11} - b_{12} + b_{22}) + a_{11}b_{11} + a_{13}b_{31} +$$

$$+ (-a_{13} + a_{22} + a_{23})(b_{23} + b_{31} - b_{33}) + (a_{13} - a_{23})(b_{23} - b_{33})$$

$$
\begin{aligned}
&= \underbrace{-a_{11}b_{12}}_{1} + \underbrace{a_{11}b_{22}}_{2} + \underbrace{a_{21}b_{12}}_{3} - \underbrace{a_{21}b_{22}}_{4} - \underbrace{a_{22}b_{11}}_{5} + \underbrace{a_{22}b_{12}}_{6} + a_{22}b_{21} - \\
&- \underbrace{a_{22}b_{22}}_{7} - \underbrace{a_{22}b_{23}}_{8} - \underbrace{a_{22}b_{31}}_{9} + \underbrace{a_{22}b_{33}}_{10} - \underbrace{a_{11}b_{11}}_{11} + \underbrace{a_{11}b_{12}}_{1} - \underbrace{a_{11}b_{22}}_{2} + \\
&+ a_{21}b_{11} - \underbrace{a_{21}b_{12}}_{3} + \underbrace{a_{21}b_{22}}_{4} + \underbrace{a_{22}b_{11}}_{5} - \underbrace{a_{22}b_{12}}_{6} + \underbrace{a_{22}b_{22}}_{7} + \underbrace{a_{11}b_{11}}_{11} + \\
&+ \underbrace{a_{13}b_{31}}_{12} - \underbrace{a_{13}b_{23}}_{13} - \underbrace{a_{13}b_{31}}_{12} + \underbrace{a_{13}b_{33}}_{14} + \underbrace{a_{22}b_{23}}_{8} + \underbrace{a_{22}b_{31}}_{9} - \underbrace{a_{22}b_{33}}_{10} + \\
&+ \underbrace{a_{23}b_{23}}_{15} + a_{23}b_{31} - \underbrace{a_{23}b_{33}}_{16} + \underbrace{a_{13}b_{23}}_{13} - \underbrace{a_{13}b_{33}}_{14} - \underbrace{a_{23}b_{23}}_{15} + \underbrace{a_{23}b_{33}}_{16} \\
&= a_{22}b_{21} + a_{21}b_{11} + a_{23}b_{31},
\end{aligned}
$$

$$
\begin{aligned}
c_{22} &= m_2 + m_4 + m_5 + m_6 + m_{20} \\
&= (a_{11} - a_{21})(-b_{12} + b_{22}) + (-a_{11} + a_{21} + a_{22})(b_{11} - b_{12} + b_{22}) + \\
&+ (a_{21} + a_{22})(-b_{11} + b_{12}) + a_{11}b_{11} + a_{23}b_{32} \\
&= -\underbrace{a_{11}b_{12}}_{1} + \underbrace{a_{11}b_{22}}_{2} + a_{21}b_{12} - \underbrace{a_{21}b_{22}}_{3} - \underbrace{a_{11}b_{11}}_{4} + \underbrace{a_{11}b_{12}}_{1} - \underbrace{a_{11}b_{22}}_{2} + \\
&+ \underbrace{a_{21}b_{11}}_{5} - \underbrace{a_{21}b_{12}}_{6} + \underbrace{a_{21}b_{22}}_{3} + \underbrace{a_{22}b_{11}}_{7} - \underbrace{a_{22}b_{12}}_{8} + a_{22}b_{22} - \underbrace{a_{21}b_{11}}_{5} + \\
&+ \underbrace{a_{21}b_{12}}_{6} - \underbrace{a_{22}b_{11}}_{7} + \underbrace{a_{22}b_{12}}_{8} + \underbrace{a_{11}b_{11}}_{4} + a_{23}b_{32} \\
&= a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32},
\end{aligned}
$$

$$
\begin{aligned}
c_{23} &= m_{14} + m_{16} + m_{17} + m_{18} + m_{21} \\
&= a_{13}b_{31} + (-a_{13} + a_{22} + a_{23})(b_{23} + b_{31} - b_{33}) + (a_{13} - a_{23})(b_{23} - b_{33}) + \\
&+ (a_{22} + a_{23})(-b_{31} + b_{33}) + a_{21}b_{13} \\
&= \underbrace{a_{13}b_{31}}_{1} - \underbrace{a_{13}b_{23}}_{2} - \underbrace{a_{13}b_{31}}_{1} + \underbrace{a_{13}b_{33}}_{3} + a_{22}b_{23} + \underbrace{a_{22}b_{31}}_{4} - \underbrace{a_{22}b_{33}}_{5} + \\
&+ \underbrace{a_{23}b_{23}}_{6} + \underbrace{a_{23}b_{31}}_{7} - \underbrace{a_{23}b_{33}}_{8} + \underbrace{a_{13}b_{23}}_{2} - \underbrace{a_{13}b_{33}}_{3} - \underbrace{a_{23}b_{23}}_{6} + a_{23}b_{33} - \\
&- \underbrace{a_{22}b_{31}}_{4} + \underbrace{a_{22}b_{33}}_{5} - \underbrace{a_{23}b_{31}}_{7} + \underbrace{a_{23}b_{33}}_{8} + a_{21}b_{13} \\
&= a_{22}b_{23} + a_{23}b_{33} + a_{21}b_{13},
\end{aligned}
$$

$$
\begin{aligned}
c_{31} &= m_6 + m_7 + m_8 + m_{11} + m_{12} + m_{13} + m_{14} \\
&= a_{11}b_{11} + (-a_{11} + a_{31} + a_{32})(b_{11} - b_{13} + b_{23}) + (-a_{11} + a_{31})(b_{13} - b_{23}) + \\
&+ a_{32}(-b_{11} + b_{13} + b_{21} - b_{22} - b_{23} - b_{31} + b_{32}) + (-a_{13} + a_{32} + a_{33}) \cdot \\
&\cdot (b_{22} + b_{31} - b_{32}) + (a_{13} - a_{33})(b_{22} - b_{32}) + a_{13}b_{31} \\
&= \underbrace{a_{11}b_{11}}_{1} - \underbrace{a_{11}b_{11}}_{1} + \underbrace{a_{11}b_{13}}_{2} - \underbrace{a_{11}b_{23}}_{3} + a_{31}b_{11} - \underbrace{a_{31}b_{13}}_{4} + \underbrace{a_{31}b_{23}}_{5} + \\
&+ \underbrace{a_{32}b_{11}}_{6} - \underbrace{a_{32}b_{13}}_{7} + \underbrace{a_{32}b_{23}}_{8} - \underbrace{a_{11}b_{13}}_{2} + \underbrace{a_{11}b_{23}}_{3} + \underbrace{a_{31}b_{13}}_{4} - \underbrace{a_{31}b_{23}}_{5} - \\
&- \underbrace{a_{32}b_{11}}_{6} + \underbrace{a_{32}b_{13}}_{7} + a_{32}b_{21} - \underbrace{a_{32}b_{22}}_{9} - \underbrace{a_{32}b_{23}}_{8} - \underbrace{a_{32}b_{31}}_{10} + \underbrace{a_{32}b_{32}}_{11} - \\
&- \underbrace{a_{13}b_{22}}_{12} - \underbrace{a_{13}b_{31}}_{13} + \underbrace{a_{13}b_{32}}_{14} + \underbrace{a_{32}b_{22}}_{9} + \underbrace{a_{32}b_{31}}_{10} - \underbrace{a_{32}b_{32}}_{11} + \underbrace{a_{33}b_{22}}_{15} + \\
&+ a_{33}b_{31} - \underbrace{a_{33}b_{32}}_{16} + \underbrace{a_{13}b_{22}}_{12} - \underbrace{a_{13}b_{32}}_{14} - \underbrace{a_{33}b_{22}}_{15} + \underbrace{a_{33}b_{32}}_{16} + \underbrace{a_{13}b_{31}}_{13} \\
&= a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31}, \\
c_{32} &= m_{12} + m_{13} + m_{14} + m_{15} + m_{22} \\
&= (-a_{13} + a_{32} + a_{33})(b_{22} + b_{31} - b_{32}) + (a_{13} - a_{33})(b_{22} - b_{32}) + \\
&+ a_{13}b_{31} + (a_{32} + a_{33})(-b_{31} + b_{32}) + a_{31}b_{12} \\
&= \underbrace{-a_{13}b_{22}}_{1} - \underbrace{a_{13}b_{31}}_{2} + \underbrace{a_{13}b_{32}}_{3} + a_{32}b_{22} + \underbrace{a_{32}b_{31}}_{4} - \underbrace{a_{32}b_{32}}_{5} + \underbrace{a_{33}b_{22}}_{6} + \\
&+ \underbrace{a_{33}b_{31}}_{7} - \underbrace{a_{33}b_{32}}_{8} + \underbrace{a_{13}b_{22}}_{1} - \underbrace{a_{13}b_{32}}_{3} - \underbrace{a_{33}b_{22}}_{6} + \underbrace{a_{33}b_{32}}_{8} + \underbrace{a_{13}b_{31}}_{2} - \\
&- \underbrace{a_{32}b_{31}}_{4} + \underbrace{a_{32}b_{32}}_{5} - \underbrace{a_{33}b_{31}}_{7} + a_{33}b_{32} + a_{31}b_{12} \\
&= a_{32}b_{22} + a_{33}b_{32} + a_{31}b_{12},
\end{aligned}
$$

and finally

$$
\begin{aligned}
c_{33} &= m_6 + m_7 + m_8 + m_9 + m_{23} \\
&= a_{11}b_{11} + (-a_{11} + a_{31} + a_{32})(b_{11} - b_{13} + b_{23}) + (-a_{11} + a_{31})(b_{13} - b_{23}) + \\
&+ (-a_{11} + a_{31})(b_{13} - b_{23}) + (a_{31} + a_{32})(-b_{11} + b_{13}) + a_{33}b_{33}
\end{aligned}
$$

$$= \underbrace{a_{11}b_{11}}_{1} - \underbrace{a_{11}b_{11}}_{1} + \underbrace{a_{11}b_{13}}_{2} - \underbrace{a_{11}b_{23}}_{3} + \underbrace{a_{31}b_{11}}_{4} - \underbrace{a_{31}b_{13}}_{5} + \underbrace{a_{31}b_{23}}_{6} +$$

$$+ \underbrace{a_{32}b_{11}}_{7} - \underbrace{a_{32}b_{13}}_{8} + a_{32}b_{23} - \underbrace{a_{11}b_{13}}_{2} + \underbrace{a_{11}b_{23}}_{3} + a_{31}b_{13} - \underbrace{a_{31}b_{23}}_{6} -$$

$$- \underbrace{a_{31}b_{11}}_{4} + \underbrace{a_{31}b_{13}}_{5} - \underbrace{a_{32}b_{11}}_{7} + \underbrace{a_{32}b_{13}}_{8} + a_{33}b_{33}$$

$$= a_{32}b_{23} + a_{31}b_{13} + a_{33}b_{33}.$$

Laderman's Algorithm produces the same result as traditional matrix multiplication.

3.3 Comparing Algorithms

3.3.1 Traditional Method

The following is an example of multiplying two 3 x 3 matrices showing the traditional method, Strassen's Algorithm and Laderman's Algorithm. Consider

$$\begin{pmatrix} -2 & 0 & 1 \\ 3 & -1 & -3 \\ 1 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 & -1 \\ -2 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}.$$

Traditional methods yield the following results:

$$\begin{array}{lcccccccl}
c_{11} & = & (-2)(3) & + & (0)(-2) & + & (1)(1) & = & -5, \\
c_{12} & = & (-2)(-1) & + & (0)(0) & + & (1)(1) & = & 3, \\
c_{13} & = & (-2)(-1) & + & (0)(2) & + & (1)(0) & = & 2, \\
c_{21} & = & (3)(3) & + & (-1)(-2) & + & (-3)(1) & = & 8, \\
c_{22} & = & (3)(-1) & + & (-1)(0) & + & (-3)(1) & = & -6, \\
c_{23} & = & (3)(-1) & + & (-1)(2) & + & (-3)(0) & = & -5, \\
c_{31} & = & (1)(3) & + & (2)(-2) & + & (-1)(1) & = & -2, \\
c_{32} & = & (1)(-1) & + & (2)(0) & + & (-1)(1) & = & -2, \\
c_{33} & = & (1)(-1) & + & (2)(2) & + & (-1)(0) & = & 3.
\end{array}$$

Replacing each entry with its value gives

$$\begin{pmatrix} -2 & 0 & 1 \\ 3 & -1 & -3 \\ 1 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 & -1 \\ -2 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -5 & 3 & 2 \\ 8 & -6 & -5 \\ -2 & -5 & 3 \end{pmatrix}.$$

Using this method requires 27 multiplication steps and 18 multiplication steps for a total of 45 operations.

3.3.2 Strassen's Algorithm

To use Strassen's Algorithm, the matrices need to be expanded to 4 x 4 matrices by adding a row of zeros and a column of zeros to each matrix:

$$\left(\begin{array}{cc|cc} -2 & 0 & 1 & 0 \\ 3 & -1 & -3 & 0 \\ \hline 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \cdot \left(\begin{array}{cc|cc} 3 & -1 & -1 & 0 \\ -2 & 0 & 2 & 0 \\ \hline 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) = \left(\begin{array}{cc|cc} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ \hline c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{array}\right).$$

Then

$$\begin{aligned}
\mathrm{I} &= \left[\begin{pmatrix} -2 & 0 \\ 3 & -1 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} 3 & -1 \\ -2 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right] \\
&= \begin{pmatrix} -3 & 0 \\ 3 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 \\ -2 & 0 \end{pmatrix} \\
&= \begin{pmatrix} -9 & 3 \\ 11 & -3 \end{pmatrix}; \\
\mathrm{II} &= \left[\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} 3 & -1 \\ -2 & 0 \end{pmatrix}\right] \\
&= \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 \\ -2 & 0 \end{pmatrix} \\
&= \begin{pmatrix} -4 & 0 \\ 0 & 0 \end{pmatrix}; \\
\mathrm{III} &= \left[\begin{pmatrix} -2 & 0 \\ 3 & -1 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} -1 & 0 \\ 2 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \begin{pmatrix} -2 & 0 \\ 3 & -1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 2 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 0 \\ -5 & 0 \end{pmatrix}; \\
\mathrm{IV} &= \left[\begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} -3 & 1 \\ 2 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\right] \\
&= \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} -2 & 2 \\ 2 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 2 & -2 \\ 0 & 0 \end{pmatrix}; \\
\mathrm{V} &= \left[\begin{pmatrix} -2 & 0 \\ 3 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right] \\
&= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; \\
\mathrm{VI} &= \left[\begin{pmatrix} 2 & 0 \\ -3 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} 3 & -1 \\ -2 & 0 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 2 & 0 \end{pmatrix}\right] \\
&= \begin{pmatrix} 3 & 2 \\ -3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & -1 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 6 & -3 \\ -6 & 3 \end{pmatrix}; \\
\mathrm{VII} &= \left[\begin{pmatrix} 1 & 0 \\ -3 & 0 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right] \\
&= \begin{pmatrix} 2 & 0 \\ -3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}
\end{aligned}
$$

$$= \begin{pmatrix} 2 & 2 \\ -3 & -3 \end{pmatrix}$$

Now substitute the above matrices into the equations for C_{11}, C_{12}, C_{21}, and C_{22}.

$$\begin{aligned} C_{11} &= \quad \mathrm{I} \quad + \quad \mathrm{IV} \quad - \quad \mathrm{V} \quad + \quad \mathrm{VII} \\ &= \begin{pmatrix} -9 & 3 \\ 11 & -3 \end{pmatrix} + \begin{pmatrix} 2 & -2 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ -3 & -3 \end{pmatrix} = \begin{pmatrix} -5 & 3 \\ 8 & -6 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} C_{12} &= \quad \mathrm{III} \quad + \quad \mathrm{V} \\ &= \begin{pmatrix} 2 & 0 \\ -5 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ -5 & 0 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} C_{21} &= \quad \mathrm{II} \quad + \quad \mathrm{IV} \\ &= \begin{pmatrix} -4 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 & -2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -2 & -2 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} C_{22} &= \quad \mathrm{I} \quad + \quad \mathrm{III} \quad - \quad \mathrm{II} \quad + \quad \mathrm{VI} \\ &= \begin{pmatrix} -9 & 3 \\ 11 & -3 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ -5 & 0 \end{pmatrix} - \begin{pmatrix} -4 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 6 & -3 \\ -6 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Now using the above values, the product results in the following 4 x 4 matrix.

$$\begin{pmatrix} -5 & 3 & 2 & 0 \\ 8 & -6 & -5 & 0 \\ -2 & -2 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Removing the superfluous row and column of zeros, the product of the original multiplication problem is found.

$$\begin{pmatrix} -2 & 0 & 1 \\ 3 & -1 & -3 \\ 1 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 & -1 \\ -2 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -5 & 3 & 2 \\ 8 & -6 & -5 \\ -2 & -2 & 3 \end{pmatrix}$$

Strassen's Algorithm has 49 multiplication steps and 198 addition steps for a total of 247 operations. It would seem as if we could reduce the number of steps needed because of the extra zero row and column. The example above does contain a zero submatrix and some steps appear unnecessary. However, this would not always be the case. Our example just happens to have a zero in the third row, third column of one of the original factors, causing this to happen. While we could predict some steps that would involve either adding zero or mutliplying by zero, we would still need to keep the submatrices of correct dimensions in order to perform the operations and those steps would be included.

3.3.3 Laderman's Algorithm

Now Laderman's Algorithm is presented. Again consider

$$\begin{pmatrix} -2 & 0 & 1 \\ 3 & -1 & -3 \\ 1 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 & -1 \\ -2 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}.$$

Then

$$\begin{aligned}
m_1 &= (a_{11} + a_{12} + a_{13} - a_{21} - a_{22} - a_{32} - a_{33}) \cdot b_{22} \\
&= [-2 + 0 + 1 - 3 - (-1) - 2 - (-1)] \cdot 0 &&= \quad 0, \\
m_2 &= (a_{11} - a_{21}) \cdot (-b_{12} + b_{22}) \\
&= (-2 - 3) \cdot (1 + 0) &&= \quad -5, \\
m_3 &= a_{22} (-b_{11} + b_{12} + b_{21} - b_{22} - b_{23} - b_{31} + b_{33}) \\
&= (-1) \cdot [-3 + (-1) + (-2) - 0 - 2 - 1 + 0 &&= \quad 9, \\
m_4 &= (-a_{11} + a_{21} + a_{22}) \cdot (b_{11} - b_{12} + b_{22}) \\
&= (2 + 3 - 1) \cdot (3 + 1 + 0) &&= \quad 16, \\
m_5 &= (a_{21} + a_{22}) \cdot (-b_{11} + b_{12}) \\
&= [3 + (-1)] \cdot [-3 + (-1)] &&= \quad -8, \\
m_6 &= a_{11} b_{11} \\
&= (-2) \cdot (3) &&= \quad -6, \\
m_7 &= (-a_{11} + a_{31} + a_{32}) \cdot (b_{11} - b_{13} + b_{23}) \\
&= (2 + 1 + 2) \cdot [3 - (-1) + 2] &&= \quad 30,
\end{aligned}$$

$$\begin{array}{lclcr}
m_8 & = & (-a_{11} + a_{31}) \cdot (b_{13} - b_{23}) & & \\
 & = & (2+1) \cdot [(-1) - 2] & = & -9, \\
m_9 & = & (a_{31} + a_{32}) \cdot (-b_{11} + b_{13}) & & \\
 & = & (1+2) \cdot [(-3) + (-1)] & = & -12, \\
m_{10} & = & (a_{11} + a_{12} + a_{13} - a_{22} - a_{23} - a_{31} - a_{32}) \cdot b_{23} & & \\
 & = & [-2 + 0 + 1 - (-1) - (-3) - 1 - 2] \cdot (2) & = & 0, \\
m_{11} & = & a_{32} (-b_{11} + b_{13} + b_{21} - b_{22} - b_{23} - b_{31} + b_{32}) & & \\
 & = & 2 \cdot [-3 + (-1) + (-2) - 0 - 2 - 1 + 1] & = & -16, \\
m_{12} & = & (-a_{13} + a_{32} + a_{33}) \cdot (b_{22} + b_{31} - b_{32}) & & \\
 & = & [-1 + 2 + (-1)] \cdot (0 + 1 - 1) & = & 0, \\
m_{13} & = & (a_{13} - a_{33}) \cdot (b_{22} - b_{32}) & & \\
 & = & [1 - (-1)] \cdot (0 - 1) & = & -2, \\
m_{14} & = & a_{13} b_{31} & & \\
 & = & (1) \cdot (1) & = & 1, \\
m_{15} & = & (a_{32} + a_{33}) \cdot (-b_{31} + b_{32}) & & \\
 & = & [2 + (-1)] \cdot (-1 + 1) & = & 0, \\
m_{16} & = & (-a_{13} + a_{22} + a_{23}) \cdot (b_{23} + b_{31} - b_{33}) & & \\
 & = & [-1 + (-1) = (-3)] \cdot (2 + 1 - 0) & = & -15, \\
m_{17} & = & (a_{13} - a_{23}) \cdot (b_{23} - b_{33}) & & \\
 & = & [1 - (-3)] \cdot (2 - 0) & = & 8, \\
m_{18} & = & (a_{22} + a_{23}) \cdot (-b_{31} + b_{33}) & & \\
 & = & [-1 + (-3)] \cdot (-1 + 0) & = & 4, \\
m_{19} & = & a_{12} b_{21} & & \\
 & = & (0) \cdot (-2) & = & 0, \\
m_{20} & = & a_{23} b_{32} & & \\
 & = & (-3) \cdot (1) & = & -3, \\
m_{21} & = & a_{21} b_{13} & & \\
 & = & (3) \cdot (-1) & = & -3, \\
m_{22} & = & a_{31} b_{12} & & \\
 & = & (1) \cdot (-1) & = & -1,
\end{array}$$

$$\begin{aligned} m_{23} &= a_{33}b_{33} \\ &= (-1)\cdot(0) = 0. \end{aligned}$$

Calculating the individual entry values has the following result:

$$\begin{aligned}
c_{11} &= m_6 + m_{14} + m_{19} \\
&= -6 + 1 + 0 &&= -5, \\
c_{12} &= m_1 + m_4 + m_5 + m_6 + m_{12} + m_{14} + m_{15} \\
&= 0 + 16 + (-8) + (-6) + 0 + 1 + 0 &&= 3, \\
c_{13} &= m_6 + m_7 + m_9 + m_{10} + m_{14} + m_{16} + m_{18} \\
&= (-6) + 30 + (-12) + 0 + 1 + (-15) + 4 &&= 2, \\
c_{21} &= m_2 + m_3 + m_4 + m_6 + m_{14} + m_{16} + m_{17} \\
&= (-5) + 9 + 16 + (-6) + 1 + (-15) + 8 &&= 8, \\
c_{22} &= m_2 + m_4 + m_5 + m_6 + m_{20} \\
&= (-5) + 16 + (-8) + (-6) + (-3) &&= -6, \\
c_{23} &= m_{14} + m_{16} + m_{17} + m_{18} + m_{21} \\
&= 1 + (-15) + 8 + 4 + (-3) &&= -5, \\
c_{31} &= m_6 + m_7 + m_8 + m_{11} + m_{12} + m_{13} + m_{14} \\
&= (-6) + 30 + (-9) + (-16) + 0 + (-2) + 1 &&= -2, \\
c_{32} &= m_{12} + m_{13} + m_{14} + m_{15} + m_{22} \\
&= 0 + (-2) + 1 + 0 + (-1) &&= -2, \\
c_{33} &= m_6 + m_7 + m_8 + m_9 + m_{23} \\
&= (-6) + 30 + (-9) + (-12) + 0 &&= 3.
\end{aligned}$$

Finally, the result is the same as using traditional methods or Strassen's Algorithm:

$$\begin{pmatrix} -2 & 0 & 1 \\ 3 & -1 & -3 \\ 1 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 & -1 \\ -2 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -5 & 3 & 2 \\ 8 & -6 & -5 \\ -2 & -2 & 3 \end{pmatrix}.$$

Laderman's method uses 23 multiplication steps along with 98 addition steps for a total of 121 mathematical operations. Table 3.1 shows each method and the amount of work involved.

Table 3.1: Operations for Traditional, Strassen, and Laderman

Traditional Method			Strassen's Algorithm			Laderman's Algorithm		
mult	add	total	mult	add	total	mult	add	total
27	18	45	49	198	247	23	98	121

3.4 Work Needed to Complete Laderman's Algorithm

We now consider the recursive formula for the amount of work needed to use Laderman's Algorithm. Consider multiplying two n x n matrices, where $n = 3^k$. Similar to partitioning matrices for Strassen's Algorithm, we partition each matrix into 9 matrices each with dimensions $\frac{n}{3}$ x $\frac{n}{3}$. To apply Laderman's Algorithm, we will need to multiply two $\frac{n}{3}$ x $\frac{n}{3}$ matrices 23 times. We will also have 98 addition steps where two $\frac{n}{3}$ x $\frac{n}{3}$ matrices are being added. Each of the 98 addition step will have $\left(\frac{n}{3}\right)^2$ steps. The formula for the amount of work needed to multiply two n x n matrices using Laderman's algorithm would be

$$T(n) = 23T\left(\frac{n}{3}\right) + 98\left(\frac{n}{3}\right)^2.$$

From previous discussion, we are most interested in knowing the number of multiplication steps necessary in multiplying matrices. Laderman's Algorithm can be used to multiply 3^k x 3^k matrices using 23^k multiplications. Similar to using Strassen's Algorithm, we see that using Laderman's Algorithm to multiply two n x n matrices, where $n = 3^k$, the number of multilplication steps is represented by $23^{\lceil \log_3 n \rceil}$. Through principles of logarithms, we see that $23^{\log_3 n} = n^{\log_3 23}$ and $\log_3 23 \approx 2.85$. Thus, $n^{\log_3 23} < n^3$.

3.5 So Which Algorithm Is Better?

In the previously mentioned example, Laderman's Algorithm does a better job of reducing the amount of work needed compared to Strassen's Algorithm for multiplying two 3 x 3 matrices. This may seem unfair to Strassen's Algorithm because we are only considering an example of multiplying 3 x 3 matrices. However, that is exactly what Laderman was trying to simplify. Consider multiplying two 9 x 9 matrices. To use Strassen's Algorithm, the matrices would need to be expanded to the next power of 2 which would be 16 x 16. Using the formula from Section 2.4, this would amount to 2401 multiplication steps. However, if Laderman's Algorithm is used the matrices are already a power of three and would only need to be divided into nine submatrices before applying the algorithm. This would amount to 529 multiplication steps. It appears that Laderman's Algorithm would be beneficial when working with 3^k x 3^k matrices.

Table 3.2: Comparison of Methods

Size of Matrix	Traditional		Strassen		Laderman	
m x m	n	n^3	n	$7^{\lceil \log_2 n \rceil}$	n	$23^{\lceil \log_3 n \rceil}$
2 x 2	2	8	2	7	–	–
3 x 3	3	27	4	49	3	21
4 x 4	4	64	4	49	9	529
5 x 5	5	125	8	343	9	529
6 x 6	6	216	8	343	9	529
7 x 7	7	343	8	343	9	529
8 x 8	8	512	8	343	9	529
9 x 9	9	729	16	2401	9	529
10 x 10	10	1000	16	2401	27	12,167
11 x 11	11	1331	16	2401	27	12,167
12 x 12	12	1728	16	2401	27	12,167
13 x 13	13	2197	16	2401	27	12,167
14 x 14	14	2744	16	2401	27	12,167
15 x 15	15	3375	16	2401	27	12,167
16 x 16	16	4096	16	2401	27	12,167
27 x 27	27	19,683	32	16,807	27	12,167
32 x 32	32	32,768	32	16,807	81	279,841
64 x 64	64	262,144	64	117,649	81	279,841
81 x 81	81	531,441	128	823,543	81	279,841

We now compare the three methods discussed and the amount of work involved. Of main importance will be the number of multiplications used as matrices get larger.

Table 3.2 shows the number of multiplication steps needed to multiply two m x m matrices using each of the three methods. Note that n represents the number of rows for the expanded matrix if one is needed to create either a 2^k x 2^k or 3^k x 3^k matrix. According to Table 3.2, Laderman's Algorithm saves a significant number of multiplication steps over both the traditional method and Strassen's Algorithm when the matrices are 3^k x 3^k. As n increases, eventually Strassen's Algorithm reduces multiplication steps over traditional methods and Laderman's Algorithm.

Chapter 4

Inversion of Matrices

4.1 Finding an Inverse by Matrix Multiplication

For the beginning algebra student, finding the inverse of a matrix can often be a grueling process. Here we see that the work of finding an inverse is not worse than the work needed to multiply matrices. While this may not be a comfort to the beginning algebra student, we can see that it may be beneficial to use matrix multiplication when turning the work over to a computer.

Since much of the work in this section will involve inverses of matrices, inverses will be denoted by A^{-1} while entries of matrices will still be denoted by A_{ij}, where i represents the row of the element and j represents the column of the entry.

In the book, *The Design and Analysis of Computer Algorithms* [AHU74], the authors show how to find the inverse of a matrix through matrix multiplication. The method that will be shown applies to all triangular matrices that are nonsingular and invertible. To study this concept, some terms need definition.

Definition 1. An m x n matrix A is *upper triangular* if $A_{ij} = 0$ whenever $1 \leq j < i \leq m$. An m x n matrix A is *lower triangular* if $A_{ij} = 0$ whenever $1 \leq i < j \leq n$. [AHU74]

In other words, for a matrix to be upper triangular means all entries below the main diagonal are zero. Likewise, for a matrix to be lower triangular means all entries above the main diagonal are zero.

The following are examples of triangular matrices.

$$\underbrace{\begin{pmatrix} a_{11} & 0_{12} & \dots & 0_{1m} & \dots & 0_{1n} \\ a_{21} & a_{22} & \dots & 0_{2m} & \dots & 0_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} & \dots & 0_{mn} \end{pmatrix} \text{ and } \begin{pmatrix} a_{11} & 0_{12} & \dots & 0_{1m} \\ a_{21} & a_{22} & \dots & 0_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix}}_{\text{Lower Triangular Matrices}};$$

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} & \dots & a_{1n} \\ 0_{21} & a_{22} & \dots & a_{2m} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0_{m1} & 0_{m2} & \dots & a_{mm} & \dots & a_{mn} \end{pmatrix} \text{ and } \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ 0_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{m1} & 0_{m2} & \dots & a_{mm} \end{pmatrix}}_{\text{Upper Triangular Matrices}}.$$

Definition 2. If A is a square upper or lower triangular matrix, then A is *nonsingular* if and only if no entry on the main diagonal is zero. [AHU74]

The following lemma is presented in *The Design and Analysis of Computer Algorithms*[AHU74].

Lemma 3. *: Let A be partitioned as*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

Suppose A_{11}^{-1} exists. Define $\Delta = A_{22} - A_{21}A_{11}^{-1}A_{12}$ and assume Δ^{-1} exists. Then

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}\Delta^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}\Delta^{-1} \\ -\Delta^{-1}A_{21}A_{11}^{-1} & \Delta^{-1} \end{pmatrix}.$$

To use this method of finding the inverse of a matrix, the matrix must be triangular and invertible. Now, consider how this lemma is developed. We first rewrite A as follows

$$\mathrm{A} = \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 \\ 0 & \Delta \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}.$$

To see why A can be represented this way, we use the fact that matrix multiplication is associative. Multiplying the last two matrices produces

$$A = \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ 0 & \Delta \end{pmatrix}.$$

Multiplying these matrices, yields the original representation of matrix A,.

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

Therefore,

$$\text{A} = \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 \\ 0 & \Delta \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}.$$

Why should we write matrix A as the product of three matrices? Where do the matrices come from?

Suppose we want to write A as the product of a lower triangular matrix and an upper triangular matrix where each entry on the diagonal of the lower triangular matrix is the identity. Let B, C, D, and E represent entries such that

$$\begin{pmatrix} I & 0 \\ B & I \end{pmatrix} \begin{pmatrix} C & D \\ 0 & E \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

Through matrix multiplication, the following statements can be made.

1.

$$C = A_{11}$$

2.

$$D = A_{12}$$

3.

$$\begin{aligned} BC &= A_{21} \\ B &= A_{21}C^{-1} \\ B &= A_{21}A_{11}^{-1} \end{aligned}$$

4.

$$\begin{aligned} BD + E &= A_{22} \\ E &= A_{22} - BD \\ E &= A_{22} - A_{21}A_{11}^{-1}A_{12} \\ E &= \Delta \end{aligned}$$

Therefore,

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ 0 & \Delta \end{pmatrix}.$$

Now suppose we want $\begin{pmatrix} A_{11} & A_{12} \\ 0 & \Delta \end{pmatrix}$ to be the product of a matrix that is both upper and lower triangular and a matrix that is upper triangular and the diagonal entries are the identity. Let

$$\begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix} \begin{pmatrix} I & D \\ 0 & I \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & \Delta \end{pmatrix}.$$

Again through matrix multiplication and matching entries, the following equations hold.

1.

$$B = A_{11}$$

2.

$$\begin{aligned} BD &= A_{12} \\ D &= B^{-1}A_{12} \\ D &= A_{11}^{-1}A_{12} \end{aligned}$$

3.

$$C = \Delta$$

This yields the following equation,

$$\begin{pmatrix} A_{11} & 0 \\ 0 & \Delta \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & \Delta \end{pmatrix}.$$

This brings us to the equation we were looking for that is a product of three triangular matrices.

$$A = \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 \\ 0 & \Delta \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}$$

Now A is represented with a *unit* lower triangular matrix, meaning the entries on the main diagonal are identities, and a unit upper triangular matrix. These matrices will be easier to work with.

To develop a representation of A^{-1} we start with the representation of A shown above. A is written as the product of three matrices that are either lower triangular, upper triangular, or both. The two matrices that are individually lower triangular and upper triangular both have diagonals with elements consisting of the identity. By breaking up the original matrix into the product of these simpler matrices we will see the benefit of finding the inverse by finding the inverses of the simpler matrices.

Consider finding the inverse of a matrix that is represented by the product of three matrices. Let $A = BCD$ where B, C, D are matrices. Then,

$$A^{-1} = (BCD)^{-1}.$$

Since matrix multiplication is associative,

$$A^{-1} = [B\,(CD)]^{-1}.$$

We know $(FG)^{-1} = G^{-1}F^{-1}$. Applying this fact to find A^{-1} we get

$$\begin{aligned} A^{-1} &= [B\,(CD)]^{-1} \\ &= (CD)^{-1}\,B^{-1} \\ &= D^{-1}C^{-1}B^{-1}. \end{aligned}$$

Since A can be represented as the product of three matrices and we have the representation of the inverse of the product of three matrices, we can apply this knowledge to find the inverse of A.

We now look to find A^{-1}. Since

$$A = \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 \\ 0 & \Delta \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix},$$

then

$$\begin{aligned} A^{-1} &= \left[\begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 \\ 0 & \Delta \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} \right]^{-1} \\ &= \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}^{-1} \begin{pmatrix} A_{11} & 0 \\ 0 & \Delta \end{pmatrix}^{-1} \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix}^{-1}. \end{aligned}$$

Looking at the first inverted matrix, let

$$\begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} = D$$

and

$$\begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}^{-1} = D^{-1} = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}.$$

Using the fact that

$$DD^{-1} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix},$$

we must find matrix D^{-1} such that

$$\begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

Using matrix multiplication, the following four equations are obtained:

$$\begin{aligned} D_{11} + A_{11}^{-1}A_{12}D_{21} &= I \\ D_{12} + A_{11}^{-1}A_{12}D_{22} &= 0 \\ D_{21} &= 0 \\ D_{22} &= I. \end{aligned}$$

Now analyze equations 1 and 2.

1.

$$\begin{aligned} D_{11} + A_{11}^{-1}A_{12}D_{21} &= I \\ D_{11} + A_{11}^{-1}A_{12} \cdot 0 &= I \\ D_{11} &= I \end{aligned}$$

2.

$$\begin{aligned} D_{12} + A_{11}^{-1}A_{12}D_{22} &= 0 \\ D_{12} + A_{11}^{-1}A_{12} \cdot I &= 0 \\ D_{12} + A_{11}^{-1}A_{12} &= 0 \\ D_{12} &= -A_{11}^{-1}A_{12} \end{aligned}$$

Thus,

$$D^{-1} = \begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}.$$

We now know the first inverse in the representation of A^{-1}. Now we will show how the second inverse is derived.

Let

$$C = \begin{pmatrix} A_{11} & 0 \\ 0 & \Delta \end{pmatrix}$$

and

$$\begin{pmatrix} A_{11} & 0 \\ 0 & \Delta \end{pmatrix}^{-1} = C^{-1} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}.$$

Then

$$\begin{pmatrix} A_{11} & 0 \\ 0 & \Delta \end{pmatrix} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

This implies the following four equations:

1.

$$\begin{aligned} A_{11}C_{11} &= I \\ C_{11} &= A_{11}^{-1} \end{aligned}$$

2.

$$\begin{aligned} A_{11}C_{12} &= 0 \\ C_{12} &= 0 \end{aligned}$$

3.

$$\begin{aligned} \Delta C_{21} &= 0 \\ C_{21} &= 0 \end{aligned}$$

4.

$$\begin{aligned} \Delta C_{22} &= I \\ C_{22} &= \Delta^{-1}. \end{aligned}$$

Therefore,

$$C^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & \Delta^{-1} \end{pmatrix}.$$

Finally, let

$$B = \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix}$$

and

$$\begin{pmatrix} I & 0 \\ A_2 1 A_{11}^{-1} & I \end{pmatrix}^{-1} = B^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$

Then,

$$\begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

This gives the next four equations.

1.
$$B_{11} = I$$

2.
$$B_{12} = 0$$

3.
$$\begin{aligned} A_{21}A_{11}^{-1}B_{11} + B_{21} &= 0 \\ A_{21}A_{11}^{-1} + B_{21} &= 0 \\ B_{21} &= -A_{21}A_{11}^{-1} \end{aligned}$$

4.
$$\begin{aligned} A_{21}A_{11}^{-1}B_{12} + B_{22} &= I \\ B_{22} &= I \end{aligned}$$

Now we have obtained the inverse of the last matrix

$$B^{-1} = \begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}.$$

By finding the inverse for each of the matrices and substituting, we can obtain the equation for the inverse of A. We can then simplify the equation into one matrix. So

$$\begin{aligned} A^{-1} &= \begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & \Delta^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix} \\ &= \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}\Delta^{-1} \\ 0 & \Delta^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix} \\ &= \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}\Delta^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}\Delta^{-1} \\ -\Delta^{-1}A_{21}A_{11}^{-1} & \Delta^{-1} \end{pmatrix}. \end{aligned}$$

This lemma for finding an inverse does not apply to all nonsingular matrices. For example, the matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

has $\det(A) \neq 0$. However, partitioning the matrix into four submatrices gives $A_{11} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, meaning A_{11}^{-1} does not exist. The lemma does apply to all nonsingular triangular (either upper triangular or lower triangular) matrices.

What if the matrix you are trying to invert does not have the number of rows and columns equal to a power of 2? Can you still use Lemma 1? Suppose A is an n x n, nonsingular, invertible, triangular matrix but n is not a power of 2. Then A can be placed in a matrix of the form

$$\begin{pmatrix} A & 0 \\ 0 & I_m \end{pmatrix},$$

where I_m represents an m x m identity matrix such that $m + n \leq 2n$ and $m + n = 2^k$. This keeps the matrix triangular, invertible and nonsingular. Therefore, Lemma 3 still applies.

4.2 Another Representation of Finding an Inverse by Matrix Multiplication

When looking at the multiplication steps involved in finding an inverse, it will be helpful to use a more practical version of Lemma 1. In the article "A Strassen-Newton Algorithm for High-Speed Parallelizable Matrix Inversion" by David H. Baily and Helaman R. P. Ferguson[BF88], Lemma 1 is presented in a different form. By substitution, we can see that it is in fact the same method of finding an inverse as Lemma 1. However, by isolating individual multiplication steps, we can avoid repeating the same multiplications. Their method showing the six multiplication steps at P_2, P_3, P_4, C_{12}, C_{21}, and C_{11} is presented below.

Lemma 4. *Let*

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}.$$

Then,

$$\begin{aligned} P_1 &= A_{11}^{-1}, \\ P_2 &= A_{21}P_1, \\ P_3 &= P_1A_{12}, \\ P_4 &= A_{21}P_3, \\ P_5 &= P_4 - A_{22}, \\ P_6 &= P_5^{-1}, \end{aligned}$$

$$\begin{aligned} C_{12} &= P_3P_6, \\ C_{21} &= P_6P_2, \\ C_{11} &= P_1 - P_3C_{21}, \end{aligned}$$

and

$$C_{22} \quad = \quad -P_6.$$

This method also involves two inverse steps at P_1 and P_6, subtraction steps at P_5 and C_{11} and a negation step at C_{22}.

4.3 An Example of Finding an Inverse

The following example demonstrates the use of the above steps. Let

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 2 & 1 & 5 & -3 \\ 0 & -1 & 3 & 0 \\ 1 & 0 & 2 & 0 \end{pmatrix}.$$

Then

$$\begin{aligned}
P_1 &= A_{11}^{-1} &&= \begin{pmatrix} 1 & 0 \\ -2 & 0 \end{pmatrix}, \\
P_2 &= A_{21}P_1 &&= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \\
& &&= \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \\
P_3 &= P_1A_{12} &&= \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 5 & -3 \end{pmatrix} \\
& &&= \begin{pmatrix} 0 & 1 \\ 5 & -5 \end{pmatrix},
\end{aligned}$$

$$\begin{aligned}
P_4 &= A_{21}P_3 &&= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 5 & -5 \end{pmatrix} \\
& &&= \begin{pmatrix} -5 & 5 \\ 0 & 1 \end{pmatrix}, \\
P_5 &= P_4 - A_{22} &&= \begin{pmatrix} -5 & 5 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 0 \\ 2 & 0 \end{pmatrix} \\
& &&= \begin{pmatrix} -8 & 5 \\ -2 & 1 \end{pmatrix}, \\
P_6 &= P_5^{-1} &&= \begin{pmatrix} \frac{1}{2} & -\frac{5}{2} \\ 1 & -4 \end{pmatrix},
\end{aligned}$$

$$\begin{aligned}
C_{12} &= P_3P_6 &&= \begin{pmatrix} 0 & 1 \\ 5 & -5 \end{pmatrix}\begin{pmatrix} \frac{1}{2} & -\frac{5}{2} \\ 1 & -4 \end{pmatrix} \\
& &&= \begin{pmatrix} 1 & -4 \\ -\frac{5}{2} & \frac{15}{2} \end{pmatrix}, \\
C_{21} &= P_6P_2 &&= \begin{pmatrix} \frac{1}{2} & -\frac{5}{2} \\ 1 & -4 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \\
& &&= \begin{pmatrix} -\frac{3}{2} & -\frac{1}{2} \\ -2 & -1 \end{pmatrix}, \\
C_{11} &= P_1 - P_3C_{21} &&= \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 5 & -5 \end{pmatrix}\begin{pmatrix} -\frac{3}{2} & -\frac{1}{2} \\ -2 & -1 \end{pmatrix}
\end{aligned}$$

$$= \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} - \begin{pmatrix} -2 & -1 \\ \frac{5}{2} & \frac{5}{2} \end{pmatrix}$$

$$= \begin{pmatrix} 3 & 1 \\ -\frac{9}{2} & -\frac{3}{2} \end{pmatrix},$$

$$C_{22} = -P_6 = -\begin{pmatrix} \frac{1}{2} & -\frac{5}{2} \\ 1 & -4 \end{pmatrix}$$

$$C_{22} = \begin{pmatrix} -\frac{1}{2} & \frac{5}{2} \\ -1 & 4 \end{pmatrix}$$

Hence,

$$A^{-1} = \begin{pmatrix} 3 & 1 & 1 & -4 \\ -\frac{9}{2} & -\frac{3}{2} & -\frac{5}{2} & \frac{15}{2} \\ -\frac{3}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{5}{2} \\ -2 & -1 & -1 & 4 \end{pmatrix}.$$

The inverse is found using six matrix multiplication steps. The reduction of work to multiply matrices would therefore be beneficial, especially when working with matrices of larger dimensions.

4.4 Work Needed to Find an Inverse

We now consider how much work is involved in finding an inverse. Suppose A is an n x n matrix, where n is a power of 2. Then A can be split into four $\frac{n}{2}$ x $\frac{n}{2}$ submatrices representing A_{11}, A_{12}, A_{21}, and A_{22}. That is,

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

We refer to Lemma 3. So

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}\Delta^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}\Delta^{-1} \\ -\Delta^{-1}A_{21}A_{11}^{-1} & \Delta^{-1} \end{pmatrix}.$$

Since A is triangular, suppose it to be upper triangular. Then $A_{21} = 0$. This means the upper left entry equals A_{11}^{-1} and the lower left entry equals an $\frac{n}{2}$ x $\frac{n}{2}$ zero matrix. Since $\Delta = A_{22} - A_{21}A_{11}^{-1}A_{12}$ and $A_{21} = 0$, then $\Delta = A_{22}$. We now have

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{pmatrix}.$$

Therefore, we need to find the inverses of two $\frac{n}{2}$ x $\frac{n}{2}$ matrices, namely A_{11} and A_{22}. If $T(n)$ represents the amount of work needed to find the inverse of an n x n matrix, then $2T\left(\frac{n}{2}\right)$ represents the amount of work needed to find the two inverses. We also still need to perform two matrix multiplications and negate the term to obtain the upper right entry of the inverse.

Let $M(n)$ be the amount of worked required to multiply two n x n matrices. Then the work for finding our two matrix multiplications would be $2M\left(\frac{n}{2}\right)$. To negate this expression would take at most $\frac{n^2}{4}$ multiplications since that would be how many terms are in the resulting $\frac{n}{2}$ x $\frac{n}{2}$ matrix. The work needed to find the upper right term, after finding A_{11}^{-1} and A_{22}^{-1}, would be $2M\left(\frac{n}{2}\right) + \frac{n^2}{4}$. Now, suppose we multiply two $\frac{n}{2}$ x $\frac{n}{2}$ matrices where one of the matrices is an identity matrix. This would be the simplest nonzero multiplication that could happen and it would take at least $\frac{n^2}{4}$ multiplications because there are at least that many entries in the product. Therefore, the amount of work needed to multiply two $\frac{n}{2}$ x $\frac{n}{2}$ must be greater than or equal to the amount of work needed if one of the matrices is an identity matrix. We say $\frac{n^2}{4} \leq M\left(\frac{n}{2}\right)$. Therefore, $2M\left(\frac{n}{2}\right) + \frac{n^2}{4} \leq 3M\left(\frac{n}{2}\right)$.

Hence, the total amount of work needed to find an inverse, $T(n)$, would satisfy

$$T(n) \leq 2T\left(\frac{n}{2}\right) + 3M\left(\frac{n}{2}\right), \text{for } n \geq 2,$$

where $T(1) = 1$.

We now try to get our statement in terms of the amount of work needed to multiply two matrices, $M(n)$.

Looking at the inequality for $T(n)$, we replace n with $\frac{n}{2}$. Then

$$\begin{aligned}
T(n) &\le 2T\left(\frac{n}{2}\right) + 3M\left(\frac{n}{2}\right) \\
&\le 2\left[2T\left(\frac{n}{4}\right) + 3M\left(\frac{n}{4}\right)\right] + 3M\left(\frac{n}{2}\right) \\
&= 3\left[M\left(\frac{n}{2}\right) + 2M\left(\frac{n}{4}\right)\right] + 4T\left(\frac{n}{4}\right) \\
&\le 3\left[M\left(\frac{n}{2}\right) + 2M\left(\frac{n}{4}\right)\right] + 4\left[2T\left(\frac{n}{8}\right) + 3M\left(\frac{n}{8}\right)\right]
\end{aligned}$$

since $T\left(\frac{n}{4}\right) \le 2T\left(\frac{n}{8}\right) + 3M\left(\frac{n}{8}\right)$.

Continuing this idea until we have $T(1)$ gives

$$\begin{aligned}
T(n) &\le 3\left[M\left(\frac{n}{2}\right) + 2M\left(\frac{n}{4}\right) + 4M\left(\frac{n}{8}\right) + \ldots + \left(\frac{n}{2}\right)M(1)\right] + nT(1) \\
&= 3\left[\sum_{k=1}^{\log_2 n} 2^{k-1}M\left(\frac{n}{2}^k\right)\right] + n \\
&= 3\left[\sum_{k=1}^{\log_2 n} 2^k \cdot 2M\left(\frac{n}{2}^k\right)\right] + n \\
&= \frac{3}{2}\left[\sum_{k=1}^{\log_2 n} 2^k M\left(\frac{n}{2^k}\right)\right] + n \\
&= \frac{3}{2}\left[\sum_{k=1}^{\log_2 n} 2^k M\left(\frac{n}{2^k}\right)\right] + n \\
&= \frac{3}{2}\left[2M\left(\frac{n}{2}\right) + 4M\left(\frac{n}{4}\right) + 8M\left(\frac{n}{8}\right) + \ldots + nM\left(\frac{n}{n}\right)\right] + n.
\end{aligned}$$

Suppose you want to multiply two $2m$ x $2m$ matrices and you partition each matrix into four m x m submatrices. If you used traditional methods of multiplication, as stated in section 2, the amount of work would be

$$\begin{aligned}
T(2m) &= [2\,(M(m))]^3 - [M(m)]^2 \\
&\le 8M(m).
\end{aligned}$$

If one of the two matrices were an identity matrix, then the least amount of work would be to muliply each m x m submatrix by the m x m identity matrix, meaning $4M(m)$. We now make a reasonable assumption that

$$4M(m) \leq M(2m).$$

Combining these two statements, we see

$$4M(m) \leq M(2m) \leq 8M(m).$$

Applying to our situation letting $m = \frac{n}{2}$ gives

$$4M\left(\frac{n}{2}\right) \leq M(n) \leq 8M\left(\frac{n}{2}\right)$$

$$M\left(\frac{n}{2}\right) \leq \frac{1}{4}M(n) \leq 2M\left(\frac{n}{2}\right).$$

Now let $m = \frac{n}{4}$ and get

$$4M\left(\frac{n}{4}\right) \leq M\left(\frac{n}{2}\right) \leq 8M\left(\frac{n}{4}\right)$$

$$M\left(\frac{n}{4}\right) \leq \frac{1}{4}M\left(\frac{n}{2}\right) \leq 2M\left(\frac{n}{4}\right).$$

Since $4M\left(\frac{n}{4}\right) \leq \frac{1}{4}M\left(\frac{n}{2}\right)$ and $M\left(\frac{n}{2}\right) \leq \frac{1}{4}M(n)$, then

$$\begin{aligned} M\left(\frac{n}{4}\right) &\leq \frac{1}{4} \cdot \frac{1}{4}M(n) \\ M\left(\frac{n}{4}\right) &\leq \frac{1}{16}M(n). \end{aligned}$$

By the same process, now let $m = \frac{n}{8}$ and we see that

$$4M\left(\frac{n}{8}\right) \leq M\left(\frac{n}{4}\right) \leq 8M\left(\frac{n}{8}\right)$$

$$M\left(\frac{n}{8}\right) \leq \frac{1}{4}M\left(\frac{n}{4}\right) \leq 2M\left(\frac{n}{8}\right).$$

Now we can see that

$$M\left(\frac{n}{8}\right) \leq \frac{1}{4} \cdot \frac{1}{16} M(n)$$

$$M\left(\frac{n}{8}\right) \leq \frac{1}{64} M(n).$$

We substitute the inequality statements into our statement of work for an infinite number of terms getting

$$\begin{aligned}
T(n) &\leq \frac{3}{2}\left[2M\left(\frac{n}{2}\right) + 4M\left(\frac{n}{4}\right) + 8M\left(\frac{n}{8}\right) + \cdots + nM\left(\frac{n}{n}\right)\right] + n \\
&\leq \frac{3}{2}\left[\frac{M(n)}{2} + 4\frac{M(n)}{16} + 8\frac{M(n)}{64} + \cdots + n\frac{M(n)}{n^2}\right] + n \\
&= \frac{3}{2}M(n)\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots + \frac{1}{n}\right) + n \\
&= \frac{3}{2}M(n)\left(1 - \frac{1}{n}\right) + n
\end{aligned}$$

because the geometric series $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots + \frac{1}{n} = 1 - \frac{1}{n}$. Therefore, since $\frac{3}{2n}M(n) \geq n$

$$\begin{aligned}
T(n) &= \frac{3}{2}M(n) - \frac{3}{2n}M(n) + n \\
&\leq \frac{3}{2}M(n)
\end{aligned}$$

Now we can state the following theorem.

Theorem 5. *Let M(n) be the amount of work required to multiply two n* x *n matrices over a ring. If for all n,* $4M(n) \leq M(2n) \leq 8M(n)$*, then there exists a constant c such that the inverse of any n* x *n nonsingular upper (lower) triangular matrix A can be computed in cM(n) amount of work*[AHU74].

The work of finding an inverse is not more complex than the amount of work needed to multiply matrices .

Chapter 5

LUP Decomposition

5.1 Presentation of Algorithm for LUP Decomposition

LUP decomposition uses some of the concepts that were used in the section 4.1. If a matrix can be separated into a product of lower triangular and upper triangular matrices, they will be easier to work with. Since some matrices may contain columns with all entries equal to zero, columns in the original matrix may need to be rearranged. This will require a permutation matrix. Hence, in the term LUP decomposition L stands for a lower triangular matrix, U stands for an upper triangular matrix and P stands for a permutation matrix. The LUP form of a matrix can be found through a process called FACTOR, resulting in the LUP decomposition of the matrix.

Doctors Aho, Hopcroft, and Ullman describe the LUP decomposition method in their book, *The Design and Analysis of Computer Algorithms.* To study this method several definitions are necessary. The following definition is given:

Definition 3. The LU decomposition of an m x n matrix A, $m \leq n$, is a pair of matrices L and U such that

$$A = LU$$

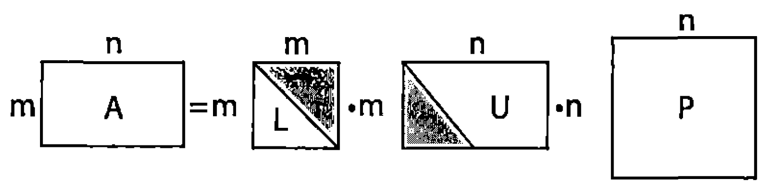
where L is m x m unit lower triangular and U is m x n upper triangular.

If the determinant of A is not equal to zero, then A is nonsingular. Not every matrix A will decompose. However, if A is nonsingular, then A can be mulitplied by a permutation matrix such that AP^{-1} has an LU decomposition. Therefore, $AP^{-1} = LU$. Multiplying by P on both sides gives $A = LUP$.

Now for any nonsingular matrix A, we can use the algorithm to find L, U, and P such that

$$A = LUP$$

This is called the LUP decomposition of A. This equation is represented in the following diagram.



The following algorithm, called FACTOR, finds the LUP decomposition for any nonsingular matirx A. It is written FACTOR(A, m, p) where A represents the matrix being written in LUP, m represents the number of rows in A with m a power of 2, and p represents the number of columns in A.

Algorithm 6. *For a nonsingular n x n matrix M, where n is a power of 2, we call the procedure* FACTOR *(shown below) to get L, U, P such that $M = LUP$ and L is unit lower triangular, U is upper triangular and P is a permutation matrix.*

FACTOR: FACTOR(A, m, p)

1. If $m = 1$ then set $L = (1)$. If $m \neq 1$ then go to step 5.

2. If the first column does not have all elements zero, then P is the p x p identity matrix. If the first column has all elements of zero, then find, if possible, a column of A, call it c, that does not have all elements zero. Let P equal the p x p identity matrix where the first column and column c are interchanged.

For example if

$$A = \begin{pmatrix} 0_{11} & 0_{12} & \dots & c_{1c} & \dots & 0_{1p} \\ 0_{21} & 0_{22} & \dots & c_{2c} & \dots & 0_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0_{m1} & 0_{m2} & \dots & c_{mc} & \dots & 0_{mp} \end{pmatrix} \text{ when } m \leq p \text{ and } c \leq p$$

and where the c-column does not have all elements equal to zero. Then

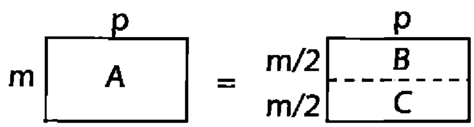
$$P = \begin{pmatrix} 0_{11} & 0_{12} & \dots & 1_{1c} & \dots & 0_{1p} \\ 0_{21} & 1_{22} & \dots & 0_{2c} & \dots & 0_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1_{c1} & 0_{c2} & \dots & 0_{cc} & \dots & 0_{cp} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_{p1} & 0_{p2} & \dots & 0_{pc} & \dots & 1_{pp} \end{pmatrix}.$$

Note: $P = P^{-1}$

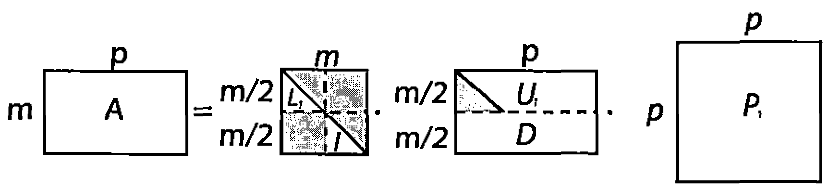
3. Let $U = AP$.

4. Multiply both sides of the equation in step 3 by P^{-1}. This gives $A = UP^{-1}$ or $A = UP$. Since $L = (1)$, the equation can be written as $A = LUP$.

5. If $m \neq 1$ then partition A into two, $m/2$ x p matrices B and C.



6. FACTOR$(B, m/2, p)$ to produce L_1, U_1, P_1. Return to step 1.



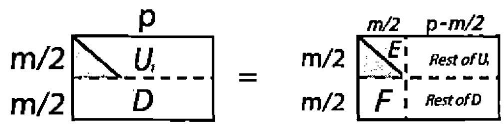
$$B = L_1 U_1 P_1$$

7. From the diagram in step 6 above, looking at the bottom half, we need

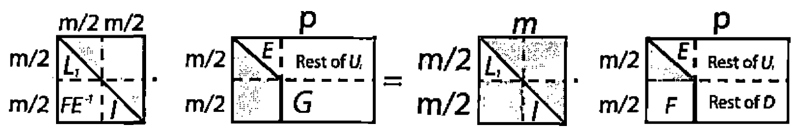
$$\begin{aligned} C &= IDP_1 \\ C &= DP_1 \\ CP_1^{-1} &= D. \end{aligned}$$

Find $D = CP_1^{-1}$.

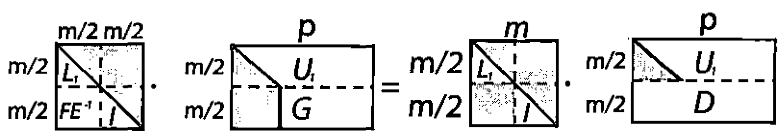
8. Let E be the first $m/2$ columns of U_1 and F be the first $m/2$ columns of D.



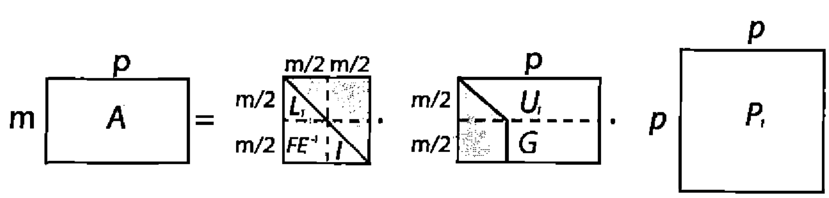
To get the matrix to be upper triangular, the box F must be a zero matrix. This is achieved by changing the lower left box of the first matrix to FE^{-1}. That is, we will find G so that the following diagram holds.



We see that the first $m/2$ columns are correct since, $FE^{-1}E + 0 = F$. The other columns will also be correct if we make the right choice for G.



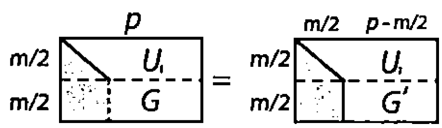
The top half is identical because of the zeros in the upper right corner of the first matrix. Looking at the bottom half of the boxes, $FE^{-1}U_1 + G = D$ or $G = D - FE^{-1}U_1$ (See step 9). Therefore,



9. Find $G = D - FE^{-1}U_1$. Note: The first $m/2$ columns of G are all zero.

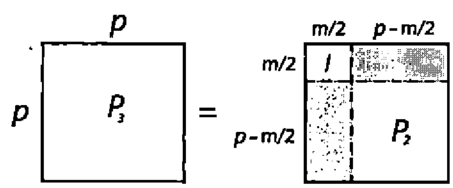
At this point, D and F were found in steps 7 and 8, respectively. E^{-1} may be found by the methods from Chapter 4 on finding inverses. Thus, G may be found by the above equation and the given matrices.

10. Let G' be the rightmost $p-m/2$ columns of G. G' = Rest of D from diagram in step 8. G' is a matrix with $m/2$ rows and $p-m/2$ columns.



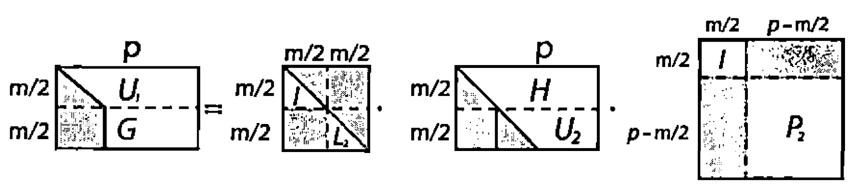
11. FACTOR $(G', m/2, p-m/2)$ to produce L_2, U_2, P_2.

12. Let P_3 be the p x p matrix with a $m/2$ x $m/2$ identity matrix in the upper left and P_2 in the lower right.

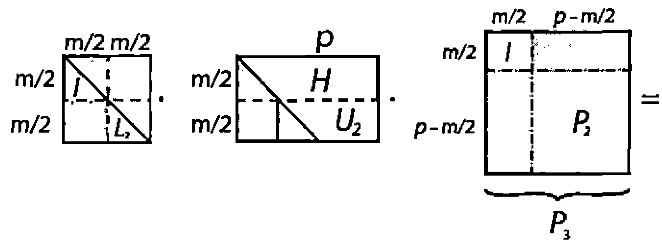


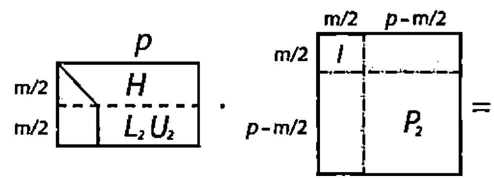
13. Find $H = U_1 P_3^{-1}$.

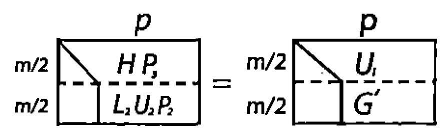
We will show the following to be true.



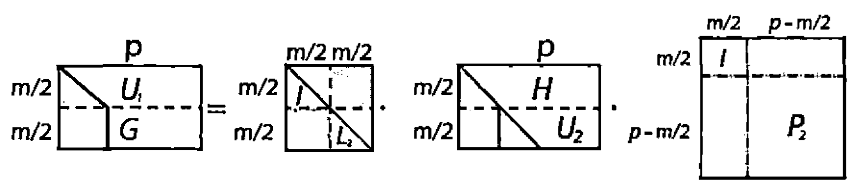
The following statements verify the above.



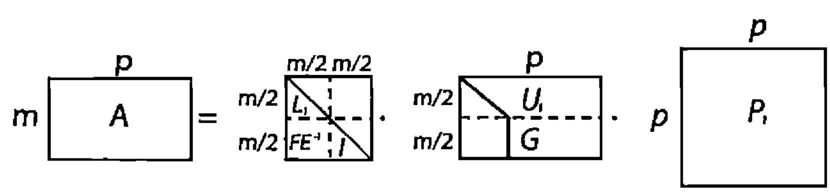




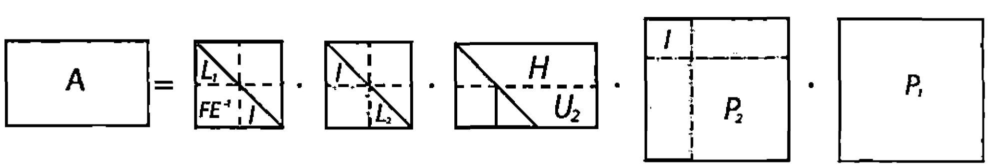
Now substitute



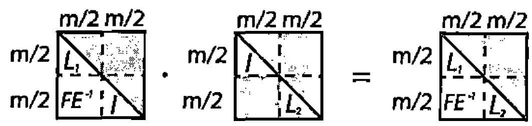
into



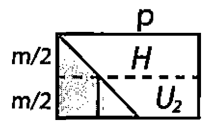
to get



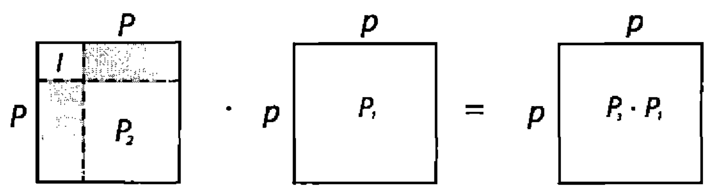
14. Let L be the first two matrices multiplied together which results in a matrix with L_1, $0_{m/2}$, FE^{-1} and L_2 such that



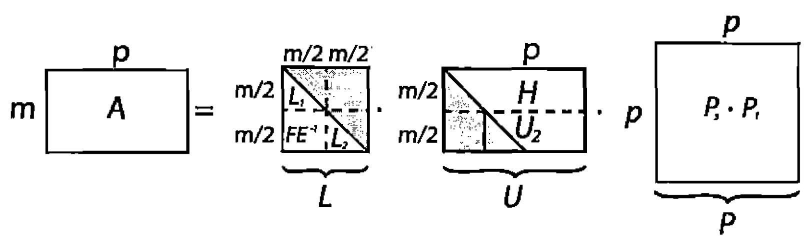
15. Let U be an m x p matrix with H in the upper half and $0_{m/2}$ and U_2 in the lower half.



16. Let P be a p x p matrix from the product of the last two matrices, P_3P_1.



17. LUP decomposition complete[AHU74].



5.2 An Example of LUP Decomposition

The following is an example of finding the LUP decomposition of a matrix. Each step is given followed by the application to the particular matrix.

We begin with FACTOR(A, m, p). Let

$$A = \begin{pmatrix} 1 & -1 & 3 & 2 \\ -2 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix}.$$

To find the LUP decomposition, we will FACTOR(A, 4, 4).

1. If $m = 1$, then set $L = (1)$. If $m \neq 1$, then go to step 5. Since $m = 4$, we will go to step 5.

5. If $m \neq 1$, then partition A into two $m/2$ x p matrices B and C.

Partition A

$$A = \left(\begin{array}{cccc} 1 & -1 & 3 & 2 \\ -2 & 0 & 1 & -1 \\ \hline 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \end{array}\right)$$

Therefore,

$$\begin{aligned} B &= \begin{pmatrix} 1 & -1 & 3 & 2 \\ -2 & 0 & 1 & -1 \end{pmatrix} \\ C &= \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix}. \end{aligned}$$

6. FACTOR($B, m/2, p$) to produce L_1, U_1, P_1 We need to FACTOR(B, 2, 4). To do this, return to step 1. Let $m' = m/2 = 2$.

6.1 If $m = 1$, then set $L = (1)$. If $m \neq 1$ then go to step 5. This time, $m = 2$ requiring step 5.

6.5 If $m \neq 1$ then partition A into two $m'/2$ x p matrices B and C.

Partition B.

$$B = \left(\begin{array}{cccc} 1 & -1 & 3 & 2 \\ \hline 2 & 0 & 1 & -1 \end{array}\right)$$

Now, $B' = \begin{pmatrix} 1 & -1 & 3 & 2 \end{pmatrix}$ and $C' = \begin{pmatrix} -2 & 0 & 1 & -1 \end{pmatrix}$.

6.6 FACTOR$(B, m'/2, p)$ to produce L_1, U_1, P_1. This means we need to FACTOR(B', 1, 4). To do this, return to step 1.

6.6.1 If $m = 1$, then set $L = (1)$. If $m \neq 1$, then go to step 5. This time, $m = m'/2 = 1$ so we follow to step 2.

6.6.2 If the first column does not have all elements zero, then P is the p x p identity matrix. If the first column has all elements of zero, then find, if possible, a column of A, call it c, that does not have all elements zero and interchange the first column with c. Let P equal the p x p identity matrix where the same two columns are interchanged.

Since the first column of B' does not have elements that are zero and $p = 4$, P is the 4 x 4 identity matrix.

6.6.3 Let $U = AP$. At this point, $A = B'$ and P is an identity matrix, so $U = B'$.

6.6.4 This gives $A = LUP$. The result is

$$B' = \begin{pmatrix} 1 & -1 & 3 & 2 \end{pmatrix} = \underbrace{(1)}_{L_1} \cdot \underbrace{\begin{pmatrix} 1 & -1 & 3 & 2 \end{pmatrix}}_{U_1} \cdot \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{P_1}$$

6.6 We have now completed FACTOR$(B, m/2, p)$ to produce L_1, U_1, P_1. FACTOR$(B', 1, 4)$ gives

$$L_1 = (1) \qquad U_1 = \begin{pmatrix} 1 & -1 & 3 & 2 \end{pmatrix} \qquad P_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

6.7 Find $D = CP_1^{-1}$. At this point in the algorithm, $D = C'$, since P_1^{-1} is an identity matrix and $C = C'$.

$$D = \begin{pmatrix} -2 & 0 & 1 & -1 \end{pmatrix}$$

6.8 Let E be the first $m'/2$ columns of U_1 and F be the first $m'/2$ columns of D.
Therefore, $E = (1)$ and $F = (-2)$.

6.9 Find $G = D - FE^{-1}U_1$. Note: The first $m'/2$ columns of G are all 0.
Here, $E^{-1} = (1)$, so

$$G = \begin{pmatrix} -2 & 0 & 1 & -1 \end{pmatrix} - (-2)(1)\begin{pmatrix} 1 & -1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 7 & 3 \end{pmatrix}.$$

6.10 Let G' be the rightmost $p - m'/2$ columns of G. Because $p = 4$ and $m' = 2$, G' will be the right three columns of G.

$$G' = \begin{pmatrix} -2 & 7 & 3 \end{pmatrix}$$

6.11 FACTOR $(G', m'/2, p - m'/2)$ to produce L_2, U_2, P_2. At this point, the call is to FACTOR$(G', 1, 3)$. Return to step 1.

6.11.1 If $m = 1$, then set $L = (1)$. If $m \neq 1$, then go to step 5. Since $m = 1$, we set $L = (1)$.

6.11.2 If the first column does not have all elements zero, then P is the p x p identity matrix. If the first column has all elements of zero, then find, if possible, a column of A, call it c, that does not have all elements zero and interchange the first column with c. Let P equal the p x p identity matrix where the same two columns are interchanged.

Since the first column of G' does not have elements that are zero and $p = 3$, P is the 3 x 3 identity matrix.

6.11.3 Let $U = AP$. At this point, $A = G'$ and P is an identity matrix, so $U = G'$.

6.11.4 This gives $A = LUP$.

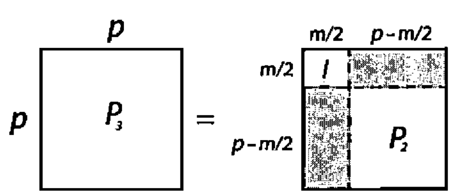
$$G' = \underbrace{(1)}_{L} \cdot \underbrace{\begin{pmatrix} -2 & 7 & 3 \end{pmatrix}}_{U} \cdot \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{P}$$

Returning to step 11.

6.11 FACTOR $(G', m/2, p - m/2)$ to produce L_2, U_2, P_2.

$$L_2 = (1) \qquad U_2 = \begin{pmatrix} -2 & 7 & 3 \end{pmatrix} \qquad P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

6.12 Let P_3 be the p x p matrix with a $m/2$ x $m/2$ identity matrix in the upper left and P_2 in the lower right.

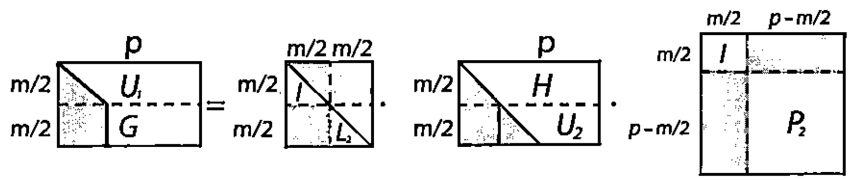


$$P_3 = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

6.13 Find $H = U_1 P_3^{-1}$.

Because $P_3^{-1} = P_3 = I$, the result is $H_1 = U_1$.

$$H_1 = \begin{pmatrix} 1 & -1 & 3 & 2 \end{pmatrix}$$



$$\begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{Q}$$

Substituting Q into R, produces the following

$$B = \begin{pmatrix} 1 & -1 & 3 & 2 \\ -2 & 0 & 1 & -1 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \cdot \underbrace{\begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \end{pmatrix}}_{R} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{Q} \cdot$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Now, reassociate the matrices to find L, U, and P for B.

$$\underbrace{\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{L}$$

$$\underbrace{\begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \end{pmatrix}}_{U}$$

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{P}$$

After multiplication, the LUP decomposition for B is

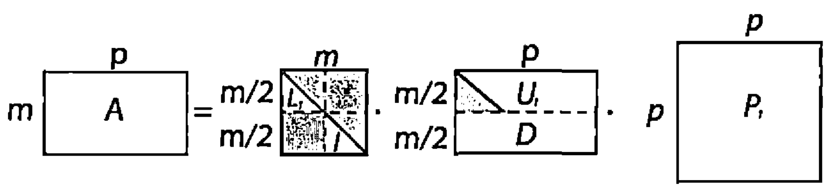
$$B = \underbrace{\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}}_{L} \cdot \underbrace{\begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \end{pmatrix}}_{U} \cdot \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{P}.$$

Remember the LUP decomposition of B results in L_1, U_1 and P_1 for FACTOR(A, 4, 4). Now we return to step 6 with the following result

$$L_1 = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \qquad U_1 = \begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \end{pmatrix} \qquad P_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

7. Find $D = CP_1^{-1}$. Since P_1^{-1} is an identity matrix, $D = C$.

$$D = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix}$$



$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

8. Let E be the first $m/2$ columns of U_1 and F be the first $m/2$ columns of D.

$$E = \begin{pmatrix} 1 & -1 \\ 0 & -2 \end{pmatrix} \qquad F = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} \qquad E^{-1} = \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & -\frac{1}{2} \end{pmatrix}$$

9. Find $G = D - FE^{-1}U_1$. Note: The first m/2 columns of G are all 0.

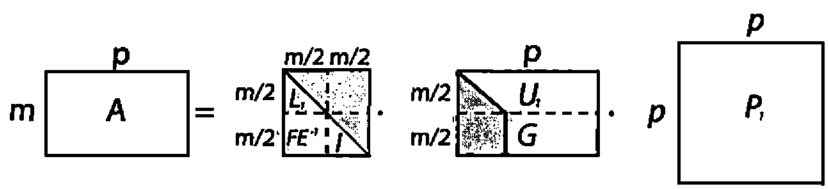
$$G = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & -\frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \end{pmatrix}$$

$$G = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & -\frac{1}{2} \\ 1 & -\frac{3}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \end{pmatrix}$$

$$G = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 1 & -\frac{7}{2} & -\frac{3}{2} \\ 1 & 2 & -\frac{15}{2} & -\frac{5}{2} \end{pmatrix}$$

$$G = \begin{pmatrix} 0 & 0 & \frac{9}{2} & \frac{5}{2} \\ 0 & 0 & \frac{15}{2} & \frac{7}{2} \end{pmatrix}$$

According to the diagram



$$A = \begin{pmatrix} 1 & -1 & 3 & 2 \\ -2 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & 0 \\ 1 & -\frac{3}{2} & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \\ 0 & 0 & \frac{9}{2} & \frac{5}{2} \\ 0 & 0 & \frac{15}{2} & \frac{7}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

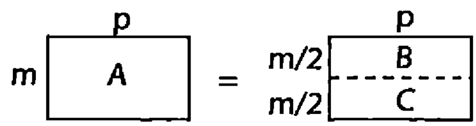
10. Let G' be the rightmost $p - m/2$ columns of G. Because $p = 4$ and $m = 4$, G' will be the right two columns of G.

$$G' = \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ \frac{15}{2} & \frac{7}{2} \end{pmatrix}$$

11. FACTOR $(G', m/2, p - m/2)$ to produce L_2, U_2, P_2. This calls for FACTOR(G', 2, 2). Return to step 1.

11.1 If $m = 1$ then set $L = (1)$. If $m \neq 1$ then go to step 5. This time, $m = 2$ requiring step 5.

11.5 If $m \neq 1$ then partition A into two ($m/2$ x p) matrices B and C.



$$G' = \left(\begin{array}{cc} \frac{9}{2} & \frac{5}{2} \\ \hline \frac{15}{2} & \frac{7}{2} \end{array}\right)$$

Partition G' giving

$$B_3 = \left(\frac{9}{2} \quad \frac{5}{2} \right) \qquad C_3 = \left(\frac{15}{2} \quad \frac{7}{2} \right)$$

11.6 FACTOR$(B, m/2, p)$ to produce L_1, U_1, P_1 Return to step 1. Here, FACTOR$(B, 1, 2)$. Return to step 1.

11.6.1 If $m = 1$ then set $L = (1)$. If $m \neq 1$ then go to step 5. Since $m = 1$, we set $L = (1)$.

11.6.2 If the first column does not have all elements zero, then P is the p x p identity matrix. If the first column has all elements of zero, then find, if possible, a column of A, call it c, that does not have all elements zero and interchange the first column with c. Let P equal the p x p identity matrix where the same two columns are interchanged.

P will be the 2 x 2 identity matrix.

11.6.3 Let $U = AP$. $U = \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \end{pmatrix}$.

11.6.4 This gives $A = LUP$. Obtaining L, U, and$P$ we go back to step 6.

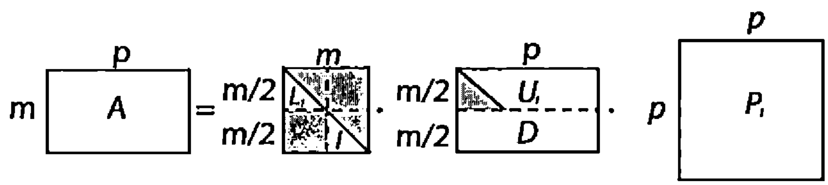
11.6 We now have

$$L_1 = (1) \qquad U_1 = \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \end{pmatrix} \qquad P_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

11.7 Find $D = CP_1^{-1}$. Since P_1^{-1} is an identity matrix, $D = C$.

$$D = \begin{pmatrix} \frac{15}{2} & \frac{7}{2} \end{pmatrix}$$

According to the diagram,



$$\begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ \frac{15}{2} & \frac{7}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ \frac{15}{2} & \frac{7}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

11.8 Let E be the first $m/2$ columns of U_1 and F be the first $m/2$ columns of D.

$$E = \begin{pmatrix} \frac{9}{2} \end{pmatrix} \qquad F = \begin{pmatrix} \frac{15}{2} \end{pmatrix} \qquad E^{-1} = \begin{pmatrix} \frac{2}{9} \end{pmatrix}$$

11.9 Find $G = D - FE^{-1}U_1$.

$$\begin{aligned}
G &= \begin{pmatrix} \frac{15}{2} & \frac{7}{2} \end{pmatrix} - \begin{pmatrix} \frac{15}{2} \end{pmatrix} \cdot \begin{pmatrix} \frac{2}{9} \end{pmatrix} \cdot \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \end{pmatrix} \\
G &= \begin{pmatrix} \frac{15}{2} & \frac{7}{2} \end{pmatrix} - \begin{pmatrix} \frac{5}{3} \end{pmatrix} \cdot \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \end{pmatrix} \\
G &= \begin{pmatrix} 0 & -\frac{2}{3} \end{pmatrix}
\end{aligned}$$

The LUP decomposition of G' is as follows

$$G' = \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ \frac{15}{2} & \frac{7}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{5}{3} & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ 0 & -\frac{2}{3} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

11.10 Let G' be the right most $p - m/2$ columns of G.

$$G' = \begin{pmatrix} -\frac{2}{3} \end{pmatrix}$$

11.11 FACTOR $(G', m/2, p - m/2)$ to produce L_2, U_2, P_2. Meaning FACTOR$(G', 1, 1)$.

$$L_2 = (1) \qquad U_2 = \begin{pmatrix} -\frac{2}{3} \end{pmatrix} \qquad P_2 = (1)$$

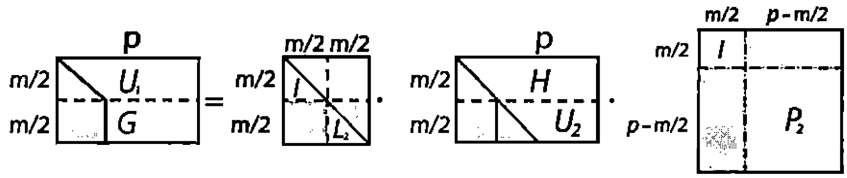
11.12 Let P_3 be the p x p matrix with a $m/2$ x $m/2$ identity matrix in the upper left and P_2 in the lower right.

Remember, we are factoring $G' = \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ \frac{15}{2} & \frac{7}{2} \end{pmatrix}$. Therefore, P_3 is the 2 x 2 identity matrix.

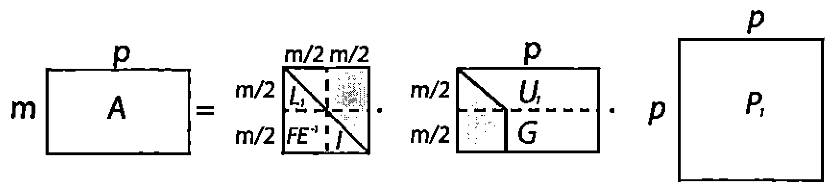
11.13 Find $H = U_1 P_3^{-1}$.

Because $P_3^{-1} = P_3 = I$, the result is $H = U_1$. Therefore,

$$H = U_1 = \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \end{pmatrix}$$



$$\begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ 0 & \neg\frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ 0 & \neg\frac{2}{3} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$



Finally through substitution we get

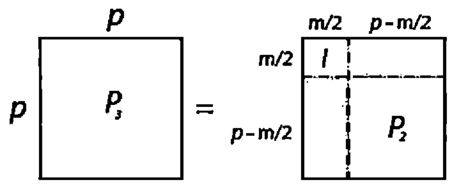
$$\begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ \frac{15}{2} & \frac{7}{3} \end{pmatrix} = \begin{pmatrix} 1 & \frac{5}{3} \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ 0 & -\frac{2}{3} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ \frac{15}{2} & \frac{7}{3} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & \frac{5}{3} \\ 0 & 1 \end{pmatrix}}_{L} \cdot \underbrace{\begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ 0 & -\frac{2}{3} \end{pmatrix}}_{U} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{P}$$

11. FACTOR $(G', m/2, p - m/2)$ to produce L_2, U_2, P_2. Return to step 11 having figured

$$L_2 = \begin{pmatrix} 1 & \frac{5}{3} \\ 0 & 1 \end{pmatrix} \qquad U_2 = \begin{pmatrix} \frac{9}{2} & \frac{5}{2} \\ 0 & -\frac{2}{3} \end{pmatrix} \qquad P_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

12. Let P_3 be the p x p matrix with a $m/2$ x $m/2$ identity matrix in the upper left and P_2 in the lower right.

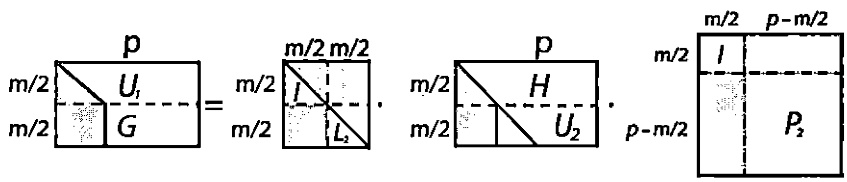


Using the diagrams, we see that P_3 is the 4 x 4 identity matrix.

13. Find $H = U_1 P_3^{-1}$.

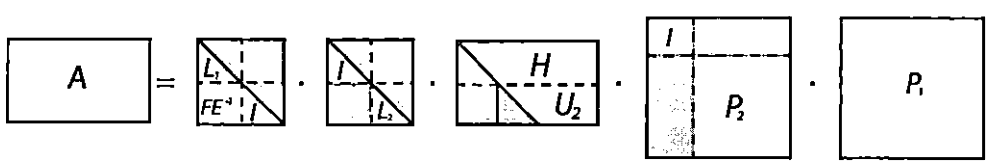
Because $P_3^{-1} = P_3 = I$, the result is $H = U_1$. Therefore,

$$H = U_1 = \begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \end{pmatrix}$$



$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{5}{3} & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \\ 0 & 0 & \frac{9}{2} & \frac{5}{2} \\ 0 & 0 & 0 & -\frac{2}{3} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{Q}$$

Finally, substituting yields the LUP decomposition for the original matrix.



$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & 0 \\ 1 & -\frac{3}{2} & 0 & 1 \end{pmatrix} \cdot Q \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$A = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & 0 \\ 1 & -\frac{3}{2} & \frac{5}{3} & 1 \end{pmatrix}}_{L} \cdot \underbrace{\begin{pmatrix} 1 & -1 & 3 & 2 \\ 0 & -2 & 7 & 3 \\ 0 & 0 & \frac{9}{2} & \frac{5}{2} \\ 0 & 0 & 0 & -\frac{2}{3} \end{pmatrix}}_{U} \cdot \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{P}$$

FACTOR is complete and we have found the LUP decomposition of A.

5.3 Work Needed to Complete *LUP* Decomposition

We now look at the amount of work involved in finding an LUP decomposition a nonsingular n x n matrix with $n = 2^k$. Let A be an m x p, with $m = 2^k$, matrix and FACTOR(A, m, p). We know from experience now that the amount of work needed will be based on the number of rows in the matrix we call to FACTOR. Let $T(m)$ represent the amount of work needed to call FACTOR(A, m, p). If $m = 1$, then only the first four steps of the algorithm are necessary to produce L, U, and P. Step 2 might require an interchange of columns. Step 3 would require AP meaning at most n operations. Therefore, to accomplish steps 1-4, when $m = 1$, would be some constant b times n.

$$T(1) = bn$$

Now we look at FACTOR when $n \geq 2$. This will involve going through the complete algorithm where FACTOR is called at steps 6 and 11. Each of those steps will require $T\left(\frac{m}{2}\right)$ amount of work. Steps 7 and 13 will require finding the inverse of a permutation matrix, which will be a function of n, meaning $O(n)$.

For step 9, we need to find E^{-1}. Using theorem 1, mentioned in section 4.1, and the fact that E is an $\frac{m}{2}$ x $\frac{m}{2}$ matrix, the work needed will be some constant c times $M\left(\frac{m}{2}\right)$. We also need to compute FE^{-1}, which will also take $M\left(\frac{m}{2}\right)$: Then we need to multiply FE^{-1} by U_1. U_1 can be at the most $\frac{m}{2}$ x n. We know n is divisible by $\frac{m}{2}$, because in the first call of FACTOR $n = m$, both n and m are powers of 2, and $m \leq n$. Think of U_1 as $\frac{m}{2}$ x $\frac{m}{2}$ blocks set next to each other. If you divide n by $\frac{m}{2}$, you would have at most $2\frac{n}{m}$ number of blocks. Therefore the amount of work would be $2\frac{n}{m}M\left(\frac{m}{2}\right)$. We consider 2 as part of the constant c giving the amount of work for step 9 to be $\frac{cn}{m}M\left(\frac{m}{2}\right)$.

The rest of the steps will have at most some constant, d times mn. This gives the following recursive formula for finding the amount of work needed to complete LUP decomposition.

$$T(m) \leq \begin{cases} 2T\left(\frac{m}{2}\right) + \frac{cn}{m}M\left(\frac{m}{2}\right) + dmn, & m > 1 \\ bn, & m = 1 \end{cases}$$

for constants b, c, and d.

We now look to find a limit to the work in terms of $M(m)$, the work of multiplying two matrices.

$$T(m) \leq 2T\left(\frac{m}{2}\right) + \frac{cn}{m}M\left(\frac{m}{2}\right) + dmn$$

Consider the constant terms $\frac{cn}{m}$ and dmn combined to be $\frac{en}{m}$. We continue in a manner similar to what was done for the work of an inverse.

$$\begin{aligned} T(m) &\leq 2T\left(\frac{m}{2}\right) + \frac{en}{m}M\left(\frac{m}{2}\right) \\ &= \frac{en}{m}M\left(\frac{m}{2}\right) + 2\left[2T\left(\frac{m}{4}\right) + \frac{2en}{m}M\left(\frac{m}{4}\right)\right] \\ &= \frac{en}{m}M\left(\frac{m}{2}\right) + \frac{4en}{m}M\left(\frac{m}{4}\right) + 4T\left(\frac{m}{4}\right) \\ &= \frac{en}{m}M\left(\frac{m}{2}\right) + \frac{4en}{m}M\left(\frac{m}{4}\right) + 4\left[2T\left(\frac{m}{8}\right) + \frac{4en}{m}M\left(\frac{m}{8}\right)\right] \end{aligned}$$

The recursive process continues until $m = 1$. This occurs when $\log m = k \log 2$ for some constant k. We use $\log m$ as the number of times the formula is applied. Therefore,

$$\begin{aligned} T(m) &\leq \frac{en}{m}M\left(\frac{m}{2}\right) + \frac{4en}{m}M\left(\frac{m}{4}\right) + \frac{16en}{m}M\left(\frac{m}{8}\right) + \ldots + \frac{4^{\log m}en}{m}M\left(\frac{m}{2^{\log m}}\right) + \\ &\quad + mT(1) \\ &= \frac{en}{4m}\left[4M\left(\frac{m}{2}\right) + 16M\left(\frac{m}{4}\right) + 64M\left(\frac{m}{8}\right) + \ldots + 4^{\log m}M(1)\right] + m(bn) \\ &= \frac{en}{4m}\left[\sum_{i=1}^{\log m} 4^i M\left(\frac{m}{2^i}\right)\right] + bmn \end{aligned}$$

Consider the condition that for some $\epsilon > 0$ then $M(2m) \geq 2^{2+\epsilon}M(m)$. To see the flow of statements, we rewrite in terms of less than.

$$2^{2+\epsilon}M(m) \leq M(2m)$$

Replacing m with $\left(\frac{m}{2}\right)$, we get

$$2^{2+\epsilon}M\left(\frac{m}{2}\right) \leq M(m).$$

If we divide $\left(\frac{m}{2}\right)$ by 2, we get

$$2^{2+\epsilon}2^{2+\epsilon}M\left(\frac{m}{4}\right) \leq M(m).$$

As we continue to divide m by 2, we see

$$\begin{aligned} \left(2^{2+\epsilon}\right)^i M\left(\frac{m}{2^i}\right) &\leq M(m) \\ 2^{2i}\cdot 2^{\epsilon i}M\left(\frac{m}{2^i}\right) &\leq M(m) \\ 4^i M\left(\frac{m}{2^i}\right) &\leq \frac{1}{2^{\epsilon i}}M(m). \end{aligned}$$

Therefore,

$$4^i M\left(\frac{m}{2^i}\right) \leq \left(\frac{1}{2^\epsilon}\right)^i M(m).$$

Because the terms are all positive, we can say

$$\sum_{i=1}^{\log m} 4^i M\left(\frac{m}{2^i}\right) \leq \sum_{i=1}^{\log m}\left(\frac{1}{2^\epsilon}\right)^i M(m).$$

We can see

$$\begin{aligned} \sum_{i=1}^{\log m}\left(\frac{1}{2^\epsilon}\right)^i M(m) &< M(m)\sum_{i=1}^{\infty}\frac{1}{2^{\epsilon i}} \\ &= M(m)\sum_{i=1}^{\infty}\left(\frac{1}{2^\epsilon}\right)^i \end{aligned}$$

since we are adding more positive terms on the right hand side, the total would be larger. We can take $M(m)$ out of the summation since it is not dependent on i.

This brings us to the following statement.

$$T(m) \leq \frac{en}{4m}M(m)\sum_{i=1}^{\infty}\left(\frac{1}{2^\epsilon}\right)^i + bmn$$

Since the summation is a geometric series with $r < 1$, we know it will equal some constant, f.

$$T(m) \leq \frac{fen}{4m} M(m) + bmn$$

Combining the constants into k, we get

$$T(m) \leq \frac{kn}{m} M(m) + bmn$$

The LUP decomposition algorithm is for an n x n matrix, the algorithm starts with $m = n$. Thus, the amount of work needed to complete the LUP decomposition for an n x n matrix is represented by

$$T(n) \leq kM(n).$$

We see that the amount of work needed is based on the work to multiply n x n matrices.

Theorem 7. *Suppose the amount of work needed to multiply two n x n matrices is represented by $M(n)$, and for all m and some $\epsilon > 0$, the following is true: $M(2m) \geq 2^{2+\epsilon} M(m)$. Then the amount of work needed to perform the LUP decomposition shown in Algorithm 3, $T(n)$, is less than some constant k times $M(n)$ for any nonsingular matrix. Meaning, $T(n) \leq kM(n)$*[AHU74].

Chapter 6

Solving Equations

6.1 Using LUP to Solve

By writing a matix in LUP decomposition, a system of linear equations can be solved without using inverses or row reduction methods. The system may be solved by using matrix multiplication and back substitution.

Consider A to be an n x n matrix. Let

$$Ax = b$$

where both x and b are n x 1 column vectors. Assume A can be factored into LUP, meaning $A = LUP$. Then

$$LUPx = b.$$

Now let

$$UPx = y$$

and

$$Ly = b.$$

By isolating L and because L is a lower triangular matrix, one variable is solved and the rest of the system can be solved by using back substitution. Therefore, the equation $Ly = b$ can be solved for y. Then use

$$UPx = y$$

and similar methods to solve for x. First, multiply UP. Then solve for x using the values that were already found for y [AHU74].

6.2 An Example of Using LUP to Solve an Equation

Looking at an example will give a visual explanation of what occurs with this process. Consider the following system of equations.

$$\begin{array}{ccccccccc} & & & & x_3 & + & 2x_4 & = & 7 \\ & & & & 3x_3 & & & = & 9 \\ x_1 & - & x_2 & & & + & x_4 & = & 3 \\ 2x_1 & & & - & x_3 & + & 3x_4 & = & 10 \end{array}$$

First, write the systems of equations as a matrix equation.

$$\begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 0 \\ 1 & -1 & 0 & 1 \\ 2 & 0 & -1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 7 \\ 9 \\ 3 \\ 10 \end{pmatrix}$$

Here

$$A = \begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 0 \\ 1 & -1 & 0 & 1 \\ 2 & 0 & -1 & 3 \end{pmatrix}.$$

It may seem simple to just find A^{-1} with the method discussed in section 3. Then solve the equation by multiplying both sides by the inverse. However, here $A_{11} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, A_{11}^{-1} does not exist and the matrix is not invertible. This suggests finding the LUP decomposition for A and using the method described above.

After using the LUP decomposition algorithm, the LUP decomposition is found for A.

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 0 & -\frac{1}{6} & 1 & 0 \\ -1 & -\frac{5}{6} & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & -6 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Therefore, the $LUPx = b$ becomes

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 0 & -\frac{1}{6} & 1 & 0 \\ -1 & -\frac{5}{6} & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & -6 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} =$$

$$= \begin{pmatrix} 7 \\ 9 \\ 3 \\ 10 \end{pmatrix}$$

Now, using $UPx = y$ and assuming

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

then $Ly = b$ becomes the following

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 0 & -\frac{1}{6} & 1 & 0 \\ -1 & -\frac{5}{6} & 2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 7 \\ 9 \\ 3 \\ 10 \end{pmatrix}.$$

Using matrix mutliplication, $y_1 = 7$. By using back substitution into the other equations obtained, the following statement can be made

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 7 \\ -12 \\ 1 \\ 5 \end{pmatrix}.$$

Now looking at the equation $UPx = y$

$$\begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & -6 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 7 \\ -12 \\ 1 \\ 5 \end{pmatrix}.$$

Multiplying UP

$$\begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -6 \\ 1 & -1 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 7 \\ -12 \\ 1 \\ 5 \end{pmatrix}.$$

Now x_2 and x_4 can be solved easily giving $x_2 = \frac{5}{2}$ and $x_4 = 2$. Through substitution, the following solution is obtained

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \frac{7}{2} \\ \frac{5}{2} \\ 3 \\ 2 \end{pmatrix}.$$

We see that the amount of work needed to use this method of solving a system of equations is based on the amount of work needed to find the LUP decomposition for A, as shown in Section 5.3 and another $M(n)$, where A is an n x n matrix. There would also be some work involved in the back substitution. However, this would be a relatively small amount of work and was not explored in this paper.

Chapter 7

Determinants

One basic operation of working with matrices is finding determinants. For beginning algebra students finding the determinant of a 2 x 2 matrix is relatively easy. Finding the determinant of a 3 x 3 matrix is very doable. However, finding determinants by hand for any matrix larger than 3 x 3 can be overwhelming. In *The Design and Analysis of Computer Algorithms*, the authors state the following lemma.

Lemma 8. *If A is a square upper triangular or lower triangular matrix then the determinant of A is equal to the product of elements on the main diagonal*[AHU74].

Proof. Suppose A represents an upper triangular matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & a_{22} & a_{23} & \ldots & a_{2n} \\ 0 & 0 & a_{33} & \ldots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & a_{nn} \end{pmatrix}$$

Using 1st column

$$a_{11} \begin{pmatrix} a_{22} & a_{23} & \ldots & a_{2n} \\ 0 & a_{33} & \ldots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix} \underbrace{\pm 0 \pm 0 \pm \cdots \pm 0}_{n-1 \text{ times}}$$

$$a_{11} \cdot a_{22} \begin{pmatrix} a_{33} & \dots & a_{3n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{pmatrix} \underbrace{\pm 0 \pm 0 \pm \cdots \pm 0}_{n-2 \text{ times}}$$

$$a_{11} \cdot a_{22} \cdot a_{33} \begin{pmatrix} a_{44} & \dots & a_{4n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{pmatrix} \underbrace{\pm 0 \pm 0 \pm \cdots \pm 0}_{n-3 \text{ times}}$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$a_{11} \cdot a_{22} \cdot a_{33} \cdots a_{(n-2)(n-2)} \begin{pmatrix} a_{(n-1)(n-1)} & a_{(n-1)n} \\ 0 & a_{nn} \end{pmatrix} \pm 0 \pm 0$$

$$a_{11} \cdot a_{22} \cdot a_{33} \cdots a_{(n-2)(n-2)} \left(a_{(n-1)(n-1)} \cdot a_{nn} - 0\right)$$

$$a_{11} \cdot a_{22} \cdot a_{33} \cdots a_{(n-2)(n-2)} \cdot a_{(n-1)(n-1)} \cdot a_{nn}$$

$$a_{11} \cdot a_{22} \cdot a_{33} \cdots a_{nn}$$

□

The determinant of an upper triangular matrix is the product of the elements on the main diagonal. As is the case with an upper triangular matrix, the determinant of a lower triangular matrix is found by multiplying the elements along the main diagonal. The proof is similar to the above proof. Therefore LUP can make it easy to find the determinants of any matrix.

By rewriting a square matrix in terms of LU we can find the determinant by multiplication. For each lower (L) and upper (U) matrix the determinant is the product of elements on the main diagonal. Therefore, the amount of work needed to find a determinant of an n x n matrix is based on the the amount of work needed to find the LUP decomposition of the matrix plus the amount of work needed to multiply the elements on the main diagonal. Thus, $T(n) \leq kM(n)$.

Chapter 8

Conclusion

When my son took algebra in high school, he came home one day and declared that "There was no need to learn matrices. You can always use other methods to solve systems of equations." Clearly, he was limited in his knowledge and how the concepts of linear algebra are used in many different fields. While we often teach the beginning concepts of linear algebra and working with matrices in our college-level algebra courses, many students are not exposed to the more advanced concepts of linear algebra and their applications. In looking at the introductory paragraphs for one linear algebra text book, it states "This course is potentially the most interesting and worthwhile undergraduate mathematics course you will complete."[Lay06] In the preface it mentions the book contains applications in the fields of engineering, computer science, mathematics, physics, biology, economics and statistics. As we look at how the study of linear algebra applies to many professional fields and we study the technology used to implement such concepts, it is useful to find ways to make the current methods more efficient. The study of the complexity of linear algebra focuses on this idea.

Strassen, Laderman and others have worked to reduce the amount of multiplications steps needed when multiplying matrices. This paper shows that the work of finding inverses, solutions to systems of equations and determinants is no more complex than a constant factor of the work of matrix multiplication. Therefore, simplifying the process of matrix multiplication will make the procedures more efficient. While this paper did not present the work of Coppersmith, his method is apparently faster than others as mentioned in the article by Bailey and Ferguson[BF88]. Further research into this method would be useful. The ability to represent a matrix in LUP form allows the matrix to be broken down into simpler matrices for ease in finding inverses, solving systems of equations and finding determinants. The complexity of these linear algebra tasks is based on the amount of work needed to multiply matrices. As computers are used to tackle large linear algebra systems, the reduction of computational steps will decrease the time needed to operate programs and increase the capacity to run larger systems.

When my brother, who is ten years older than me, took his first computer class in college I remember him telling about his experiences using punch cards and waiting long periods of time to have the opportunity to run his stack of cards. Things have changed greatly over the last 40 years. Our systems will continue to change as methods improve and the study of efficiency continues. People still look for ways to solve systems of equations and perform other linear algebra operations with greater speed and more efficiency. The study of the complexity of linear algebra continues.

Bibliography

[AHU74] Alfred V. Aho, John E. Hopcroft, and Jeffrey D. Ullman. *The Design and Analysis of Computer Algorithms.* Addison-Wesley Publishing Company, Reading, MA, 1974.

[BF88] David H. Bailey and Helaman R. P. Ferguson. A strassen-newton alogrithm for high-speed parallelizable matrix inversion. *Supercomputing,* November 1988.

[Lad76] Julian D. Laderman. A noncommutative algorithm for multiplying 3 x 3 matrices using 23 multiplications. *Bulletin of the American Mathematical Society,* January 1976.

[Lad79] Julian D. Laderman. Computational complexity: Algorithms with fewest operations. `http://onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.1979.tb14107.x`, 1979.

[Lan08] J. M. Landsberg. Geometry and the complexity of matrix multiplication. *Bulletin (new series) of the American Mathematical Society,* April 2008.

[Lay06] David C. Lay. *Linear Algebra and Its Applications, Third Edition Update.* Pearson-Addison-Wesley, Boston, MA, 2006.

[Wei99] Eric W. Weisstein. Strassen formulas. `http://mathworld.wolfram.com/StrassenFormulas.html`, 1999.