

Arithmetic

This learning module covers the following ideas. When you make your lesson plan, it should explain and contain examples of the following:

1. Be able to use and understand vector and matrix notation, matrix addition, scalar multiplication, matrix multiplication, and matrix transposing.
2. Use Gaussian elimination to solve systems of linear equations. Become familiar with the words homogeneous, nonhomogeneous, row echelon form, and reduced row echelon form.
3. Find the rank of a matrix. Be able to determine if a collection of vectors is linearly independent. For a set of linearly dependent vectors, be able to write vectors as linear combinations of the vectors which precede it.
4. For square matrices, compute determinants, inverses, eigenvalues, and eigenvectors. The next unit will focus on applications of these ideas.

Following you will find suggested homework, and then my best attempt at condensing the information we are learning into concise notes. These are not intended to be a complete resource. As you read my notes, please do the examples yourself and then find examples like them in the homework. Reading the text (especially the solved problems) will expand your knowledge of the material you see here, help you become more familiar with the language of linear algebra, and make it easier for you to succeed in future courses. With these notes I am trying follow the model found in “Preach My Gospel,” where the gospel is taught in a complete concise manner, with suggestions provided for additional study and practice.

1 Preparation and Suggested Homework

For each day of class, I will provide a list of preparation problems. To be prepared for class, you should attempt all the problems and complete at least one of them. If you do not understand a problem, that is OK; it means you are learning and have questions that we can answer in class. During class, we will spend time in groups teaching each other, using problems similar to what you prepared. You will occasionally attempt a problem which is entirely new to you. When this occurs, please look for similar examples in the notes and read the solved problems to learn how to do this problem. You will be exercising your faith as you come to class and teach your peers something you have never before seen modeled, which will cause your confidence to grow. As a group, I strongly suggest that you assign each other problems so that at least one person in your group will be prepared with each type of problem.

Preparation Problems	
Day 1	1.4, 5.3, 2.7, 2.8,
Day 2	6.16, 6.22, 1.32 (use computer), 4.5, 4.7
Day 3	4.9, 7.1, 7.5, 7.6,

In each section, I will give you minimum goal of how many homework problems you should do (that goal will appear under the suggestions). This may not be enough to master what we are learning, rather it is a minimum guideline. You are free to work on the types of problems which will help you strengthen your weaknesses. I provide the following list to help you guide your learning. Here are some comments regarding the homework list:

1. The suggested problems are a list of key problems. These are suggestions to help you make sure you are grasping the key ideas. As you work these problems and find some that are more difficult for you, please pick more problems like these difficult ones to help you. In other words, spend you time where you will benefit the most.
2. The “worked problems” in the text count as homework, provided you work through the problem (just reading the solution does not count). I strongly suggest reading or skimming most of the relevant “worked problems,” and then taking some time to do a few.
3. Use Maple as much as you can to simplify calculations which you know you have already mastered. I strongly suggest doing “worked problems” with Maple because then you have a solution in front of you, and your job is to learn how to get the computer to give the same solution.
4. Always check your answers. If you find yourself making mistakes, pinpoint where you are making the mistakes, learn how to prevent them, and find ways to check your work.

5. If you get stuck on a problem, then you are human. If you have been working on a problem for 5-10 minutes, it might be time to skip it and move on to another. Letting a problem stew in your subconscious for a day may be the only way to eventually solve it.
6. Manage your time appropriately. Do not spend so much time on one section that you do not get to the other sections. After you have done 2 or 3 problems from one section, I suggest that you move to a new section to try those problems as well. Your brain needs time to digest the ideas we are learning, and the sooner you see all the ideas, the better. You will often find that practicing problems from later in the unit will solidify the ideas from earlier in the unit.

Concept	Sec.	Suggestions	Relevant Problems
Basic Notation	1	1,2,4,7,21,22,35	1-7,19-26,33-35
Determinants	5	1,3,5,21	1-5,21-23,27,28
Gaussian Elimination	2	7,8,13,18,19,24	1-8,13-25,28-30
Rank and Independence	6	16,19,22,23,26	15-27
	1	29,30,32	11-17,27-32
Inverses	4	1,5,6,7,9,25,28	1,5-9,15-29
Eigenvalues	7	1,3,5,6,24,26	1-6,18-40

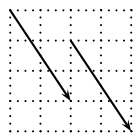
Minimum Homework Goal: 28 - Minimum Technology Goal: 8 (You should do at least 28 problems, at least 8 of which you solved using Maple. You will report what you have done via a stewardship report.)

2 Basic Matrix and Vector Notation

A matrix of size m by n has m rows and n columns. We normally write matrices using capital letters, and use the notation

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = [a_{jk}],$$

where a_{jk} is the entry in the j th row, k th column (rows always precede columns).



A row vector is a $1 \times n$ matrix, and a column vector is an $m \times 1$ matrix. The entries of a vector are called the components of a vector. Textbooks often write vectors using bold face font. By hand we add an arrow above them. The notation $\vec{v} = \langle v_1, v_2, v_3 \rangle$ can represent either row or column vectors. To draw a vector $\langle v_1, v_2 \rangle$, one option is to draw an arrow from the origin (the tail) to the point (v_1, v_2) (the head). However, the tail does not have to be placed at the origin. Both vectors on the left represent $\langle 2, -3 \rangle$.

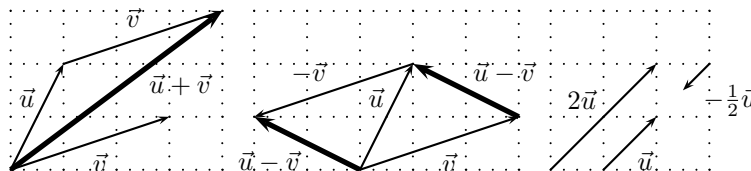
We say two matrices A and B are equal if $a_{jk} = b_{jk}$ for all j and k . We add and subtract matrices of the same size entry wise, and perform scalar multiplication cA by multiplying every entry in the matrix by the scalar c . If the number of rows and columns are equal, then we say the matrix is square. The main diagonal of a square $(n \times n)$ matrix consists of the entries $a_{11}, a_{22}, \dots, a_{nn}$, and the trace is the sum of the entries on the main diagonal $(\sum a_{jj})$. As examples, if $A = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -1 \\ 0 & 4 \end{bmatrix}$, then

$$A - 2B = \begin{bmatrix} 1 - 2 \cdot 3 & 3 - 2 \cdot (-1) \\ 0 - 2 \cdot 0 & 2 - 2 \cdot 4 \end{bmatrix} = \begin{bmatrix} -5 & 5 \\ 0 & -6 \end{bmatrix}, \text{ and the trace of } B \text{ is } 3 + 4 = 7.$$

The same principles apply to vectors. For example $\langle 1, 3 \rangle - 2 \langle -1, 2 \rangle + \langle 4, 0 \rangle = \langle 1 - 2(-1) + 4, 3 - 2(2) + 0 \rangle = \langle 7, -1 \rangle$. It is often convenient to write row vectors using column notation as in

$$\begin{bmatrix} 1 \\ 3 \end{bmatrix} - 2 \begin{bmatrix} -1 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 - 2(-1) + 4 \\ 3 - 2(2) + 0 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}.$$

Vector addition can be performed geometrically by placing the tail of the second vector at the head of the first. The resultant vector is the vector which starts at the tail of the first and ends at the head of the second. This is called the parallelogram law of addition. Scalar multiplication is equivalent to stretching a vector by the scalar, and if the scalar is negative then the vector turns around to point in the opposite direction. Geometric arithmetic is illustrated below. Take some time to practice drawing vectors and performing addition, subtraction, and scalar multiplication geometrically. Notice that vector subtraction $\vec{u} - \vec{v}$ yields a vector whose head is at the head of \vec{u} and tail at the head of \vec{v} .



The transpose of a matrix $A = [a_{jk}]$ is a new matrix $A^T = [a_{kj}]$ where you switch the rows and columns:

$$A = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 2 \end{bmatrix} \quad A^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -1 & 2 \end{bmatrix}.$$

3 The Dot Product and Matrix Multiplication

The dot product of two vectors $\vec{u} = \langle u_1, u_2, \dots, u_n \rangle$ and $\vec{v} = \langle v_1, v_2, \dots, v_n \rangle$ of the same size is the scalar $\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2 + \dots + u_nv_n = \sum u_i v_i$. Just multiply corresponding components and then add the products. For example $[1 \ 3 \ -2] \cdot [2 \ -1 \ 4] = (1)(2) + (3)(-1) + (-2)(4) = -9$. The dot product is connected to length and angles. For now, we will use the dot product to multiply matrices.

The product AB of two matrices $A_{m \times n}$ and $B_{n \times p}$ is a new matrix $C_{m \times p} = [c_{jk}]$ where $c_{jk} = \sum_{i=1}^n a_{ji} b_{ik}$. This sum is exactly the same as the dot product of the j th row of A and the k th column of B . Notice that the length of a row of A must equal the length of a column of B (in other words the number of columns of A and rows of B must be equal). Here are some examples.

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = [1, 2] \cdot \langle 5, 6 \rangle = [17] \quad (\text{a row vector times a column vector is equivalent to the dot product})$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 6 & 1 \end{bmatrix} = \begin{bmatrix} [1 \ 2] \begin{bmatrix} 5 \\ 6 \end{bmatrix} & [1 \ 2] \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ [3 \ 4] \begin{bmatrix} 5 \\ 6 \end{bmatrix} & [3 \ 4] \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 17 & 2 \\ 39 & 4 \end{bmatrix}.$$

$$A = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad B = [3 \ 4] \quad AB = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix} \quad BA = [11] \quad C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 2 \end{bmatrix} \quad CD = \begin{bmatrix} 2 & 1 & 3 \\ 4 & 3 & 5 \end{bmatrix}$$

The product DC is not defined since D has 3 columns but C has only 2 rows.

The identity matrix is a square matrix which has only 1's along the diagonal, and zeros everywhere else. We often use I_n to mean the n by n identity matrix. The identity matrix is like the number 1, in that $AI = A$ and $IA = A$ for any matrix A . If A is a 2 by 3 matrix, then $AI_3 = A$ and $I_2A = A$ (notice that the size of the matrix changes based on the order of multiplication. If A is a square matrix, then $AI = IA = A$).

4 Determinants

Associated with every square matrix is a number, called the determinant, which is related to length, area, volume, and used to generalize these quantities to higher dimensions. Determinants are only defined for square matrices. The determinant of a 2×2 and 3×3 matrix can be computed as follows:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \det \begin{bmatrix} d & f \\ g & i \end{bmatrix} + c \det \begin{bmatrix} d & e \\ g & h \end{bmatrix}$$

Notice the negative sign on the middle term of the 3×3 determinant.

The determinant in general is defined in terms of minors and cofactors. A minor M_{jk} of a matrix A is the determinant of the the matrix formed by removing row j and column k from A . We define a cofactor to be $C_{jk} = (-1)^{j+k} M_{jk}$. The determinant in general is computed using the formula $\sum_{i=1}^n a_{ik} C_{ik}$ or $\sum_{i=1}^n a_{ji} C_{ji}$, where $j, k \leq n$. In other words, you can pick any row or column you want, and then find the determinant by multiplying each entry of that row or column by its cofactor, and then adding the result. The sign matrix on the right keeps track of the $(-1)^{j+k}$ term in the cofactor.

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

sign matrix

All you have to do is determine if the first entry of your expansion has a plus or minus, and then alternate the sign as you expand. The examples below find the determinant by using a cofactor expansion (1) along the top row (so the first term is positive) and (2) along the second column (so the first term is negative).

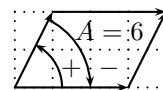
$$\begin{aligned} \det \begin{bmatrix} 1 & 2 & 0 \\ -1 & 3 & 4 \\ 2 & -3 & 1 \end{bmatrix} &= (-1)^{1+1} 1 \det \begin{bmatrix} 3 & 4 \\ -3 & 1 \end{bmatrix} + (-1)^{1+2} 2 \det \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix} + (-1)^{1+3} 0 \det \begin{bmatrix} -1 & 3 \\ 2 & -3 \end{bmatrix} \\ &= 1(3 + 12) - 2(-1 - 8) + 0(3 - 6) = 33 \\ \det \begin{bmatrix} 1 & 2 & 0 \\ -1 & 3 & 4 \\ 2 & -3 & 1 \end{bmatrix} &= (-1)^{1+2} 2 \det \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix} + (-1)^{2+2} 3 \det \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} + (-1)^{3+2} (-3) \det \begin{bmatrix} 1 & 0 \\ -1 & 4 \end{bmatrix} \\ &= -(2)(-1 - 8) + (3)(1 - 0) - (-3)(4 - 0) = 33 \end{aligned}$$

4.1 Geometric Interpretation of the determinant

Consider the 2 by 2 matrix $\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ whose determinant is $3 \cdot 2 - 0 \times 1 = 6$. Draw the column

vectors $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ with their base at the origin. These two vectors give the edges of a parallelogram whose area is the determinant 6. If I swap the order of the two vectors in the

matrix, the the determinant of $\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ is -6 . The reason for the difference is that the determinant not only keeps track of area, but also order. Starting at the first vector, if you can turn counterclockwise through an acute angle to obtain the second vector, then the determinant is positive. If you have to turn clockwise through an acute angle to obtain the second vector, then the determinant is negative. This is often termed “the right hand rule,” as placing your index finger on the first vector and middle finger on the second vector will cause your thumb to point up if the determinant is positive and down if the determinant is negative.



5 Gaussian Elimination - Solving systems of equations

A linear system of equations (such as $2x + y - z = 2$, $x - 2y = 3$, $4y + 2z = 1$) is a system of equations where each variable in the system appears in its own term and is multiplied by at most a constant (called a coefficient). Rather than using x, y, z , we use x_1, x_2, x_3 , so that we can generalize these ideas to any dimension without having to worry about what to call each additional variable. A linear system can be written in terms of vectors sums $x_1 \vec{v}_1 + x_2 \vec{v}_2 + x_3 \vec{v}_3 = \vec{b}$ or in terms of matrix multiplication as $A\vec{x} = \vec{b}$ (where the columns of A are the vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$). The matrix A is called the coefficient matrix of the system of equations. Adding the column vector \vec{b} to the right of A gives what is called an augmented matrix. The linear system at the start of this paragraph can also be expressed in the following 4 ways.

x_1, x_2, x_3 system	vector equation
$\begin{aligned} 2x_1 + x_2 - x_3 &= 2 \\ x_1 - 2x_2 &= 3 \\ 4x_2 + 2x_3 &= 1 \end{aligned}$	$x_1 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$
matrix equation	augmented matrix
$\begin{bmatrix} 2 & 1 & -1 \\ 1 & -2 & 0 \\ 0 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$	$\left[\begin{array}{ccc c} 2 & 1 & -1 & 2 \\ 1 & -2 & 0 & 3 \\ 0 & 4 & 2 & 1 \end{array} \right]$

A linear equation in two variables $ax+by = c$ represents a line. A linear equation in 3 variables $ax+by+cz = d$ represents a plane in space. In a system of equations with only two variables, the system graphically represents multiple lines in the plane. A solution to such a system is the intersection of the lines. In our 3D system above, a solution graphically can be found by finding the intersection of all three planes. The possible solutions are

1. A single point (the ideal solutions)
2. Infinitely many solutions (the three planes meet in a line, or each equation represents the same plane)
3. No solution (two of the planes are parallel and never intersect).

With any system of linear equations, there will always be either a single solution, infinitely many solutions, or no solution. If the system has a solution, it is called consistent. If there is no solution, it is called inconsistent. For a linear system $A\vec{x} = \vec{b}$, we say that the system is a homogeneous system if $\vec{b} = \vec{0}$. If $\vec{b} \neq \vec{0}$, we say the system is nonhomogeneous. You will see these words often throughout the course.

Gaussian elimination is an efficient algorithm used to solve systems of equations. The main idea is to eliminate each variable from all but one row (if possible), using at most three operations:

1. Multiply an equation (or row of a matrix) by a nonzero constant,
2. Add a nonzero multiple of any equation (or row) to another equation,
3. Interchange two equations (or rows).

These three operations are the operations learned in college algebra when solving a system using a method of elimination. Gaussian elimination streamlines elimination methods to solve generic systems of equations of any size. The process involves a forward reduction and (optionally) a backward reduction. The forward reduction creates zero in the lower left corner of the matrix. The backward reduction puts zeros in the upper right corner of the matrix. We eliminate the variables in the lower left corner of the matrix, starting with column 1, then column 2, and proceed column by column until all variables which can be eliminated (made zero) have been eliminated. Before formally stating the algorithm, let's look at a few examples.

We will start with a system of 2 equations and 2 unknowns. Solve

$$\begin{array}{rcl} x_1 - 3x_2 & = & 4 \\ 2x_1 - 5x_2 & = & 1 \end{array} \quad \left[\begin{array}{cc|c} 1 & -3 & 4 \\ 2 & -5 & 1 \end{array} \right]$$

We now subtract twice the first row from the second row.

$$\begin{array}{rcl} x_1 - 3x_2 & = & 4 \\ x_2 & = & -7 \end{array} \quad \left[\begin{array}{cc|c} 1 & -3 & 4 \\ 0 & 1 & -7 \end{array} \right]$$

The matrix at the right is said to be in “row echelon form.” This means that each row is either a zero row or has a 1 (called a leading 1, or a pivot) as its first nonzero entry, and that the leading 1 in a each row must occur further to the right than a leading 1 in the row above. The column containing a pivot is called a pivot column. At this point we can use “back-substitution” to get $x_2 = -7$ and $x_1 = 4 + 3x_2 = 4 - 21 = -17$. Alternatively (and often much easier by hand) we can continue the elimination process by eliminating the terms above each pivot, starting on the right and working backwards. This will result in a matrix where all the pivot columns contain all zeros except for the pivot. If we add 3 times the second row to the first row, we obtain.

$$\begin{array}{rcl} x_1 & = & -17 \\ x_2 & = & -7 \end{array} \quad \left[\begin{array}{cc|c} 1 & 0 & -17 \\ 0 & 1 & -7 \end{array} \right]$$

The matrix on the right is said to be in “reduced row echelon form.” A matrix is in reduced row echelon form if it is in echelon form and each pivot column contains all zeros except for the pivot. You can easily read solutions to systems of equations directly from a matrix which is in “reduced row echelon form.”

Let's solve a nonhomogeneous (the right side is not zero) system with 3 equations and 3 unknowns:

$$\begin{array}{rcl} 2x_1 + x_2 - x_3 & = & 2 \\ x_1 - 2x_2 & = & 3 \\ 4x_2 + 2x_3 & = & 1 \end{array} \quad . \text{ To simplify the writing, we'll just use matrices this time. To keep track of each step,}$$

I will write the row operation next to the row I will replace. Remember that the 3 operations are (1) multiply a row by a nonzero constant, (2) add a multiple of one row to another, (3) interchange any two rows. If I write $R_2 + 3R_1$ next to R_2 , then this means I will add 3 times row 1 to row 2. If I write $2R_2 - R_1$ next to R_2 , then I have done two row operations, namely I multiplied R_2 by 2, and then added (-1) times R_1 to the result (replacing R_2 with the sum). To follow the word below, read left to right, top to bottom. In order to avoid fractions, I wait to divide each row until the last step.

$$\begin{aligned}
 &\Rightarrow \left[\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 1 & -2 & 0 & 3 \\ 0 & 4 & 2 & 1 \end{array} \right] \begin{array}{l} \\ 2R_2 - R_1 \\ \\ \end{array} \Rightarrow \left[\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 0 & -5 & 1 & 4 \\ 0 & 4 & 2 & 1 \end{array} \right] \begin{array}{l} \\ \\ 5R_3 + 4R_2 \end{array} \\
 &\Rightarrow \left[\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 0 & -5 & 1 & 4 \\ 0 & 0 & 14 & 21 \end{array} \right] \begin{array}{l} \\ R_3/7 \\ \\ \end{array} \Rightarrow \left[\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 0 & -10 & 2 & 8 \\ 0 & 0 & 2 & 3 \end{array} \right] \begin{array}{l} 2R_1 + R_3 \\ R_2 - R_3 \\ \\ \end{array} \\
 &\Rightarrow \left[\begin{array}{ccc|c} 4 & 2 & 0 & 7 \\ 0 & -10 & 0 & 5 \\ 0 & 0 & 2 & 3 \end{array} \right] \begin{array}{l} R_2/5 \\ \\ \\ \end{array} \Rightarrow \left[\begin{array}{ccc|c} 4 & 2 & 0 & 7 \\ 0 & -2 & 0 & 1 \\ 0 & 0 & 2 & 3 \end{array} \right] \begin{array}{l} \\ R_1 + R_2 \\ \\ \end{array} \\
 &\Rightarrow \left[\begin{array}{ccc|c} 4 & 0 & 0 & 8 \\ 0 & -2 & 0 & 1 \\ 0 & 0 & 2 & 3 \end{array} \right] \begin{array}{l} R_1/4 \\ R_2/-2 \\ R_3/2 \end{array} \Rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1/2 \\ 0 & 0 & 1 & 3/2 \end{array} \right]
 \end{aligned}$$

Writing the final matrix in terms of a system, we have the solution $x_1 = 2, x_2 = -1/2, x_3 = 3/2$.

The following steps describe the Gaussian elimination algorithm that we used above. Please take a moment to compare what is written below with the example above. Most of the problems in this unit can be solved using Gaussian elimination, so we will practice it as we learn a few new ideas.

1. Forward Phase (row echelon form) - The following 4 steps should be repeated until you have mentally erase all the rows or all the columns. In step 1 or 4 you will erase a column and/or row from the matrix.
 - (a) Consider the first column of your matrix. Interchange rows (if needed) to place a nonzero entry in the first row. If all the elements in the first column are zero, then ignore that column in future computations (mentally erase the column) and begin again with the smaller matrix which is missing this column. If you erase the last column, then stop.
 - (b) Divide the first row (of your possibly smaller matrix) row by it's leading entry so that you have a leading 1. This entry is a pivot, and the column is a pivot column. [When doing this by hand, it is often convenient to skip this step and do it at the very end so that you avoid fractional arithmetic. If you can find a common multiple of all the terms in this row, then divide by it to reduce the size of your computations.]
 - (c) Use the pivot to eliminate each nonzero entry below the pivot, by adding a multiple of the top row (of your smaller matrix) to the nonzero lower row.
 - (d) Ignore the row and column containing your new pivot and return to step one (mentally cover up or erase the row and column containing your pivot). If you erase the last row, then stop.
2. Backward Phase (reduced row echelon form) - At this point each row should have a leading 1, and you should have all zeros to the left and below each leading 1. If you skipped step 2 above, then at the end of this phase you should divide each row by its leading coefficient to make each row have a leading 1.
 - (a) Starting with the last pivot column, use the pivot in that column to eliminate all the nonzero entries above it.
 - (b) Work from right to left, using each pivot to eliminate the nonzero terms above it. Working from right to left will help you reduce the matrix in the least number of steps.

Numerical algorithms written on a computer often find the largest nonzero element in each column before it creates a pivot.

5.1 Reading Reduced Row Echelon Form - rref

If you row reduce a matrix to reduced row echelon form, then you can read the solution immediately from the matrix. Here are some typical examples of what you will see when you reduce a system that does not have a unique solution, together with their solution. The explanations which follow illustrate how to see the solution immediately from the matrix.

$$\begin{array}{ccc} \left[\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right] & \left[\begin{array}{ccc|c} 1 & 0 & 4 & -5 \\ 0 & 1 & -2 & 3 \end{array} \right] & \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right] \\ (1 - 2x_2, x_2, 3) & (-5 - 4x_3, 3 + 2x_3, x_3) & \text{no solution, } 0 \neq 1 \end{array}$$

In the first example, the first and third columns are pivot columns, while the second is not a pivot column. When a column is not a pivot column, we say the corresponding variable is a free variable, because we can pick the free variable to be arbitrary and still obtain a solution to the system. Writing the first matrix in system form, we have $x_1 + 2x_2 = 1, x_3 = 3$, or $x_1 = 1 - 2x_2, x_3 = 3$. If we add to this system the redundant equation $x_2 = x_2$, then we have in vector form $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 - 2x_2 \\ x_2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$. Sometimes we may find it useful to use a different variable, such as $x_2 = t$, and then write the solution in parametric form as $(1 - 2t, t, 3)$, where t is any real number.

In the second example, the first and second columns are pivot columns, so x_3 is a free variable. The solution can be written in the form $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -5 - 4x_3 \\ 3 + 2x_3 \\ x_3 \end{bmatrix} = \begin{bmatrix} -5 \\ 3 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -4 \\ 2 \\ 1 \end{bmatrix}$. Again the last column appears in the solution, and then the opposite of the third column together with a 1 in the third spot gives us the vector which is multiplied by x_3 .

In the third example, the first, second, and fourth columns are pivot columns. Because the last column is a pivot column, we obtain the equation $0 = 1$, which is absurd. Hence the system is inconsistent and has no solution.

The ideas above generalize to higher dimensions. Here are two large rref matrices and their solutions.

$$\begin{array}{ccc} \left[\begin{array}{ccccc|c} 0 & 1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] & \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 4 \end{bmatrix} + x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ -2 \\ -3 \\ 1 \\ 0 \end{bmatrix} & \\ \left[\begin{array}{cccccc|c} 0 & 1 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] & \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ -2 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ -4 \\ 1 \end{bmatrix} & \end{array}$$

Each non pivot column corresponds to one of the vectors in the sum above.

6 Rank and Linear Independence

The rank of a matrix is the number of pivot columns of the matrix. To find the rank of a matrix, you reduce the matrix using Gaussian elimination until you discover the pivot columns.

A linear combination of vectors $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ is an expression of the form $c_1\vec{a}_1 + c_2\vec{a}_2 + \dots + c_n\vec{a}_n$, where c_i is a constant for each i . In other words, a linear combination of vectors is a sum of scalar multiples of the vectors. The span of a set of vectors is all possible linear combinations of the vectors. The span of the vector $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ is vectors of the form $a \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \end{bmatrix}$ which is the x axis, a line through the origin. The span of the vectors $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$ is the set of vectors in 3D of the form $a \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} a & b & 0 \end{bmatrix}$

which is the xy plane, a plane through the origin. Geometrically you can obtain the span of vectors by adding together all possible stretches of the given vectors. The span of a set of vectors will be a line, a plane, or some higher dimensional version of these objects (called a hyperplane) which passes through the origin.

We say that a set of vectors is linearly independent if the only solution to the homogeneous system $c_1\vec{a}_1 + c_2\vec{a}_2 + \dots + c_n\vec{a}_n = \vec{0}$ is the trivial solution $c_1 = c_2 = \dots = c_n = 0$. Otherwise we say the vectors are linearly dependent, and it is possible to write one of the vectors as a linear combination of the others. We say the vectors are dependent because one of them depends on (can be obtained as a linear combination of) the others. An easy way to test if vectors are linearly independent is to create a matrix A where each column represents one of the vectors. The vectors are linearly independent if $A\vec{c} = \vec{0}$ means $\vec{c} = \vec{0}$. Row reduce the matrix to reduced row echelon form. The vectors are linearly independent if and only if each column of A is a pivot column (forcing \vec{c} to be zero). If a column is not a pivot column, then the vector corresponding to that column can be written as a linear combination of the preceding vectors using the coefficients in that column. For example, the vectors $\begin{bmatrix} 1 & 3 & 5 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 3 & 1 \end{bmatrix}$ are linearly independent, as the reduced

row echelon form of $\begin{bmatrix} 1 & -1 & 0 \\ 3 & 0 & 3 \\ 5 & 1 & 1 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, and each column is a pivot column. However, the vectors

$\begin{bmatrix} 1 & 3 & 5 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 6 & 11 \end{bmatrix}$ are linearly dependent, as the reduced row echelon form of $\begin{bmatrix} 1 & -1 & 1 \\ 3 & 0 & 6 \\ 5 & 1 & 11 \end{bmatrix}$

is $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, and column 3 is not a pivot column. This immediately means that the third vector (of the original matrix) is a linear combination of the preceding two. Using the coefficients 2 and 1 from the third column, we can write $\begin{bmatrix} 1 & 6 & 11 \end{bmatrix} = 2\begin{bmatrix} 1 & 3 & 5 \end{bmatrix} + 1\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$.

6.1 Linear Combinations, Spans, and Solutions to Systems via RREF

Solving a system of equations such as $x + 3y = 0, 2x - y = 5$ is equivalent to solving the vector equation $x\begin{bmatrix} 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$. This however is equivalent to the following questions.

- Is $\begin{bmatrix} 0 \\ 5 \end{bmatrix}$ a linear combination of the vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$?
- Is $\begin{bmatrix} 0 \\ 5 \end{bmatrix}$ in the span of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$?

To answer these questions, we reduce the matrix $\begin{bmatrix} 1 & 3 & 0 \\ 2 & -1 & 5 \end{bmatrix}$ to reduced row echelon form.

$$\begin{bmatrix} 1 & 3 & 0 \\ 2 & -1 & 5 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 3 & 0 \\ 0 & -7 & 5 \end{bmatrix} \xrightarrow{3R_2 + 7R_1} \begin{bmatrix} 7 & 0 & 15 \\ 0 & -7 & 5 \end{bmatrix} \xrightarrow{\begin{matrix} R_1/7 \\ R_2/-7 \end{matrix}} \begin{bmatrix} 1 & 0 & 15/7 \\ 0 & 1 & -5/7 \end{bmatrix}.$$

This means that a solution to our system is $x = 15/7, y = -5/7$, and we can write $\begin{bmatrix} 0 \\ 5 \end{bmatrix} = \frac{15}{7}\begin{bmatrix} 1 \\ 2 \end{bmatrix} - \frac{5}{7}\begin{bmatrix} 3 \\ -1 \end{bmatrix}$, which means that the answer to both questions is “Yes.” The example above generalizes to show that a system $A\vec{x} = \vec{b}$ has a solution if and only if \vec{b} is a linear combination of the columns of A , if and only if \vec{b} is in the span of the columns of A .

7 Matrix Inverse

For the rest of this unit, we will assume that the matrix A is a square matrix (in other words we are solving a system where the number of equations and number of unknowns are the same). Recall that the identity matrix I behaves in matrix multiplication like the number 1 behaves in regular multiplication. When we solve the equation $ax = b$ with numbers, we divide both sides by a to obtain $x = \frac{1}{a}b$. We have been studying

linear systems of the form $A\vec{x} = \vec{b}$. It would be nice if we could just divide both sides by A , but there is no such thing as division by a matrix in general. If we look only at square matrices, then sometimes it is possible to find a matrix B such that $BA = AB = I$, the identity matrix. If such a matrix B exists, then multiplying both sides of $A\vec{x} = \vec{b}$ on the left by the matrix B yields $BA\vec{x} = I\vec{x} = \vec{x} = B\vec{b}$. The matrix B is called the inverse of A , and is written A^{-1} . In such cases the solution to $A\vec{x} = \vec{b}$ is $\vec{x} = A^{-1}\vec{b}$.

If an inverse exists, then write $A^{-1} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$. Then the equation $AA^{-1} = I$ requires that the 3 matrix equations $A \begin{bmatrix} c_{11} \\ c_{21} \\ c_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $A \begin{bmatrix} c_{12} \\ c_{22} \\ c_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $A \begin{bmatrix} c_{13} \\ c_{23} \\ c_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ each have a solution, or that $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ are all linear combinations of the columns of A . These three systems can be solved simultaneously by putting the augmented matrix $\left[\begin{array}{ccc|ccc} a_{11} & a_{12} & a_{13} & 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 1 \end{array} \right]$, or $[A \mid I]$, in reduced

row echelon form. The rref form will tell us the coefficients we need to write $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ as linear combinations of the columns of A , and these coefficients will be the entries of A^{-1} . In other words,

an inverse exists of this system reduces to $\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & c_{11} & c_{12} & c_{13} \\ 0 & 1 & 0 & c_{21} & c_{22} & c_{23} \\ 0 & 0 & 1 & c_{31} & c_{32} & c_{33} \end{array} \right]$, or more simply $[I \mid A^{-1}]$ (where the

inverse matrix appears on the left). If the left block of the augmented matrix does not reduce to the identity matrix, then the matrix does not have an inverse. Notice that a matrix has an inverse if and only if the columns of the matrix are linearly dependent. Later we will show that having an inverse is equivalent to having a nonzero determinant and having a rank which equals the number of columns. Remember that this is true for square matrices, or systems where we have the same number of equations as unknowns.

The inverse of the matrix $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is found by row reducing $\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array} \right]$. The reduced row echelon form is $\left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{array} \right]$. The inverse is hence $\begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}$. You can always check your result by

computing $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Using this inverse, the solution to the system $\begin{cases} 1x + 2y = 4 \\ 3x + 4y = 0 \end{cases}$

is $A^{-1} \begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} -8 \\ 6 \end{bmatrix}$.

8 Eigenvalues and Eigenvectors

Let A be a square $n \times n$ matrix. An eigenvector is a nonzero vector \vec{x} such that $A\vec{x} = \lambda\vec{x}$ (matrix multiplication reduces to scalar multiplication) for some scalar λ which is called an eigenvalue. This can be rewritten as $\vec{0} = A\vec{x} - \lambda\vec{x} = A\vec{x} - \lambda I\vec{x} = (A - \lambda I)\vec{x}$. This means that the columns of $A - \lambda I$ are linearly dependent, which means $A - \lambda I$ has no inverse, and it means that $\det(A - \lambda I) = 0$. The expression $\det(A - \lambda I)$ is called the characteristic polynomial of A . It is a polynomial of degree n , and hence there are at most n eigenvalues (which correspond to the zeros of the polynomial). As a way to check your work, the sum of the eigenvalues will always equal the trace of the matrix (the sum of the diagonal elements), and the product of the eigenvalues will always equal the determinant. We will eventually discover why this is true, but for now you can just use these facts to help you check your work.

Let's look at an example. To find the eigenvalues of $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, first subtract λ from each diagonal entry $\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}$, and then find the determinant. Factor to get $(2-\lambda)(2-\lambda) - 1 = \lambda^2 - 4\lambda + 3 = (\lambda-1)(\lambda-3)$. The zeros are 1 and 3, so the eigenvalues are $\lambda = 1, 3$. As a check, the trace of this matrix is $2 + 2 = 4$ and the sum of the eigenvalues is $1 + 3$. In addition, the determinant is $4 - 1 = 3$ which equals the product of the eigenvalues.

When $\lambda = 1$, we compute $A - \lambda I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. The equation $(A - \lambda I)\vec{x} = 0$ is solved using Gaussian Elimination, and row reduction gives the augmented matrix $\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right]$, which means all eigenvectors are of the form $x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ for some $x_2 \neq 0$ (remember you can write $x_1 + x_2 = 0$ and $x_2 = x_2$ since it is a free variable which gives $x_1 = -x_2, x_2 = x_2$). There are infinitely many eigenvectors corresponding to $\lambda = 1$. A particular one is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and all the rest are a linear combination of this one.

When $\lambda = 3$, we compute $A - \lambda I = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$. The equation $(A - \lambda I)\vec{x} = 0$ is solved using Gaussian Elimination, and row reduction gives the augmented matrix $\left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array} \right]$, which means all eigenvectors are of the form $x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ for some $x_2 \neq 0$. A particular eigenvector corresponding to $\lambda = 3$ is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and all others are a linear combination of this one.

Finding eigenvalues and eigenvectors requires that we compute determinants, find zeros of polynomials, and then solve homogeneous systems of equations. You know you are doing the problem correctly if you get infinitely many solutions to the system $(A - \lambda I) = 0$ for each lambda.

Eigenvalues and eigenvectors show up in many different places in engineering, computer science, physics, chemistry, and mathematics. We will be exploring how eigenvectors appear everywhere in linear algebra as the semester progresses. For know it is just crucial that you can find them. In the next unit we will apply them.

As a final example, let's find the eigenvalues and eigenvectors of the matrix $A = \begin{bmatrix} 2 & 1 & 4 \\ 1 & 2 & 4 \\ 0 & 0 & 1 \end{bmatrix}$. Subtracting

λ from the diagonals gives $A = \begin{bmatrix} 2-\lambda & 1 & 4 \\ 1 & 2-\lambda & 4 \\ 0 & 0 & 1-\lambda \end{bmatrix}$. The determinant of this matrix is easiest to compute

if you expand along the third row to get $0 - 0 + (1 - \lambda) \det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}$. This becomes $(1 - \lambda)[(2 - \lambda)(2 - \lambda) - 1] = (1 - \lambda)(\lambda - 1)(\lambda - 3)$, so the eigenvalues are 1, 1, 3 (1 is a repeated root). Their sum is 5 and the trace is $2 + 2 + 1 = 5$ as well.

When $\lambda = 1$, we compute $A - \lambda I = \begin{bmatrix} 1 & 1 & 4 \\ 1 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}$. The equation $(A - \lambda I)\vec{x} = 0$ is solved using Gaussian

Elimination, and row reduction gives the augmented matrix $\left[\begin{array}{ccc|c} 1 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$, which has two free variables.

The solution is $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}$. The eigenvalues are linear combinations of $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}$.

When $\lambda = 3$, we compute $A - \lambda I = \begin{bmatrix} -1 & 1 & 4 \\ 1 & -1 & 4 \\ 0 & 0 & -2 \end{bmatrix}$. The equation $(A - \lambda I)\vec{x} = 0$ is solved using

Gaussian Elimination. Reducing $\left[\begin{array}{ccc|c} -1 & 1 & 4 & 0 \\ 1 & -1 & 4 & 0 \\ 0 & 0 & -2 & 0 \end{array} \right]$ gives the augmented matrix $\left[\begin{array}{ccc|c} -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$, which

has only one free variable. The solution is $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. The eigenvalues are linear combinations of $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.