Properties

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

1. Create examples to illustrate mathematical definitions. Know these definitions.
2. Create examples to illustrate matrix properties, and construct proofs of many of these theorems. You do not have to memorize all the properties, rather when one is given make sure you can construct examples of it.
3. When a square system has a unique solution, this is equivalent to many other facts about the coefficient matrix. Illustrate this equivalence with examples, and prove that these ideas are equivalent.
4. Define a vector space, and be able to show why something is or is not a vector space. Give examples of vector spaces and subspaces we have encountered previously in this class and other classes, and explain how to find a basis for the vector space and its dimension.

Following you will find preparation assignments and suggested homework, and then my best attempt at condensing the information we are learning into a concise set of notes. The text contains possible proofs and explanations of the ideas we are learning. Once we hit vector spaces, we will be shifting the the "Beginning Linear Algebra" textbook, where we will spend most of the remainder of the semester.

1 Preparation and Suggested Homework

	Preparation Problems
Day 1	2, 7, 8, 12
Day 2	16, 24, 21, 23
Day 3	To be announced

The homework in this unit comes from a handout which is stapled to the back of this packet. Your assignment is to do 7 problems per day of class, at least 2 of which included using Maple. When constructing examples, if the computations are time intensive, use Maple to simplify your work (this would be a good use of Maple, so that you can quickly create matrices and see patterns as you let Maple perform complex computations).

2 Definitions, Theorems, and Proofs

Mathematics and technical writing are often written in a style that is different than most prose. Learning to read technical writing takes time and practice, but can be learned. Our main goal in this unit is to becomes more familiar with properties of matrices and vector spaces. Along the way, we will learn some techniques which can be used to help read difficult technical passages.

Mathematical writing often begins with definitions. Those definitions are followed by theorems and proofs. If space permits, examples are often given to illustrate the definitions and theorems. If appropriate, an application is also given to show how a theorem is important. Those who write mathematical papers and textbooks around the nation have become accustomed to writing papers for peers in journals. Most of the time, these papers assume a lot of knowledge from the reader, and because of space limitations the papers are missing lots of examples. In order to read this king of mathematics, you have to learn how to create examples to illustrate definitions and theorems whenever you see them. Similar writing appears in engineering, physics, computer science, economics, and other branches of science.

In this unit, most of the definitions and theorems you have seen and used already in the previous units. The next three subsections present definition, theorems, and then proofs. In the definitions and theorems section, I will provide examples of some of the ideas. Your homework is to create examples to illustrate the other ideas. In the proofs section, I will illustrate a few common techniques that are used to prove theorems. Your homework is to use these techniques to prove many of the remaining theorems. The proofs in linear algebra are often fairly short, and much easier to follow than proofs in other branches of mathematics. Those of you who plan to pursue any type of graduate study in engineering, physics, computer science, or economics, statics, and more will be expected to give proofs of the ideas you discover. In this class, you will gain some skill with simpler proofs.

2.1 Definitions

Whenever you see a new definition, one of the most important things you can do is create an example which illustrates that definition. As you read each definition below, create an example which illustrates that definition. Most of these words are review, a few are new. Make sure you know these definitions.

Definition 1. The **product** of two matrices $A_{mn} = \{a_{ij}\}$ and $B_{np} = \{b_{jk}\}$ is a new matrix $AB = C$ where the ith row and kth column is the dot product of the ith row of A and the kth column of B, or $c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$.

Definition 2. We say that two matrices are **row equivalent** if one can be obtained from the other by a sequence of row operations. The three row operations are (1) interchange two rows, (2) multiply a row by a nonzero constant, (3) add a multiple of one row to another.

Definition 3. The **determinant** of a 1×1 matrix equals the number in the matrix. A **minor** M_{ij} of an $n \times n$ matrix is the determinant of the $(n-1) \times (n-1)$ submatrix that remains after removing the ith row and jth column. A **cofactor** of a matrix A is the product $C_{ij} = (-1)^{i+j} M_{ij}$. The determinant of any square $n \times n$ matrix $(n \geq 2)$ is found by multiplying each entry a_{1i} on the first row of the matrix by its corresponding cofactor, and then summing the result. In other words, the determinant equals $a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n} = \sum_{j=1}^{n} a_{1j}C_{1j}$.

Let's look at an example. Notice how I specifically point out in the example where each cofactor and minor appear. Notice that an alternating sign pattern appears because the $(-1)^{i+j}$ in each cofactor alternates between positive and negative as the column k increases.

$$\begin{aligned}
\det \begin{bmatrix} 1 & 2 & 0 \\ -1 & 3 & 4 \\ 2 & -3 & 1 \end{bmatrix} &= 1C_{11} && +2C_{12} && +0C_{13} \\
&= 1(-1)^{1+1}M_{11} && +2(-1)^{1+2}M_{12} && +0(-1)^{1+3}M_{13} \\
&= 1 \det \begin{bmatrix} 3 & 4 \\ -3 & 1 \end{bmatrix} && -2 \det \begin{bmatrix} -1 & 4 \\ 2 & 1 \end{bmatrix} && +0 \det \begin{bmatrix} -1 & 3 \\ 2 & -3 \end{bmatrix} \\
&= 1(3+12) && -2(-1-8) && +0(3-6) \\
&= 33
\end{aligned}.$$

Definition 4. A **linear combination of vectors** $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is a sum $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$ for some scalars $c_1, c_2, \ldots, c_n$. The **span of a set of vectors** $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is all possible linear combinations of those vectors.

The vector $\begin{bmatrix} 3 & 2 \end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix} 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 \end{bmatrix}$, since $\begin{bmatrix} 3 & 2 \end{bmatrix} = 2\begin{bmatrix} 1 & 0 \end{bmatrix} + 1\begin{bmatrix} 1 & 2 \end{bmatrix}$. The span of the vector $\begin{bmatrix} 1 & 0 \end{bmatrix}$ is the x-axis. The span of $\begin{bmatrix} 1 & 2 \end{bmatrix}$ is a line through the origin that passes through $(1, 2)$. Since every vector in $\mathbb{R}^2$ is a linear combination of $\begin{bmatrix} 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 \end{bmatrix}$, the span of these two vectors is the entire plane $\mathbb{R}^2$. In general, the span of a collection of vectors in 2D and 3D will be a line, plane, or all of space.

Definition 5. We say a set $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ of m-dimensional vectors is **linearly independent** if the only solution to the homogeneous equation $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = \vec{0}$ is the trivial solution $c_1 = c_2 = \cdots = c_n = 0$. If there is a nonzero solution to this homogeneous equation, then we say the set of vectors is **linearly dependent**, or we just say the vectors are linearly dependent (leaving of the words "set of").

Consider the vectors $\vec{v}_1 = \begin{bmatrix} 1 & 3 & 5 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 & 3 & 1 \end{bmatrix}$. The reduced row echelon form of (augment by zero) $\left[\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 5 & 1 & 1 & 0 \end{array}\right]$ is $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$, which means the only solution to $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$ is $c_1 = c_2 = c_3 = 0$. This means the three vectors are linearly independent.

Now consider the vectors $\vec{v}_1 = \begin{bmatrix} 1 & 3 & 5 \end{bmatrix}$ $\vec{v}_2 = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$ $\vec{v}_3 = \begin{bmatrix} 1 & 6 & 11 \end{bmatrix}$. The reduced row echelon form of $\left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 3 & 0 & 6 & 0 \\ 5 & 1 & 11 & 0 \end{array}\right]$ is $\left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$, and column 3 is not a pivot column. Since c_3 is a free variable, there are infinitely many solutions to $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}$, which means the three vectors are linearly dependent. In particular, if we choose $c_3 = -1$ then we have $c_2 = 1$ and $c_1 = 2$ which means $2\vec{v}_1 + 1\vec{v}_2 - 1\vec{v}_3 = \vec{0}$, or $\vec{v}_3 = 2\vec{v}_1 + 1\vec{v}_2$. This last equation shows that the third vector is a linear combination of the preceding two, and we write $\begin{bmatrix} 1 & 6 & 11 \end{bmatrix} = 2\begin{bmatrix} 1 & 3 & 5 \end{bmatrix} + 1\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$.

Definition 6. The **column space** of a matrix is the span of the column vectors. A **basis for the column space** is a set of linearly independent vectors whose span is the column space. The **dimension** of the column space is the number of vectors in a basis for the column space. The **rank** of a matrix is the dimension of the column space.

The reduced row echelon form of $\begin{bmatrix} 1 & -1 & 1 \\ 3 & 0 & 6 \\ 5 & 1 & 11 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. A basis for the column space is the first two columns (the pivot columns). The dimension of the column space is 2, which is the rank of the matrix. The column space is a plane in 3D which passes through the origin and the points $(1, 3, 5)$ and $(-1, 0, 1)$. The 3D vector grapher in maple will draw the column space for you.

Definition 7. The **row space** of a matrix is the span of the row vectors. A **basis for the row space** is a set of linearly independent vectors whose span is the row space. The **dimension** of the row space is the number of vectors in a basis for the row space.

Definition 8. A **diagonal matrix** is a square matrix where the entries off the main diagonal are zero ($a_{ij} = 0$ if $i \neq j$). The **identity matrix** is a diagonal matrix where each entry on the main diagonal is 1 ($a_{ii} = 1$). We often write I to represent the identity matrix, and I_n to represent the $n \times n$ identity matrix when specifying the size is important. An **upper triangular matrix** is a square matrix where every entry below the main diagonal is zero. A **lower triangular matrix** is a square matrix where every entry above the main diagonal is zero.

$$\underset{\text{diagonal}}{\begin{bmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 8 \end{bmatrix}} \quad \underset{\text{identity}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \quad \underset{\text{upper-triangular}}{\begin{bmatrix} 1 & 3 & 2 \\ 0 & 4 & 0 \\ 0 & 0 & -1 \end{bmatrix}} \quad \underset{\text{lower-triangular}}{\begin{bmatrix} 1 & 0 & 0 \\ 3 & 4 & 0 \\ 2 & 0 & -1 \end{bmatrix}}$$

Definition 9. The **inverse** of a square matrix A is a matrix B such that $AB = I$ and $BA = I$, where I is the identity matrix. If a square matrix does not have an inverse, we say the matrix is **singular**.

Definition 10. The **transpose** of a matrix $A_{m\times n} = \{a_{ij}\}$ is a new matrix $A^T_{n\times m}$ where the ith column of A is the ith row of A^T. A **symmetric matrix** is a matrix such that $A = A^T$.

Definition 11. An **eigenvector** of a square matrix A is a nonzero vector $\vec{x}$ with the property that $A\vec{x} = \lambda\vec{x}$ for some scalar λ. We say that $\vec{x}$ is an **eigenvector** corresponding to the **eigenvalue** λ.

2.2 Theorems

The definitions above lead to many properties. The list of properties below will help you keep track of the different properties that we will be using throughout the semester. Each is called a theorem because it is a statement that requires justification (a proof). Many of these properties we have been using without specifically making mention of them, while a few are new. In the homework, your job is to create key examples to illustrate each of these properties. You do not have to memorize all of these properties, rather you should be able to create examples to illustrate them. In the next section, we will prove some of them.

2.2.1 System Properties

Theorem 1. *Every matrix has a unique reduced row echelon form. In particular, this implies that the location of pivot columns is well defined, and row equivalent matrices have the same reduced row echelon form.*

Theorem 2. *A linear equation $A\vec{x} = \vec{b}$ has a solution if and only if the number of pivot columns of A equals the number of pivot columns of the augmented matrix $[A|b]$.*

Theorem 3. *A linear equation has no, one, or infinitely many solutions. In particular, if a consistent linear equation $A\vec{x} = \vec{b}$ has n unknowns and k leading ones, then it has $n - k$ free variables.*

Theorem 4. *For a homogeneous system $A\vec{x} = \vec{0}$, any linear combination of solutions is again a solution.*

Theorem 5. *For a non homogeneous system $A\vec{x} = \vec{b}$, the difference between any two solutions is a solution to the homogeneous system $A\vec{x} = \vec{0}$.*

2.2.2 Matrix Multiplication Properties

Theorem 6. *The product $A\vec{x}$ of a matrix A and column vector x is a linear combination of the columns of A.*

Theorem 7. $(AB)^T = B^T A^T$

Theorem 8. *The product $\vec{x}A$ of a row vector $\vec{x}$ and matrix A is a linear combination of the rows of A.*

Theorem 9. *Consider the matrices $A_{m\times n} = \begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n\end{bmatrix}$ and $B_{n\times p} = \begin{bmatrix}\vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_p\end{bmatrix}$, where $\vec{a}_i$ and $\vec{b}_i$ are the columns of A and B. The product AB equals $AB = \begin{bmatrix}A\vec{b}_1 & A\vec{b}_2 & \cdots & A\vec{b}_p\end{bmatrix}$. In other words, every column of AB is a linear combination of the columns of A.*

Theorem 10. *Every row of AB is a linear combination of the rows of B.*

2.2.3 Linear Independence Properties

Theorem 11. *The rank of a matrix is the number of pivot columns, which equals the number of leading 1's. The columns are linearly independent if and only if every column is a pivot column.*

Theorem 12. *Consider a matrix A whose columns are $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$. Let R be the reduced row echelon form of A with any rows of zeros removed. If $\vec{v}_k$ is not a pivot column, then $\vec{v}_k$ is a linear combination of the preceding vectors. In particular, if C is the matrix obtained from A by erasing the non pivot columns, and $\vec{r}_k$ is the kth column of R, then $C\vec{r}_k = v_k$. In other words, the columns of R tell us how to linearly combine the pivot columns of A to obtain columns of A.*

Theorem 13. *A set (of 2 or more vectors) is linearly dependent if and only if at least one of the vectors can be written as a linear combination of the others. In particular, two vectors are linearly dependent if and only if one is a multiple of the other.*

Theorem 14. *The dimension of the column space (the rank) equals the dimension of the row space. In particular, the pivot columns of a matrix form a basis for the column space and the nonzero row vectors in reduced row echelon form form a basis for the row space.*

Theorem 15. *The system $A\vec{x} = \vec{b}$ is consistent if and only if $\vec{b}$ is in the column space of A.*

2.2.4 Inverse Properties

Theorem 16. *The inverse of a matrix is unique.*

Theorem 17. *The inverse of A^{-1} is A, which means $(A^k)^{-1} = (A^{-1})^k$ and $(cA)^{-1} = \frac{1}{c}A^{-1}$.*

Theorem 18. $(A^T)^{-1} = (A^{-1})^T$.

Theorem 19. $(AB)^{-1} = B^{-1}A^{-1}$.

Theorem 20. *The solution to $A\vec{x} = \vec{b}$ is $\vec{x} = A^{-1}\vec{b}$ if A is invertible.*

2.2.5 Determinant Properties

Theorem 21. *The determinant can be computed by using a cofactor expansion along any row or any column.*

Theorem 22. *The determinant of a triangular matrix is the product of its entries along the main diagonal.*

Theorem 23. *The determinant of AB is $|AB| = |A||B|$. (Postpone the proof of this one.)*

Theorem 24. $|A^{-1}| = |A|^{-1}$*, and* $|A^T| = |A|$

Theorem 25. *A matrix is invertible if and only if its determinant is not zero.*

2.2.6 Eigenvalues and Eigenvectors

Theorem 26. *If λ is an eigenvalue of A and $\vec{x}$ is an eigenvector, then $(A - \lambda I)\vec{x} = \vec{0}$, and* $\det(A - \lambda I) = 0$.

Theorem 27. *For an n by n matrix, the characteristic polynomial is an nth degree polynomial. This means there are n eigenvalues (counting multiplicity and complex eigenvalues).*

Theorem 28. *For a triangular matrix (all zeros either above or below the diagonal), the eigenvalues appear on the main diagonal.*

Theorem 29. *The eigenvalues of A^T are the same as the eigenvalues of A.*

Theorem 30. *If the sum of every row of A is the same, then that sum is an eigenvalue corresponding to the eigenvector* $(1, 1, \ldots, 1)$*. If the sum of every column of A is the same, then that sum is an eigenvalue of A as well (because it is an eigenvalue of A^T). This is why a Markov process always has 1 as an eigenvector.*

2.3 Proof Techniques - An introduction to proving theorems.

In this section I will illustrate various methods of justifying a theorem (giving a proof). You do not have to be able to prove every one of the theorems from the previous section, rather focus on learning to prove the ones I give here, and the ones I refer to in the homework.

2.3.1 Prove by using a definition

To prove theorem 19 ($(AB)^{-1} = B^{-1}A^{-1}$), we just use the definition of an inverse. It often helps to rewrite a definition. The inverse of a matrix C is a matrix D such that $CD = DC = I$ (multiplying on either side gives the identity). We need to show that multiplying AB on the both the left and right by $B^{-1}A^{-1}$ results in the identity. We multiply AB on the left by $B^{-1}A^{-1}$ and compute

$$(B^{-1}A^{-1})(AB) = B^{-1}A^{-1}AB = B^{-1}IB = BB^{-1} = I.$$

Similarly we multiply AB on the right by $B^{-1}A^{-1}$ and compute

$$(AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I.$$

Because multiplying AB on both the left and right by $B^{-1}A^{-1}$ gives us the identity matrix, we know $B^{-1}A^{-1}$ is the inverse of AB. All we did was use the definition of an inverse.

2.3.2 Proving something is unique

To show something is unique, one common approach is to assume there are two possible answers and then show the answers must be the same. We'll do this to show that inverses are unique (Theorem 16). Suppose A is a square matrix with an inverse B and an inverse C. This means that $AB = BA = I$, and $AC = CA = I$. I now need to show that $B = C$. If we multiply both sides of the equation $AB = I$ on the left by C, then we obtain $C(AB) = CI$. Since matrix multiplication is associative, we can rearrange the parentheses on the left to obtain $(CA)B = C$. Since C is an inverse for A, we know $CA = I$, which means $IB = C$, or $B = C$. Hence inverses are unique.

2.3.3 Prove by using another theorem

Sometimes using others theorem can quickly prove a new theorem. Theorem 8 says that the product $\vec{x}A$ is a linear combination of the rows of A. The two theorems right before this one state that $A\vec{x}$ is a linear combination of the columns of A, and that $(AB)^T = B^TA^T$. Rather than prove Theorem 8 directly, we are going to use these other two theorems. The transpose of $\vec{x}A$ is $A^T\vec{x}^T$, which is a linear combination of the columns of A^T. However, the columns of A^T are the transposed rows of A, so $A^T\vec{x}^T$ is a linear combination of the transposed rows of A. This means that $\vec{x}A$ (undoing the transpose) is a linear combination of the non transposed rows of A. Theorems which give information about columns often immediately give information about rows as well.

2.3.4 If and only if

The words "if and only if" require that an "if-then" statement work in both directions. The following key theorem (13) has an if and only if statement in it. A set of 2 or more vectors is linearly dependent if and only if one of the vectors is a linear combination of others. To prove this, we need to show 2 things.

1. If the vectors are linearly dependent, then one is a linear combination of the others.
2. If one of the vectors is a linear combination of the others, then the vectors are linearly independent.

We'll start by proving the first item. If the vectors are dependent, then there is a nonzero solution to the equation $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = \vec{0}$. Pick one of the nonzero constants, say c_k. Subtract this term to the other side and divide by $-c_k$ (which isn't zero) to obtain

$$\frac{c_1}{-c_1}\vec{v}_1 + \cdots + \frac{c_{k-1}}{-c_k}\vec{v}_{k+1} + \frac{c_{k+1}}{-c_k}\vec{v}_{k+1} + \cdots + \frac{c_n}{-c_k}\vec{v}_n = \vec{v}_k,$$

which means that v_k is a linear combination of the other vectors.

Now let's prove the reverse condition. Suppose one of the vectors is a linear combination of the others (suppose it is the kth). Write $v_k = c_1\vec{v}_1 + \cdots + c_{k-1}\vec{v}_{k+1} + c_{k+1}\vec{v}_{k+1} + \cdots + c_n\vec{v}_n$. Subtracting v_k from both sides gives a nonzero solution ($c_k = -1$) to the equation $\vec{0} = c_1\vec{v}_1 + \cdots + c_{k-1}\vec{v}_{k+1} - v_k + c_{k+1}\vec{v}_{k+1} + \cdots + c_n\vec{v}_n$, which means the vectors are linearly dependent.

2.3.5 Proof by contradiction

Sometimes in order to prove a theorem, we make a contradictory assumption and then show that this assumption lead to an error (which means the assumption cannot be correct). We will do this to show that if A is invertible, then $|A| \neq 0$. We will also use the fact that the determinant of a product is the product of the determinants ($|AB| = |A||B|$). We are going to start with a matrix A which has an inverse and then assume that $|A| = 0$. The inverse of A is A^{-1} and $AA^{-1} = I$. Taking determinants of both sides gives $|AA^{-1}| = 1$, or $|A||A^{-1}| = 1$. If $|A| = 0$, then $0 = 1$. Since this is not possible, we must have $|A| \neq 0$.

2.3.6 Prove $a = b$ by showing $a \leq b$ and $b \leq a$

If $a \leq b$ and $b \leq a$, then $a = b$. Sometimes it is easier to show an inequality between two numbers than it is to show the numbers are equal. We will use this to show that the rank of a matrix (the dimension of the column space) equals the dimension of the row space. Let the rank of A equal r, and let the dimension of the row space equal s. Write $A = CR$ where C is the pivot columns of A and R is the nonzero rows of the RREF of A. The rank r equals the number of columns in C. Since every row of A can be written as a linear combination of the rows of R, we know that a basis for the row space cannot contain more than r vectors, or $s \leq r$. We now repeat the previous argument on the transpose of A. Write $A^T = C'R'$ where C' is the pivot columns of A^T and R' is the nonzero rows of the RREF of A^T. The columns of C' are the linearly independent rows of A, so C' has s columns, and R' has s rows. Since every row of A^T can be written as a linear combination of the rows of R', a basis for the row space of A^T (i.e. the column space of A) cannot have more than s vectors, or $r \leq s$. Since $r \leq s$ and $s \leq r$, we have $r = s$.

3 A huge equivalence - Connecting all the ideas

Theorem 31. *For an n by n matrix A, the following are all equivalent (meaning you can put an if and only if between any two of these statements):*

1. *The system $A\vec{x} = \vec{0}$ has only the trivial solution.*
2. *The system $A\vec{x} = \vec{b}$ has a unique solution for every $\vec{b}$.*
3. *The reduced row echelon form of A is the identity matrix.*
4. *Every column of A is a pivot column.*
5. *The columns of A are linearly independent. (The dimension of the column space is n, and the span of the columns of A is $\mathbb{R}^n$.)*
6. *The rank of A is n.*
7. *The rows of A are linearly independent. (The dimension of the row space is n, and the span of the rows of A is $\mathbb{R}^n$.)*
8. *The vectors $\vec{e}_1 = (1,0,0,\dots,0), \vec{e}_2 = (0,1,0,\dots,0), \dots, \vec{e}_n = (0,0,0,\dots,1)$ are in the column space of A.*
9. *A has an inverse.*
10. *$|A| \neq 0$.*
11. *The eigenvalues of A are nonzero.*

We have been using this theorem periodically without knowing it. Because this theorem is true, there is an if and only if between any pair of items. In addition, because there is an if and only if between every item, we can also negate every single statement and obtain an equivalence as well:

Theorem 32. *If $A\vec{x} = 0$ has more than one solution, then the system $A\vec{x} = \vec{b}$ does not have a unique solution (it may not have one at all), the reduced row echelon form of A is not the identity matrix, at least one column will not be a pivot column, the columns will be linearly dependent (the dimension of the column space will be less than n), the rank will be less than n, the rows will be linearly dependent, there will not be an inverse (A is singular), the determinant will be zero, and zero will be an eigenvector.*

The results of this theorem involve 2 results (an if and only if) between every pair of statements (making 110 different theorems altogether). We can prove the theorem much more quickly. If we want to show 4 things are equivalent, then we can show that (1) implies (2), (2) implies (3), (3) implies (4), and (4) implies (1). Then we will have shown that (1) also implies (3) and (4) because we can follow a circle of implications to obtain an implication between any two statements. All we have to do to prove the theorem above is show a chain of 11 implications that eventually circles back on itself.

Proof. [(1) implies (2)] If the system $A\vec{x} = \vec{0}$ has only the trivial solution, then the reduced row echelon form of $[A|\vec{0}]$ is $[I|\vec{0}]$. The row operations used to reduce this matrix will be the same as the row operations used to reduce $[A|\vec{b}]$. Hence the RREF of $[A|\vec{b}]$ will be $[I|\vec{x}]$, where $A\vec{x} = \vec{b}$, so there will be a unique solution.

[(2)implies (3)] With a unique solution to $A\vec{x} = \vec{b}$ for any $\vec{b}$, the RREF of $[A|\vec{0}]$ is $[I|\vec{0}]$. Since removing the last column of zeros won't change any row operations, the RREF of A is I.

[(3) implies (4)] Since the reduced row echelon form of A is I, then each column contains a leading 1, and hence is a pivot column

[(4) implies (5)] Since each column is a pivot column, the only solution to $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = \vec{0}$ (where $\vec{v}_i$ is the ith column of A) is the trivial solution $c_1 = c_2 = \cdots = c_n = 0$. This is because in order for each column of A to be a pivot column, each row of A must have a leading 1 (since there are n rows as well as n columns). Notice that we have just shown (4) implies (1) as well, which means that (1) through (4) are all equilvalent.

[(5) implies (6)] The definition of rank is the dimension of the column space. Since there are n pivot columns, the rank is n.

[(6) implies (7)] This is the result of Theorem 14, which also shows that (7) implies (5).

[(7) implies (8)] Since (7) is equivalent to (5), we know that the span of the column space is all of $\mathbb{R}^n$. This means in particular the the vectors listed are in the span of the columns of A.

[(8) implies (9)] To find the inverse of A, we need to solve $AB = I$ or $[A\vec{b}_1\ A\vec{b}_2\ \ldots A\vec{b}_n] = [\vec{e}_1\ \vec{e}_2\ \ldots \vec{e}_n]$. This means we need to solve the equation $A\vec{b}_i = \vec{e}_i$ for $\vec{b}_i$ for each i, and this can be done because each $\vec{e}_i$ is in the column space of A. The inverse of A is thus $[\vec{b}_1\ \vec{b}_2\ \ldots \vec{b}_n]$.

[(9) implies (10)] We already showed that if a matrix has an inverse, then the determinant is not zero (otherwise $|A||A^{-1}| = |I|$ gives $0 = 1$).

[(10) implies (11)] If the determinant is not zero, then $|A - 0I| \neq 0$, which means 0 is not an eigenvalue.

[(11) implies (1)] If zero is not an eigenvalue, then the equation $A\vec{x} = 0\vec{x}$ cannot have a nonzero solution. This means that the only solution is the trivial solution. □

At this point, if we want to say that (10) implies (9) (determinant nonzero implies there is an inverse), then the proof above shows that (10) implies (11) which implies (1) which implies (2) and so on until it implies (9) as well. By creating a circle of implications, all of the ideas above are equivalent.

In the homework, I ask you to illustrate this theorem with a few matrices. Your job is to determine show that the matrix satisfies every single one of the statements above, or to show that it satisfies the opposite of every single statement above.

4 Vector Spaces

The notion of a vector space is the foundational tool upon which we will build for the rest of the semester. Vectors in the plane, vectors in space, and planes in space through the origin are all examples of vector spaces. However, they encompass much more. Continuous and differentiable functions form vector spaces. Matrices of the same size form a vector space. Mathematicians discovered some common properties about all of these kinds of objects, and created the word "Vector Space" to allow them to focus on the key commonalities. The results which sprang from the following definition created linear algebra as we know it today.

Definition 12. A vector space is a set V together with two operations:

1. $+$ vector addition assigns to any pair $\vec{u}, \vec{v} \in V$ another vector $\vec{u} + \vec{v} \in V$,
2. $\cdot$ scalar multiplication assigns to any $\vec{v} \in V$ and $c \in \mathbb{R}$ another vector $c\vec{v}\vec{V}$.

These two operations satisfy the following axioms ($c, d \in \mathbb{R}$ and $u, v, w \in V$):

(A_1) Vector addition is associative: $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$.

(A_2) Vector addition is commutative: $\vec{u} + \vec{v} = \vec{u} + \vec{v})$.

(A_3) There is a zero vector $\vec{0}$ in V which satisfies $\vec{0} + \vec{v} = \vec{v} + \vec{0} = \vec{v}$.

(A_4) Every $\vec{v} \in V$ has an additive inverse, called $-\vec{v}$, which satisfies $\vec{v} + (-\vec{v}) = \vec{0}$.

(M_1) Scalar multiplication distributes across vector addition: $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$.

(M_2) Scalar multiplication distributes across scalar addition: $(c + d)\vec{v} = c\vec{v} + d\vec{v}$.

(M_3) Scalar multiplication is associative: $(cd)\vec{v} = c(d\vec{v})$

(M_4) Scalar multiplication by 1 does nothing: $1\vec{v} = \vec{v}$

(In the book, they talk about a field K. Just ignore this and replace it with $\mathbb{R}$ for now. Scalar multiplication does not have to be done over the real numbers (it could be done over the complex numbers). For simplicity we will stick to the real numbers for now.)

The definition above is the full technical definition of a vector space. The definition was created to generalize properties we already use about vectors in 2 and 3 dimensions. When working with a vector space, this definition essentially means you can add and multiply as you would normally expect, no surprises will occur.

The definition also is created so that the span of any set of vectors will always be a vector space. Essentially, in a vector space every linear combination of the vectors in that space must also be in the space. This means that lines and planes in 3D which pass through the origin are examples of vector spaces. As the semester progresses, we will discover that our intuition about what happens in 2D and 3D generalizes to all dimensions, and can help us understand any kind of vector space. For now, let's look at some examples of vector spaces, and explain why they are vector spaces.

1. Examples of vector spaces
   (a) Both $\mathbb{R}^2$ and $\mathbb{R}^3$ are vector spaces. The collection of all n dimensional vectors, $\mathbb{R}^n$, is a vector space. Vector addition is defined component wise. Addition is associative, commutative, the zero vector is $(0, 0, \ldots, 0)$, and $(v_1, v_2, \ldots, v_n) + (-v_1, -v_2, \ldots, -v_n) = \vec{0}$. Also, scalar multiplication distributes across vector and scalar addition, it is associative, and $1\vec{v} = \vec{v}$. Essentially the definition of a vector space was defined to capture the properties of $\mathbb{R}^n$ which are useful in other settings.
   (b) The set of m by n matrices, written M_{mn}, with matrix addition is a vector space.
   (c) The set of all polynomials, written $P(x)$, is a vector space.
   (d) The set of continuous functions on an interval $[a, b]$, written $C[a, b]$, with function addition is a vector space.
   (e) The set of continuously differentiable functions on an interval (a, b), written $C^1(a, b)$, with function addition is a vector space.
2. Examples that are not vector spaces
   (a) A line that does not pass through the origin. The zero vector is not on the line.
   (b) Polynomials of degree n. There is no zero vector.
   (c) Invertible Matrices. There is no zero matrix.
   (d) The nonnegative x axis. There is no additive inverse for $(1, 0)$, as $(-1, 0)$ is not in the space.
   (e) The line segment from $(-1, 0)$ to $(1, 0)$. The product $2(1, 0) = (2, 0)$ is not in the space.

4.1 Subspaces

The notion of a subspace allows us to quickly find new vector spaces inside of some we already understand.

Definition 13. Suppose V is a vector space. If U is a set of vectors in V such that U is a vector space itself (using the addition and multiplication in V), then we say that U is a subspace of V and write $U \subset V$.

As an example, we know that the vectors in the plane form a vector space. Every line through the origin forms a vector subspace of the plane. Every line and plane in 3D which passes through the origin forms a vector subspace of $\mathbb{R}^3$.

Theorem 33. *Suppose U is a set of vectors in a vector space V. If*

1. 0
   inU
2. $\vec{u} + \vec{v} \in U$ *for every* $\vec{u}, \vec{v} \in U$
3. $c\vec{u} \in U$ *for every* $\vec{u} \in U$ *and* $c \in \mathbb{R}$

Then U is a vector subspace of V. In other words, if every linear combination of vectors in U is also in U, then U is a vector subspace of V.

The key value of this theorem is that it allows us to quickly show that many spaces are vector spaces. Let's start with some examples, and then we'll end with a proof.

1. The span of a set of vectors will always be a subspace of the overriding vector space. This is because every linear combination is by definition in the subspace. The key use we will focus on for now is that the row space and column space are both vector subspaces (as they are defined as spans).
2. The set of upper triangular m by n matrices, is a subspace of M_{mn}. So are symmetric matrices.
3. The set of polynomials of degree less than or equal to n, written $P_n(x)$, is a subspace of $P(x)$.
4. More examples are given on page 160 (examples 4.4,4.5,4.6) in your Lipschutz book.

Proof. The proof of this theorem is rather quick. We are assuming that vector addition and scalar multiplication give us vectors in U. Because every vector in U is also a vector in V, we immediately know that vector addition is associative (A_1) and commutative (A_2), and that scalar multiplication distributes across vector (M_1) and scalar addition (M_2), is associative (M_3), and multiplication by 1 does nothing (M_4). By assumption there is a zero vector (A_3). We just need to show there is are additive inverses (A_4). If $\vec{u} \in U$, then because $c\vec{u} \in U$ for every c we can pick $c = -1$ to obtain $-1\vec{u} \in U$. This means that $u + (-1\vec{u}) = (1-1)\vec{u} = 0\vec{u} = \vec{0}$, which means every $u \in U$ has an additive inverse $-1\vec{u} \in U$ (A_4). This complete the proof. □

4.2 Basis and Dimension

To introduce the words basis and dimension, let's start with an example. This example was shown next to the definition of column space in the first section. The reduced row echelon form of $\begin{bmatrix} 1 & -1 & 1 \\ 3 & 0 & 6 \\ 5 & 1 & 11 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. The reduced row echelon form tells us that the third column is 2 times the first plus the second. The column space is the vector space spanned by the column vectors $\vec{u}_1 = \begin{bmatrix} 1 & 3 & 5 \end{bmatrix}^T$, $\vec{u}_2 = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}^T$, and $\vec{u}_3 = \begin{bmatrix} 1 & 6 & 11 \end{bmatrix}^T$. Since the third column is already a linear combination of the first two columns, it does not contribute any new vectors to the span of just the first two vectors. Hence, the column space is the span of $\vec{u}_1$ and $\vec{u}_2$. Because these vectors are linearly independent, I cannot remove either without reducing the number of vectors in their span. Hence we say the vectors $\vec{u}_1$ and $\vec{u}_2$ form a basis for the column space. A basis for the column space is a minimal set of vectors whose span is the the column space. Because $\vec{u}_1$ and $\vec{u}_2$ are linearly independent, and they span the column space, they form a basis for the column space. The dimension of the column space is the number of vectors in a basis for the column space. To find a basis, I just reduce the matrix and pick out the pivot columns. (The 3D vector grapher in the Technology Introduction will help you visualize this).

Definition 14. A **basis for a vector space** is a set of linearly independent vectors whose span is the vector space. The **dimension of a vector space** is the number of vectors in a basis for the vector space. We say the dimension of the zero vector space is zero.

Let's look at a few more examples.

1. The standard basis for $\mathbb{R}^2$ is $\vec{e}_1 = [1,0], \vec{e}_2 = [0,1]$, so the dimension is 2. Similarly, $\vec{e}_1 = [1,0,\ldots,0], \vec{e}_2 = [0,0,\ldots,0], \ldots, \vec{e}_2 = [0,0,\ldots,1]$ is a basis for $\mathbb{R}^n$, so the dimension is n.
2. The span of the polynomials $\{1, x, x^2\}$ is $a + bx + cx^2$, which is every polynomial of degree less than or equal to 2. Hence these polynomials form a basis for $P_2(x)$. Notice that dimension of $P_2(x)$ is 3 (one more than the degree of the polynomial).
3. The polynomials $1 + x, x + x^2$ span a 2 dimensional vector subspace of $P_2(x)$.

When working with vector spaces that are not $\mathbb{R}^n$, the easiest way to deal with them is to purposefully create a correspondence between vectors in that space and column vectors in $\mathbb{R}^n$. For example, the vector space $P_2(x)$ consists of all polynomials of the form $a + bx + cx^2$. If we just use $[a, b, c]^T$ to keep track of the coefficients, then it is easy to see how $P_2(x)$ has dimension 3, and the polynomials $1 = [1,0,0]^T, x = [0,1,0]^T, x^2 = [0,0,1]^T$ form a basis for $P_2(x)$.

Using the correspondence above, show the vectors $1 = [1,0,0]^T, 1+x = [1,1,0]^T, 1+x+x^2 = [1,1,1]^T$ form a basis for $P_2(x)$, and write $2+3x-4x^2 = [2,3,-4]$ as a linear combination of these vectors. To do this, we just reduce the augmented matrix $\begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & -4 \end{bmatrix}$ to $\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & -4 \end{bmatrix}$. Since the first 3 columns are pivot columns, they are independent (so the 3 polynomials are independent). The last column of RREF gives us the following linear combination $2+3x-4x^2 = -1(1) + 7(1+x) - 4(1+x+x^2)$.

As a final example, let's show how we can create this correspondence when working with matrices. The matrices $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ form a basis for the vector space M_{22} of 2 by 2 matrices whose dimension is 4. Any matrix $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ in this space can be written in the form $[a,b,c,d]^T$. This allows us to work with 2 by 2 matrices as a vector space by considering column vectors in $\mathbb{R}^4$. Whenever we encounter vector spaces of finite dimension, we will almost alway convert them to column vectors of $\mathbb{R}^n$ before doing any calculations. This means that once we understand what happens in $\mathbb{R}^n$, we can understand what happens in any vector space of dimension n.