SPRING 2024: MATH 590 DAILY UPDATE

Monday, April 29. The first fourteen minutes of class were devoted to Quiz 12. The remainder of the class period was spent working on practice problems for the final exam.

Friday, April 26. We continued our discussion of the JCF for 3×3 non-diagonalizable matrices A focusing on the case that $p_A(x) = (x - \lambda)^3$. We saw that this lead to two possible JCFs, namely $\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$, when $\dim(E_\lambda) = 2$, or $\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$, when $\dim(E_\lambda) = 1$. We then worked through an example of each case. When $\dim(E_\lambda) = 2$, the process was as follows: (i) Find $v_2 \notin E_\lambda$, and set $v_1 := (A - \lambda I)v_2$ (which will be an eigenvector for λ). Then choose v_3 in E_λ not a multiple of v_1. Upon letting P be the 3×3 matrix whose columns are v_1, v_2, v_3 we saw that $P^{-1}AP = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$.

The second example was a 3×3 matrix with $p_A(x) = (x - \lambda)^3$ and $\dim(E_\lambda) = 1$. We first calculated $(A - \lambda I)^2$ and took a vector v_3 such that $(A - \lambda I)^2 v_3 \neq 0$. We then set $v_2 := (A - \lambda I)v_3$ and $v_1 = (A - \lambda I)v_2$. Upon doing so, we found that if P is the matrix whose columns are v_1, v_2, v_3, then $P^{-1}AP = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$.

Wednesday, April 24. We continued our discussion of the JCF, reviewing the definition and both, the form of the JCF we can put the matrix into and the two facts (a) and (b) from the previous lecture. We then easily observed that if A is a non-diagonalizable 2×2 matrix, the JCF must be of the form $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. We then found the JCF of the 2×2 matrix $\begin{pmatrix} 0 & -9 \\ 1 & 6 \end{pmatrix}$, as well as the change of basis matrix P by following the steps: (i) Find λ; (ii) Calculate $\dim(E_\lambda)$; (iii) Find v_2 such that $v_2 \notin E_\lambda$; (iv) Take $v_1 := (A - \lambda I)v_2$ and $P = [v_1\ v_2]$.

We next discussed in detail the case that A is a non-diagonalizable 3×3 matrix with $p_A(x) = (x - \lambda_1)^2(x - \lambda_2)$, so that $\dim(E_{\lambda_1}) = 1$. In this case, facts (a) and (b) from the previous lecture immediately show that the JCF of A has the form $\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$. In this case, to find P, the change of basis matrix with columns v_1, v_2, v_2, one must: (i) Find a vector v_2 such that $(A - \lambda_1 I)^2 v_2 = 0$, but v_2 not an eigenvector for λ_1, so that $(A - \lambda_1 I)v_2 \neq 0$; (ii) Set $v_1 := (A - \lambda_1 I)v_2$; (iii) Take v_3 any eigenvector of λ_1. We then showed why this works. First, we have $(A - \lambda_1 I)v_1 = (A - \lambda_1 I)^2 v_2 = \vec{0}$, which shows that $Av_1 = \lambda_1 v_1$. Then, by definition, $v_1 = (A - \lambda_1 I)v_2$, so that $Av_2 = v_1 + \lambda_1 v_2$. And $Av_3 = \lambda_2 v_3$. Thus,

$$A \cdot [v_1\ v_2\ v_3] = [Av_1\ Av_2\ Av_3] = [\lambda_1 v_1\ \ v_1 + \lambda_1 v_2\ \ \lambda_2 v_3] = [v_1\ v_2\ v_3] \cdot \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

Once one shows that $P = [v_1\ v_2\ v_3]$ are linearly independent, then P is invertible, and $P^{-1}AP = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$.

Added after Bonus Problem 6 was turned in. To see that v_1, v_2, v_3 are linearly independent, suppose $av_1 + bv_2 + cv_3 = \vec{0}$. Apply the matrix $(A - \lambda_1 I)^2$ to both sides of this equation to get:

$$a(A-\lambda_1 I)^2 v_1 + b(A-\lambda_1 I)^2 v_2 + c(A-\lambda_1 I)v_3 = \vec{0}.$$

By definition, $(A - \lambda_1 I)v_2 = 0$. And $(A - \lambda_1 I)^2 v_1 = \vec{0}$, because we showed above that $(A - \lambda_1 I)v_1 = \vec{0}$. Thus, $c(A - \lambda_1 I)v_2 = \vec{0}$. However,

$$\begin{aligned}(A-\lambda_1 I)^2 v_2 &= (A-\lambda_1 I)\{(A-\lambda_1 I)v_2\}\\ &= (A-\lambda_1 I)\{Av_2 - \lambda_1 v_2\} = (A-\lambda_1 I)\{(\lambda_2-\lambda_1)v_2\}\\ &= (\lambda_2-\lambda_1)(A-\lambda_1 I)(v_2) = (\lambda_2-\lambda_1)(\lambda_2-\lambda_1)v_2.\end{aligned}$$

Thus, from $c(A - \lambda_1 I)v_2 = \vec{0}$, we have $c(\lambda_2 - \lambda_1)^2 v_2 = 0$. Since $\lambda_2 - \lambda_1 \neq 0$, we must have $c = 0$. Thus, $av_1 + bv_2 = 0$. Now apply $A - \lambda_1 I$ to this equation. Since $(A - \lambda_1 I)v_1 = \vec{0}$ and $(A - \lambda_1)v_2 \neq \vec{0}$, we have $b(A - \lambda_1 I)v_2 = \vec{0}$ which implies $b = 0$. Finally, this leaves $av_1 = \vec{0}$, so $a = 0$. Therefore, v_1, v_2, v_3 are linearly independent.

Monday, April 22. The first fifteen minutes of class were devoted to Quiz 11, followed by seven minutes allotted for the online course survey. We then began our discussion of the *Jordan Canonical Form* (JCF) for linear transformations and matrices. We noted that the JCF always exists when the transformation or matrix in question has all of its eigenvalues in $F = \mathbb{R}$ or $\mathbb{C}$. In particular, the JCF always exists when working over $\mathbb{C}$. For the scalar λ, we then defined the *Jordan block* of size s, $J(\lambda, s)$, to be the $s \times s$ matrix with λ down the diagonal, 1s on the diagonal above the main diagonal and 0s elsewhere. So for example, when $s = 3$, we have the Jordan block $\begin{pmatrix}\lambda & 1 & 0\\ 0 & \lambda & 1\\ 0 & 0 & \lambda\end{pmatrix}$. We then stated the following theorem:

Jordan Canonical Form Theorem. Let $T : V \to V$ be a finite dimensional vector space over $F = \mathbb{R}$ or $\mathbb{C}$. If $F = \mathbb{R}$, assume that $p_T(x)$ has it roots in $\mathbb{R}$. Then there exists a basis $\alpha \subseteq V$ such that $[T]_\alpha^\alpha = J$, where $J = \begin{pmatrix}J_1 & 0 & \cdots & 0\\ 0 & J_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & J_k\end{pmatrix}$ is block diagonal and each block J_i is a Jordan block. Equivalently, if A is an $n \times n$ matrix over F and $p_A(x)$ has its roots in F, then there is an invertible matrix P, with entries in F, such that $P^{-1}AP = J$, for J as above. In each case, we call J the *Jordan canonical form* of T or A.

We then noted the following for J the JCF of T or A:

(i) All of the eigenvalues appear among the entries of the J_i and the same eigenvalue can appear more that one J_i
(ii) We can assume Jordan blocks with the same eigenvalue are adjacent in the matrix J
(iii) We can assume the Jordan blocks associated with the same value appear in decreasing size.
(iv) We call the submatrix consisting of all Jordan blocks associated to a given eigenvalue λ the *Jordan box* associated with λ.

We also recorded the following facts that completely determine the JCF for 2×2 and 3×3 matrices. Let A be an $n \times n$ matrix over F (so that A might be $[T]_\beta^\beta$ for some basis $\beta \subseteq V$). Suppose $p_A(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$.

(a) The Jordan box associated to each λ_i is an $e_i \times e_i$ matrix.
(b) The number of Jordan blocks in the Jordan box associated to λ_i is $\dim(E_{\lambda_i})$.

We ended class showing that the JCF of the matrix $\begin{pmatrix}2 & -1\\ 1 & 0\end{pmatrix}$ is the matrix $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$.

Friday, April 19. After recalling the definition of the adjoint of an $n \times n$ complex matrix, A, namely $A^* = (\overline{A})^t = \overline{A^t}$, we discussed (but did not prove) the following properties of the adjoint:

Properties of the adjoint. Let A be a complex $n \times n$ matrix.

(i) $(A^*)^* = A$.
(ii) $(AB)^* = B^*A^*$.
(iii) $\langle Av, w\rangle = \langle v, A^*w\rangle$, for all $v, w \in \mathbb{C}^n$.
(iv) A^*A and AA^* are self-adjoint.
(v) The following are equivalent:
  (a) $AA^* = I_n$.
  (b) $A^*A = I_n$.
  (c) The columns of A form an orthonormal basis for $\mathbb{C}^n$.
  (d) The rows of A form an orthonormal basis for $\mathbb{C}^n$.

We noted that a complex matrix P satisfying the conditions in (v) above is called a *unitary* matrix. Such a matrix is the complex analogue of a real orthogonal matrix. We then defined A is be *normal* if $A^*A = AA^*$ and noted (but did not prove) the following

Properties of a normal matrix. Suppose A is an $n \times n$ complex matrix satisfying $A^*A = AA^*$,
(i) If $\lambda \in \mathbb{C}$ is an eigenvalue of A, then $\overline{\lambda}$ is an eigenvalue of A^*.
(ii) $||A^*v|| = ||Av||$, for all $v \in \mathbb{C}^n$.
(iii) If v_1, v_2 are eigenvectors of A corresponding to distinct eigenvalues, then $\langle v_1, v_2\rangle = 0$.

We noted that just as in the Spectral Theorem for real symmetric matrices, item (iii) above for normal matrices plays a *crucial* role in the complex spectral theorem. We then stated, but did not prove any cases of the:

Complex Spectral Theorem. Let A be an $n \times n$ complex matrix. Then A is normal if and only if it is orthogonally diagonalizable, i.e., A is normal if and only if there is a unitary matrix P such that $P^*AP = D$, where D is a diagonal matrix. In particular, a self-adjoint complex matrix is orthogonally diagonalizable.

We then considered the normal matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and found a unitary matrix P such that $P^*AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$. We ended class by noting that the singular value decomposition of a complex matrix A is obtained by the same process used in the case of real matrices, only one starts by finding the eigenvalues of A^*A or AA^*.

Wednesday, April 17. In preparation for stating the Spectral Theorem for complex matrices, we reviewed various properties of complex numbers, complex conjugation, and the inner (dot) product of vectors in $\mathbb{C}^n$. In particular:

Properties of complex numbers. Addition and multiplication of complex numbers are both commutative and associative; complex multiplication distributes over addition; every complex number has an additive inverse; every non-zero complex number has a multiplicative inverse.

Properties of conjugation. For $z = a + bi$, $\overline{z} = a - bi$ denotes its conjugate.
(i) $\overline{z_1z_2} = \overline{z_1}\,\overline{z_2}$ for all $z_1, z_2 \in \mathbb{C}$.
(ii) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$, for all $z_1, z_2 \in \mathbb{C}$.
(iii) If $z = a + bi$, $z\overline{z} = a^2 + b^2 \in \mathbb{R}$ and equals zero if and only if $z = 0$.
(iv) The *modulus* or *absolute value* of $z = a + bi$, is $|z| := \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$.

Properties of the inner product of complex vectors. Suppose v, w are column (or row) vectors in $\mathbb{C}^n$, with coordinates $\alpha_1, \dots, \alpha_n$ and $\beta_1, \dots, \beta_n$. Then inner product of $\langle v, w\rangle$ of v and w is defined as $\langle v, w\rangle := \alpha_1\overline{\beta_1} + \cdots + \alpha_n\overline{\beta_n}$. We discussed the following properties:
(i) $\overline{\langle w, v\rangle} = \langle v, w\rangle$.
(ii) $\langle \lambda v, w\rangle = \lambda\langle v, w\rangle$ and $\langle v, \lambda w\rangle = \overline{\lambda}\langle v, w\rangle$, for all $\lambda \in \mathbb{C}$.
(iii) $\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle$ and $\langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_2\rangle$.
(v) $\langle v, v\rangle$ is a real number greater than or equal to zero and $\langle v, v\rangle = 0$ if and only if $v = 0$.
(vi) The length of v is defined to be $\sqrt{\langle v, v\rangle}$.
(vii) v is defined to be orthogonal to w if and only if $\langle v, w\rangle = 0$.

We ended class by defining the *adjoint*[1] A^* of an $n \times n$ matrix with complex entries to be the conjugate transpose of A, i.e., $A^* = (\overline{A})^t = \overline{A^t}$. We noted in the special case of 2×2 matrices, if the columns of P are an orthonormal basis for $\mathbb{C}^2$, then P^* is the inverse of P.

Monday, April 15. The first 16 minutes of class were devoted to Quiz 10. We then We began class by restating the Singular Value Decomposition Theorem for an $m \times n$ matrix A over $\mathbb{R}$. After this, we sketched a proof of the existence of the existence of the SVD by giving the following steps (with justification) by applying the Spectral Theorem to A^tA.

Steps to the SVD. Let A be an $m \times n$ matrix over $\mathbb{R}$.

(i) Let $\lambda_1 \geq \cdots \geq \lambda_r > 0$ be the non-zero eigenvalues of A^tA. Here $r = \text{rank}(A)$.
(ii) Let P be the $n \times n$ orthogonal matrix that diagonalizes A^tA, so that $P^{-1}(A^tA)P = D$, where D is the diagonal matrix whose diagonal entries are $\lambda_1, \ldots, \lambda_r, 0, \ldots, 0$.
(iii) For $1 \leq i \leq r$, set $\sigma_i = \sqrt{\lambda_i}$.
(iv) For $1 \leq i \leq r$, set $v_i := \frac{1}{\sigma_i}u_i$, where $u_1, \ldots, n_n \in \mathbb{R}^n$ are the columns of P.
(v) Since $v_1, \ldots, v_r$ form an orthonormal system, extend this set of vectors to an orthonormal basis $v_1, \ldots, m_m$ of $\mathbb{R}^m$.
(vi) Letting Q be the orthogonal matrix whose columns are $v_1, \ldots, v_m$, we have $A = Q\sum P^t$, where $\sum$ is the $m \times n$ diagonal matrix whose non-zero entries are $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_r > 0$.

We ended class by noting that a few key facts to the proof sketch we offered are among the properties of A^tA and AA^t listed below. We did not prove these facts.

(i) A^tA and AA^t are symmetric matrices, and hence their eigenvalues are in $\mathbb{R}$.
(ii) The eigenvalues of A^tA and AA^t are all greater than or equal to zero.
(iii) The matrices A^tA and A have the same null space. Similarly, AA^t and A^t have the same null space.
(iv) The matrices A, A^t, A^tA, and AA^t all have the same rank.
(v) A^tA and AA^t have the same non-zero eigenvalues with the same algebraic multiplicities.

It follows that we could also find the SVD of A by starting with the matrix AA^T and follow the same steps as above, except in step (iv), we take $v_i := \frac{1}{\sigma_i}A^tu_i$, for $1 \leq i \leq r$.

Friday, April 12. We spent the class discussing and working an example illustrating the:

Singular Value Decomposition. Let A be an $m \times n$ matrix with entries in $\mathbb{R}$. Then over $\mathbb{R}$ there exists an orthogonal $m \times m$ matrix Q, an orthogonal $n \times n$ matrix P, and an $m \times n$ diagonal matrix $\sum$ such that $A = Q\sum P^t$ and the non-zero diagonal entries of $\sum$ are real numbers $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_r > 0$, where r is the rank of A. The real numbers $\sigma_1, \ldots, \sigma_r$ are called the *singular values of* A.

After stating the SVD, we had a lengthy discussion concerning the following comments:

(i) The SVD gives a *pseudo-diagonalization*: $Q^{-1}AP = \sum$.
(ii) The SVD is obtained by applying the Spectral Theorem to either A^tA or AA^t.
(iii) An orthogonal matrix corresponds to either a reflection or rotation.
(iv) Thus, multiplication by **any** $m \times n$ matrix A gives linear transformation from $\mathbb{R}^n \to \mathbb{R}^m$ which is a rotation or reflection of $\mathbb{R}^n$, followed by an elongation, followed a reflection or rotation of $\mathbb{R}^m$.
(v) The SVD gives a *pseudo-inverse* of A, namely $A^+ = P\sum^+ Q^t$, where $\sum$ is the $n \times m$ diagonal matrix with diagonal entries $\sigma_1^{-1}, \ldots, \sigma_r^{-1}$. From this, one gets a solution to the following **least squares problem**: Given $w \in \mathbb{R}^m$, find a vector $v \in \mathbb{R}^n$ so that the length $||Av - w||$ is minimal, as $v \in \mathbb{R}^n$ varies over all possible column vectors. The answer is $v_0 = A^+w$.
(vi) Other applications include: data compression, image restoration, noise removal.

We then spent the rest of the class finding the SVD for $A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, from the homework of Wednesday, April 10. The important point here was to follow the general outline presented in that assignment.

Wednesday April 10. We began class by restating the Spectral Theorem for $n \times n$ matrices over $\mathbb{R}$ and noting three key facts established along the way: (i) $Av \cdot w = v\dot{A}w$, for column vectors $v, w \in \mathbb{R}^n$; (ii) $p_A(x)$ has

[1]IMPORTANT NOTE: The adjoint of A is not the same as the *classical adjoint* of A.

all of its roots in $\mathbb{R}$; (iii) If $v_1, \dots, v_r \in \mathbb{R}$ are eigenvectors corresponding to distinct eigenvalues of A, then $v_1, \dots, v_r$ are mutually orthogonal.

We then worked through finding an orthogonal matrix diagonalizing $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, taking care to note that E_0 is the orthogonal complement of E_3. This was to illustrate the general reduction process: For an $n \times n$ real symmetric matrix A, first find a single eigenvalue λ_1 and unit length eigenvector u_1 for λ_1. Upon setting $W := \text{Span}\{u_1\}$, find an orthonormal basis $u_2, \dots, u_n$ for $W^\perp$. Upon setting $P := [u_1\ u_2\ \cdots\ u_n]$, one has $P^{-1}AP = \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & B \end{pmatrix}$, where $\mathbf{0}$ denotes the zero row or column vector of length $n-1$ and B is a symmetric $(n-1) \times (n-1)$ matrix. On then repeats the process on B and continues until A has been orthogonally diagonalized.

We then started with the matrix $A := \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$, and showed that there exists an orthogonal 3×3 matrix Q and a 2×2 orthogonal matrix P such that $Q^{-1}AP = \sum$, where $\sum = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. Here $3, 1$ are the eigenvalues of the symmetric matrix A^tA. It followed that $A = Q\Sigma P^t$, which is a special case of the *Singular Value Decomposition* of the matrix A.

Monday, April 8. We began class with a few comments on Exam 2. We also stated that students who did not do well on the exam (or the first midterm exam) will have the option of replacing their **lowest** midterm exam score, with their final exam score, assuming it helps their grade.

We then discussed how one reduces the general $n \times n$ case of the Spectral Theorem to the $(n-1) \times (n-1)$ case. The point is to follow what we did in the lecture of April 5 in reducing the 3×3 case to the 2×2 case. This reduction relied upon the following observations for $n \times n$ real symmetric matrices A:

(ii) If $\lambda_, \lambda_2$ are distinct eigenvalues of A with $Av_1 = \lambda v_2$ and $Av_2 = \lambda_2 v_2$, then $v_1 \cdot v_2 = 0$,
(ii) A has all of its eigenvalues in $\mathbb{R}$.

Item (i) above followed from the fact, that if $Av_1 \cdot v_2 = v_1 \cdot Av_2$ (dot product) for v_1, v_2 column vectors in $\mathbb{R}^n$. Item (ii) followed first by using the Fundamental Theorem of Algebra, which implies that every eigenvalue of A belongs to $\mathbb{C}$, and then by noting that if $(A - \lambda I)(v) = 0$, with $\lambda = a + bi \in \mathbb{C}$, then applying $A - \overline{\lambda} I$ to $(A - \lambda I)(v) = \vec{0}$ leads to $(A^2 - aI)^2 v + b^2 b = \vec{0}$ and dotting with v, ultimately shows that $b = 0$, so $\lambda \in \mathbb{R}$. Here we used $\overline{\lambda}$ is the complex conjugate of λ.

Friday, April 5. Exam 2.

Wednesday, April 3. The class worked on practice problems for Exam 2.

Monday, April 1. The first fifteen minutes of class were devoted to Quiz 9. We then considered the vector space V over $\mathbb{R}$ with inner product $\langle\ ,\ \rangle$ and a subspace $W \subseteq V$. We defined $W^\perp$ ("W perp"), the *orthogonal complement of* W, as $W^\perp := \{v \in V \mid \langle v, w \rangle = 0, \text{ for all } w \in W\}$. We noted that, for example, the z-axis in $\mathbb{R}^3$ is the orthogonal complement of the xy-plane in $\mathbb{R}^3$. Similarly, if $W \subseteq \mathbb{R}^3$ is the plane defined by $ax + by + cz = 0$, we noted that $W^\perp$ is the line through the origin containing the vector (a, b, c) and that perpendicular lines through the origin in $\mathbb{R}^2$ give subspaces W and $W^\perp$. This was followed by discussing and giving a proof of the following proposition:

Proposition. In the notation above, we have:

(i) $W^\perp$ is a subspace of V.
(ii) If $\{u_1, \dots, u_r\}$ is an orthonormal basis for W and $\{u_{r+1}, \dots, u_n\}$ is such that $\{u_1, \dots, u_n\} \subseteq V$ is an orthonormal basis for V, then $\{u_{r+1}, \dots, u_n\}$ is an orthonormal basis for $W^\perp$.

The proof of (ii) required extending $u_1, \dots, u_r$ to a basis $u_1, \dots, u_r, v_{r+1}, \dots, v_v$ for V and then applying Gram-Schmidt, to get the orthonormal basis $u_1, \dots, u_n$ for V. After concluding the proof of part (ii) we then noted that the proposition shows that $V = W \oplus W^\perp$.

We ended class with a brief discussion of the QR factorization of an $m \times n$ matrix A over $\mathbb{R}$. This result states that if the columns of A are linearly independent (and hence $n \leq m$), then there are real matrices Q, R such that: (i) Q is an $m \times n$ matrix whose columns form an orthonormal basis for the column space of A and (ii) R is an $n \times n$ upper triangular matrix whose diagonal entries are all positive. In particular, if A is an $n \times n$ invertible matrix, then Q is an orthogonal matrix, i.e., $QQ^t = I_n$. We mentioned that the QR factorization makes solving least squares problems of the type $A\mathbf{x} = \mathbf{b}$ easier, when the system of equations $A\mathbf{x} = \mathbf{b}$ does not have a solution.

Friday, March 29. We discussed at length the Gram-Schmidt process and worked a couple of examples illustrating it. We began the discussion by seeing how to start with two linearly independent vectors v_1, v_2 and construct orthogonal vectors w_1, w_2 so that $\text{Span}\{v_1, v_2\} = \text{Span}\{w_1, w_2\}$. We then looked at the case of three vectors and were able to derive a formula for the process in general.

Gram-Schmidt Process. Let V be a vector space with inner product $\langle \ , \ \rangle$ and suppose $\{v_1, \ldots, v_r\}$ is a linearly independent set of vectors. Then there exists an orthogonal set of vectors $\{w_1, \ldots, w_r\}$ such that $\text{Span}\{w_1, \ldots, w_r\} = \text{Span}\{v_1, \ldots, v_r\}$. More over, the vectors $w_1, \ldots, w_r$ can be constructed inductively as follows:

(i) $w_1 := v_1$.
(ii) If $w_1, \ldots, w_i$ have been constructed so that $\text{Span}\{w_1, \ldots, w_i\} = \text{Span}\{v_1, \ldots, v_i\}$ and $w_1, \ldots, w_i$ are mutually orthogonal then taking

$$w_{i+1} = v_{i+1} - \frac{\langle v_{i+1}, w_1 \rangle}{\langle w_1, w_1 \rangle} \cdot w_1 - \cdots - \frac{\langle v_{i+1}, w_i \rangle}{\langle w_i, w_i \rangle} \cdot w_i,$$

we have that $\text{Span}\{w_1, \ldots, w_{i+1}\} = \text{Span}\{v_1, \ldots, v_{i+1}\}$ and $\{w_1, \ldots, w_{i+1}\}$ is an orthogonal set of vectors.

It followed immediately from the Gram-Schmidt process that if V is any finite dimensional vector space (for now, over $\mathbb{R}$) with an inner product, then V admits an orthonormal basis. We ended class by applying the Gram-Schmidt process to the following basis for $\mathbb{R}^3$: $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

Wednesday, March 27. We continued our discussion of inner products by looking at the following examples:

(i) $V = \mathbb{R}^n$ with the usual dot product as the inner product.
(ii) $V =$ the space of real polynomials of degree less than or equal to n with $\langle f(x), g(x) \rangle := \int_a^b f(x)g(x)\, dx$.
(iii) $V =$ the space on $n \times n$ matrices over $\mathbb{R}$ with $\langle A, B \rangle := \text{trace}(A^t \cdot B)$.
(iv) $V = \mathbb{R}^n$ (column vectors) and A, a positive definite symmetric $n \times n$ real matrix, with $\langle v, w \rangle := v^t A w$.

For the examples above, we worked through (iii) for the case of 2×2 matrices over $\mathbb{R}$. We then gave the following definition, emphasizing that equipping the vector space V with an inner product enables us to establish the following concepts in very general situations.

Definitions. Let V be a vector space over $\mathbb{R}$ with inner product $\langle \ , \ \rangle$.

(ii) $v, w \in V$ are *orthogonal* if $\langle v, w \rangle = 0$.
(ii) The *length* of $v \in V$ is given by $||v|| := \sqrt{\langle v, v \rangle}$.
(iii) The set of vectors $\{v_1, \ldots, v_r\} \subseteq V$ is an *orthogonal set* if $\langle v_i, v_j \rangle = 0$, for all $v_i \neq v_j$.
(iv) The set of vectors $\{u_1, \ldots, u_r\}$ is *orthonormal* if it is orthonormal and $||u_j|| = 1$, for all $1 \leq j \leq r$.

With these definitions in hand, we were able to establish the following:

Two important facts. Let V be a real vector space with an inner product.

(i) If $\{v_1, \ldots, v_r\}$ is an orthogonal subset of V, then $v_1, \ldots, v_r$ are linearly independent.
(ii) If $\{u_1, \ldots, u_n\}$ is an orthonormal basis for V and $v \in V$, then $v = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_n \rangle u_n$.

We ended class by observing that $u_1 := \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\\0\end{pmatrix}, u_2 := \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, u_3 := \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\-2\end{pmatrix}$ is an orthonormal basis for $\mathbb{R}^3$ and using Fact (ii) above to quickly find $x, y, z \in \mathbb{R}$ such that $\begin{pmatrix}2\\7\\13\end{pmatrix} = xu_1 + yu_2 + zu_3$.

Monday, March 25. The first thirteen minutes of class were devoted to Quiz 8. We then reviewed the basis properties of the dot product in $\mathbb{R}^3$, noting that for vectors $v, w \in \mathbb{R}^3$ and $\lambda \in \mathbb{R}$,

(i) $v \cdot w = w \cdot v$.
(ii) $(v_1 + v_2) \cdot w = v_1 \cdot w + v_2 \cdot w$ and $v \cdot (w_1 + w_2) = v \cdot w_1 + v \cdot w_2$.
(iii) $(\lambda v) \cdot w = \lambda(v \cdot w) = v \cdot (\lambda w)$.
(iv) $v \cdot v \geq 0$ and $v \cdot v = 0$ if and only if $v = \vec{0}$.

Two further properties were noted:

(v) $||v|| = \sqrt{v \cdot v}$, where $||v||$ denotes the length of v.
(vi) The angle between v, w is given by $\cos^{-1}(\frac{v \cdot w}{||v||||w||})$, so that $v \cdot w = 0$ if and only v and w are orthogonal.

We then defined the dot product of vectors $v = (a_1, \ldots, a_n)$ and $w = (b_1, \ldots, b_n)$ in $\mathbb{R}^n$ as

$$v \cdot w := a_1b_1 + \cdots + a_nb_n,$$

and similarly if we take column vectors in $\mathbb{R}^n$. We then noted (but did not prove) that properties (i)-(vi) hold for the dot product in $\mathbb{R}^n$. This lead to the following definitions.

Definitions. Let V be a vector space over $\mathbb{R}$.

(A) A function from $V \times V$ to $\mathbb{R}$ taking the pair of vectors (v, w) to the real number $\langle v, w\rangle$ is an *inner product* on V if it satisfies properties (i)-(iv) above, namely, for all $v, w \in V$ and $\lambda \in \mathbb{R}$,
  (i) $\langle v, w\rangle = \langle w, v\rangle$.
  (ii) $\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle$.
  (iii) $\langle \lambda v, w\rangle = \lambda\langle v, w\rangle = \langle v, \lambda w\rangle$.
  (iv) $\langle v, v\rangle \geq 0$ and $\langle v, v, \rangle = 0$ if and only if $v = \vec{0}$.
(B) The *length* $||v||$ of the vector v is given by $\sqrt{\langle v, v\rangle}$.
(C) The *angle* between non-zero vectors $v, w \in V$ is given by $\cos^{-1}(\frac{\langle v,w\rangle}{||v||||w||})$. In particular v and w are orthogonal if $\langle v, w\rangle = 0$.

We ended class by noting that the dot product on $\mathbb{R}^n$ is an inner product and that for V, the vector space of real polynomial of degree less than n, $\langle f(x), g(x)\rangle := \int_0^1 f(x) \cdot g(x)\, dx$ defines an inner product on V. We also noted, but did not fully verify, that if V is the vector space of real 2×2 matrices, then $\langle A, B\rangle := A^tB$, for all $A, B \in V$ is an inner product on V.

Friday, March 22. We continued our discussion of the main theorem from the previous lecture. In particular, we focused on the implication (ii) implies (i) in the case that A is a 7×7 matrix whose characteristic polynomial $p_A(x) = (x - \lambda_1)^2(x - \lambda_2)^2(x - \lambda_3)^2$. The crucial point in our analysis was the following: From the conditions $\dim(E_{\lambda_1}) = 2, \dim(E_{\lambda_2}) = 3, \dim(E_{\lambda_3}) = 3$ we showed that putting together bases for each eigenspace E_{λ_j} gives a basis for $\mathbb{R}^7$, leading quickly to the diagonalizability of A.

We then applied the theorem to show that the matrix $A = \begin{pmatrix}1&0&2\\0&2&-1\\0&0&3\end{pmatrix}$ is diagonalizable and found the diagonalizing matrix P and the matrix $B = \begin{pmatrix}1&1&0\\0&1&0\\0&0&2\end{pmatrix}$ is not diagonalizable.

We ended class by noting the following computation criteria for an $n \times n$ matrix A to be diagonalizable, assuming $p_A(x)(x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$: A is diagonalizable if and only if for each $1 \leq i \leq r$, the reduced row echelon form of $A - \lambda_i I_n$ has rank $n - e_i$ if and only if for each $\leq i \leq r$, the number of zero rows in reduced row echelon form of $A - \lambda_i I_n$ equals e_i.

Wednesday, March 20. The first fifteen minutes of class were devoted to Quiz 7.After reviewing some of the material presented in the previous class, we began a discussion of the following important theorem.

Theorem. Let $T : V \to V$ be a linear transformation with V an n dimensional vector space over F and A an $n \times n$ matrix with entries in F. Write $p_T(x)$, for the characteristic polynomial of T and $p_A(x)$ for the characteristic polynomial of A. Suppose that $\lambda_1, \ldots, \lambda_r \in F$ are distinct.

(A) The following are equivalent for T:
- (i) T is diagonalizable.
- (ii) $p_T(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$ and $\dim(E_{\lambda_i}) = e_i$, for $1 \leq i \leq r$.
- (iii) $p_T(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$ and $\dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_r}) = n$.

(B) The following are equivalent for A:
- (i) A is diagonalizable.
- (ii) $p_A(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$ and $\dim(E_{\lambda_i}) = e_i$, for $1 \leq i \leq r$.
- (iii) $p_A(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$ and $\dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_r}) = n$.

We noted that the eigenspaces E_{λ_i} in part (A) are subspaces of V, while the eigenspaces in (B) consist of column vectors in F^n. Rather than giving a formal proof of the theorem we did an in depth analysis of the implication (i) implies (ii) in part (B) by looking at a diagonalizable 7×7 matrix A satisfying $P^{-1}AP = D(\lambda_1, \lambda_1, \lambda_2, \lambda_2, \lambda_2, \lambda_3, \lambda_3)$. We ended class by beginning an analysis of (ii) implies (i) for a 7×7 matrix A satisfying $p_A(x) = (x - \lambda_1)^2 (x - \lambda_2)^3 (x - \lambda_3)^2$.

Monday, March 18. We continued our discussion of eigenvectors and eigenvalues, both for a linear transformation $T : V \to V$ ($\dim(V) = n$), and A, an $n \times n$ matrix. After reviewing the definitions from the previous lectures, we discussed the following important facts.

Fact 1. Suppose $v_1, \ldots, v_r \in V$ are eigenvectors of T with distinct eigenvalues $\lambda_1, \ldots, \lambda_r$, then $v_1, \ldots, v_r$ are linearly independent.

Before proving Fact 1, we illustrated this fact with the matrix $\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$, which we ultimately showed was diagonalizable. The general proof of Fact 1was by contradiction: Assuming the fact was false, by looking at a non-trivial dependence relation of shortest length among the v_j, we were able to derive a shorter dependence relation, thereby deriving a contradiction. We then noted that Fact 1 has the following **important consequence.**

Corollary. Suppose $T : V \to V$ has n distinct eigenvalues. Then T is diagonalizable.

Before moving on to Fact 2, we noted that Fact 1 and its corollary have easy to obtain analogues for matrices.

Fact 2. Suppose the characteristic polynomial of T or A has the form $(x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$. Then for each $1 \leq i \leq r$, E_{λ_i}, the eigenspace of λ_i, has dimension less than or equal to e_i. We noted that e_i is often called the *algebraic multiplicity* of λ_i and the dimension of E_{λ_i} is often called the *geometric multiplicity* of λ_i.

We sketched a proof of the general argument. The idea being that if, say E_{λ_1} contains $e_1 + 1$ independent vectors, then these vectors can be extended to a basis α of V. The matrix $[T]_\alpha^\alpha$ is then in block form, with an upper block being a $(e_1 + 1) \times (e_1 + 1)$ identity matrix with zeros below this block. Using this matrix to calculate $p_T(X)$ gives that $p_T(x)$ is divisible by $(x - \lambda_1)^{e_1+1}$, which cannot happen. Thus, the dimension of E_{λ_1} can be no more than e_1.

We ended class by stating, but not yet proving, that the matrix A, or the transformation T, is diagonalizable if and only if $p_A(X)$ or $P_T(x)$ factors completely as a product of linear polynomials and for each eigenvalue λ_i, the geometric multiplicity of λ_i equals the algebraic multiplicity of λ_i.

Friday, March 8. We began class by reviewing the definitions of *eigenvalue* and *eigenvector*, both for A, an $n \times n$ matrix over F, and a linear transformation $T : V \to V$: Given $\lambda \in F$, λ is an eigenvalue of A or T if there exists $0 \neq v \in F^n$ or $0 \neq v \in V$ such that $AV = \lambda v$ or $T(v) = \lambda v$. Given A, we set $p_A(x) := \left|A - xI_n\right|$, the *characteristic polynomial of* A. We then noted that λ is an eigenvalue of A if and only if $p_A(\lambda) = 0$ if and only if the nullspace of $A - \lambda I_n$ is non-zero, and $0 \neq v$ is an eigenvector associated to λ if and only if v belongs to the nullspace of $A - \lambda I_n$. The nullspace of $A - \lambda I_n$ is called the *eigenspace of* λ and is denoted E_λ. The foregoing also applies to T, since if A and B are two matrices representing T, $p_A(x) = p_B(x)$, so that λ is an eigenvalue of T if and only if $p_A(\lambda) = 0$, for any matrix A representing T. In this case, the eigenspace associated to λ is $E_\lambda := \ker(T - \lambda \cdot I_V)$.

We then defined A to be *diagonalizable (over F)* if there exists an invertible $n \times n$ matrix P with entries in F such that $P^{-1}AP = D$, a diagonal matrix. The linear transformation T is *diagonalizable (over F)* if there exists a basis α of V such that $[T]_\alpha^\alpha = D$. In each case, we noted that the diagonal entries of D are necessarily the eigenvalues of A or T. We concluded the lecture by analyzing what happens with a 2×2 matrix A over $\mathbb{R}$:

(i) If $p_A(x)$ has no roots in $\mathbb{R}$, then A is not diagonalizable over $\mathbb{R}$.
(ii) If $p_A(x)$ has two distinct roots in $\mathbb{R}$, then A is diagonalizable.
(iii) If $p_A(x)$ has a repeated root λ, i.e., $p_A(x) = (x - \lambda)^2$, then either $A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ or A is not diagonalizable, because in the latter case case, E_λ is a one-dimensional subspace of $\mathbb{R}^2$.

As an example, we noted that the matrix $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ is not diagonalizable.

Wednesday, March 6. We began class by recalling that in the previous lecture we showed that if A is a 2×2 real symmetric matrix, then $Av_1 \cdot v_2 = v_2 \cdot Av_2$, for all column vectors $v_1, v_2 \in \mathbb{R}^2$. This lead to the following definition:

Definition. Let $R : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation. Then T is *symmetric* if for all column vectors $v_1, v_2 \in \mathbb{R}^2$, $T(v_1) \cdot v_2 = v_1 \cdot T(v_2)$.

We then presented the following Proposition, proving the only if direction.

Proposition. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation. Then T is symmetric if and only if for every *orthonormal basis* $\alpha \subseteq \mathbb{R}^2$, $[T]_\alpha^\alpha$ is a symmetric matrix.

We noted that the matrix of a symmetric linear transformation with respect to a basis that is not an orthonormal basis need not be symmetric, as the following example shows.

Example. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(x, y) := (x + 2y, 2x + y)$. Then, if E denotes the standard basis of $\mathbb{R}^2$, $[T]_E^E = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, so the matrix of T with respect to the orthonormal basis E is symmetric. On the other hand, consider the basis $B = \{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$, which is not an orthonormal basis. Then it is easy to check that $[T]_B^B = \begin{pmatrix} 3 & 2 \\ 0 & -1 \end{pmatrix}$, which is not symmetric.

This was followed by observing that the converse of the spectral theorem holds, namely, if A is a 2×2 matrix over $\mathbb{R}$ and there exists an orthogonal matrix P such that $P^{-1}AP$ is a diagonal matrix, then A is symmetric. We ended class by defining the general concepts of eigenvalues and eigenvectors, both for for $n \times n$ matrices with entries in $F = \mathbb{R}$ or $\mathbb{C}$ and linear transformations $T : V \to V$, where V is an n-dimensional vector space over F.

Monday, March 4. The first fifteen minutes of class were devoted to Quiz 6. We then continued our discussion of the *Spectral Theorem*, turning our attention to an arbitrary 2×2 real symmetric matrix $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Through a series of straightforward calculations, we were able to show:

(i) If $p(x) = \begin{vmatrix} x - a & -b \\ -b & x - c \end{vmatrix}$ is the characteristic polynomial of A, then its roots are in $\mathbb{R}$. That is, the eigenvalues of A are real numbers.

(ii) If $p(x)$ has a repeated root, then $A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$, and the standard basis for $\mathbb{R}^2$ is already an orthonormal basis consisting of eigenvalues for A.
(iii) If A has distinct eigenvalues, say $\lambda_1 \neq \lambda_2$ with eigenvectors v_1, v_2, respectively, then $v_1 \cdot v_2 = 0$, i.e., v_1 and v_2 are orthogonal.
(iv) For $u_1 := \frac{1}{||v_1||} \cdot v_1$ and $u_2 := \frac{1}{||v_2||} \cdot v_2$, $\{u_1, u_2\}$ is an orthonormal basis for $\mathbb{R}^2$ consisting of eigenvalues for A.
(v) If P is the 2×2 matrix whose columns are u_1, u_2, then P is an orthogonal matrix (so $P^{-1} = P^t$) and $P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

Friday, March 1. We began class by stating, but not proving, the following **important fact.** If A and B are $n \times n$ matrices over F, then $|AB| = |A| \cdot |B|$. We indicated roughly why this formula holds, as follows: We began by defining an elementary matrix E to be one obtained from I_n be applying one of the three standard elementary row operations. We then observed (for 2×2 matrices) that EA is the matrix obtained by applying the corresponding elementary row operation directly to A. From our previous discussions, we have $|EA| = |E| \cdot |A|$. Since an invertible matrix row reduces to the identity matrix, if A is invertible, we can write $A = E_1 \cdots E_r$ as a product of elementary matrices. Thus,

$$|AB| = |E_1 \cdots E_r B| = |E_1| \cdots |E_r| \cdot |B| = |E_1 \cdots E_r| \cdot |B| = |A| \cdot |B|.$$

We also noted that if A or B has non-zero nullspace, the same applies to AB so both sides of the equation $|AB| = |A| \cdot |B|$ are zero.

We then began our discussion of the *Spectral Theorem* over $\mathbb{R}$ which states that if A is an $n \times n$ symmetric matrix, then there is an *orthonormal basis* of $\mathbb{R}^n$ consisting of eigenvectors of A, which is equivalent to saying that there exists an *orthogonal matrix* P such that $P^{-1}AP = D$, where D is the diagonal matrix having the eigenvalues of A down its main diagonal. We noted that an orthogonal matrix is one whose columns form an orthonormal basis for $\mathbb{R}^n$. We ended class by finding an orthonormal basis consisting of eigenvectors for the matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. A key ingredient was that the eigenvectors associated to the two eigenvalues of A were orthogonal. We also found the orthogonal matrix P satisfying $P^{-1}AP = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$.

Wednesday, February 28. We continued our discussion of determinants, beginning with recalling the effect elementary row or column operations have on calculating the determinant of an $n \times n$ matrix. We then used elementary row operations to calculate the determinant of a 3×3 matrix.

We then discussed the adjoint formula, $A \cdot A' = |A| \cdot I_n = A' \cdot A$, where A is an $n \times n$ matrix over F and $A' = C^t$, for C the $n \times n$ matrix whose $(i, j)^{\text{th}}$-entry is $(-1)^{i+j}|A_{ij}|$, and illustrated this formula by calculating a few entries in AA', when A is an arbitrary 3×3 matrix. We noted that it follows immediately from the classical adjoint formula that A is invertible with $A^{-1} = \frac{1}{|A|} \cdot A'$, if $|A| \neq 0$. We then derived:

Cramer's Rule. Let A be an $n \times n$ matrix with coefficients in F, and $A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ be a system of n equations in n unknowns. For each $1 \leq i \leq n$ let B_i be the matrix obtained fro A by replacing its ith column by $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$. Then, for each $1 \leq i \leq n$, $x_i = \frac{|B_i|}{|A|}$.

We ended class by discussing (but not formally proving) the connections between the conditions in the following theorem:

Theorem. Let A be n $n \times n$ matrix with entries in F. The following re equivalent:

(i) $|A| \neq 0$.
(ii) A is invertible.

(iii) The null space of A is zero, i.e., if $v \in F^n$ and $Av = \vec{0}$,then $v = \vec{0}$.
(iv) A reduces to I_n via elementary row operations.
(vi) The rows (respectively, columns) of A are linearly independent.
(vii) The rows (respectively, columns) of A span F^n.
(viii) The rows (respectively, columns) of A form a basis for F^n.
(ix) The linear transformation $T_A : F^n \to F^n$ is one-to-one and onto.
(x) T_A is 1-1.
(xi) T_A is onto.
(xii) Any $n \times n$ system of linear equation with coefficient matrix A has a unique solution.

Monday, February 26. We began class by briefly discussing Exam 1 and how students might prepare better for the next exam. We then began our discussion of determinants. After calculating a few examples of determinants of matrices of different sizes, we gave a formal definition:

Definition. Let $A = (a_{ij})$ be and $n \times n$ matrix with entries in F. Then the *determinant* of A, denoted $|A|$ or $\det(A)$, is defined by the following equations:

$$\begin{aligned} |A| &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \cdot |A_{ij}| \qquad \text{(expansion along the } i\text{th row)} \\ &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \cdot |A_{ij}| \qquad \text{(expansion along the } j\text{th column)}, \end{aligned}$$

where A_{ij} denotes the $(n-1) \times (n-1)$ matrix obtained from A by deleting its ith row and jth column. We emphasized that the fact that the different expansions of the determinant always give the same answer is not an easy fact to prove, and we will just assume that all expansions in the definition give the same result.

We then discussed the following properties of the determinant, thinking of the determinant as a function of its rows. Letting A denote an $n \times n$ matrix over F:

(i) If A' is obtained form A by multiplying a row (or columns) of A times $\lambda \in F$, then $|A'| = \lambda \cdot |A|$.
(ii) If A' is obtained from A by interchanging two rows (or tow columns), then $|A'| = -|A|$.
(iii) If a row (or column) of A consists entirely of 0s, then $|A| = 0$.
(iv) If tow rows (or columns) of A are the same, then $|A| = 0$.
(v) If A' is obtained from A by adding a multiple of one row of A to **another** row, then $|A'| = |A|$.
(vi) If A is an upper or lower triangular matrix, then $A|$ is the product of the diagonal entries of A.
(vii) The determinant is a linear function of its rows (or columns)

We ended class by using elementary row operations to calculate the determinant of a 4×4 matrix.

Friday, February 23. Exam 1.

Wednesday, February 21. We began class with a discussion and proof of the following theorem.

Theorem. Let V be a finite dimensional vector space and $T : V \to V$ a linear transformation. The following statements are equivalent:

(i) T is one-to-one and onto.
(ii) T is one-to-one.
(iii) T is onto.

We noted that this theorem is similar in spirit to our previous theorem stating that for n vectors in an n-dimensional vector space the following are equivalent: (i) The vectors are linearly independent and span; (ii) the vectors are linearly independent; (iii) the vectors span. The rest of the class was devoted to group work on practice problems for Exam 1.

Monday, February 19. The first eighteen minutes of class were devoted to Quiz 5. We then began by recalling the definitions of the kernel and image of a linear transformation. The majority of the rest of the class was devoted to a careful proof of the Rank plus Nullity Theorem, as stated in the previous lecture. We ended class by showing that for the linear transformation $T : V \to W$, T is 1-1 if and only if $\ker(T) = \{\vec{0}\}$ and noting that T is onto if and only if $\text{im}(T) = W$. Consequently, if $T : V \to V$ and $\dim(\text{im}(T)) = \dim(V)$, then T is 1-1.

Friday, February 16. We began class by reviewing the change of basis theorem presented in the previous lecture. We then discussed, and ultimately proved, a more general change of basis result:

General Change of Basis Theorem. Let $T : V \to W$ be a linear transformation. Assume α_1, α_2 are bases for V and β_1, β_2 are bases for W respectively. Then:

$$[T]_{\alpha_2}^{\beta_2} = [I_W]_{\beta_1}^{\beta_2} \cdot [T]_{\alpha_1}^{\beta_1} \cdot [I_V]_{\alpha_2}^{\alpha_1},$$

where I_V denotes the identity transformation on V and I_W denotes the identity transformation on W. Writing $A := [T]_{\alpha_1}^{\beta_1}, B := [T]_{\alpha_2}^{\beta_2}, P := [I_V]_{\alpha_2}^{\alpha_1}$ and $Q := [I_W]_{\beta_2}^{\beta_1}$, the change of basis formula given above takes a more familiar form: $B = Q^{-1}AP$. Again, the proof of the theorem was an application of the second important formula from the lecture of February 12. We also noted that if V has dimension n, and W has dimension m, then: A and B are $m \times n$ matrices, P is an $n \times n$ matrix and Q is an $m \times m$ matrix.

We next considered $T : V \to W$ and gave the following definitions:

Definitions. (i) The *kernel of* T (sometimes called the *null space of* T), denoted $\ker(T)$, is the set of vectors $v \in V$ satisfying $T(v) = \vec{0}$.

(ii) The *image of* T (sometimes called the *range of* T), denoted $\text{im}(T)$, is the set of all vectors $w \in W$ such that $w = T(v)$, for some $v \in V$.

We then showed that $\ker(T)$ is a subspace of V and indicated that $\text{im}(T)$ is a subspace of W. This was followed by stating the following **very important** theorem. Our book refers to this as the *Dimension Theorem.* More commonly, this theorem is referred to as the *Rank plus Nullity Theorem*, since the dimension of $\ker(T)$ is often called the *nullity of* T and the dimension of $\text{im}(T)$ is called the *rank of* T.

Rank plus Nullity Theorem. Let $T : V \to W$ be a linear transformation between the finite dimensional vector spaces V and W. Then:

$$\dim(V) = \dim(\text{im}(T)) + \dim(\ker(T)).$$

We ended class by verifying the rank plus nullity theorem for the linear transformation $T_A : \mathbb{R}^3 \to \mathbb{R}^3$ for A given by $\begin{pmatrix} 1 & -2 & 1 \\ 3 & -1 & -1 \\ 4 & -3 & 0 \end{pmatrix}$.

Wednesday, February 14. We we began by reviewing the Formulas 2 and 3 from the previous lecture. We also gave a proof of Formula 3 and illustrated Formula 3 with an example. where $T : \mathbb{R}^2 \to \mathbb{R}^3$ was given by $T(x, y) = (x + y, y - x, 2x)$ and $S : \mathbb{R}^3 \to \mathbb{R}^2$ was given by $S(x, y, z) = (x + 2y + 3z, -y)$¡ using the standard bases for $\mathbb{R}^3$ and $\mathbb{R}^2$.

We followed this by discussing and proving the following

Change of Basis Theorem. Let $T : V \to V$ be a linear transformation, and suppose α_1 and α_2 are bases for V. Then

$$[T]_{\alpha_2}^{\alpha_2} = [I]_{\alpha_1}^{\alpha_2} \cdot [T]_{\alpha_1}^{\alpha_1} \cdot [I]_{\alpha_2}^{\alpha_1},$$

where $I : V \to V$ denotes the identity transformation.

We noted that $[I]_{\alpha_2}^{\alpha_1}$ is the matrix obtained by expressing the vectors in α_2 in terms of the vectors in α_1 and that the matrices $[I]_{\alpha_2}^{\alpha_1}$ and $[I]_{\alpha_1}^{\alpha_2}$ are inverses of one another. We also noted that if we write $A := [T]_{\alpha_1}^{\alpha_1}$, $B := [T]_{\alpha_2}^{\alpha_2}$ and $P := [I]_{\alpha_2}^{\alpha_1}$, then the change of basis theorem takes the familiar form $B = P^{-1}AP$.

We ended class by verifying the change of basis theorem for $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x, y) := (x+2y, 2x-y)$, $\alpha_1 = E := \{e_1, e_2\}$, the standard basis, and $\alpha_2 = F := \{f_1, f_2\}$, for $f_1 := (-1, 1)$ and $f_2 := (1, 1)$.

Monday, February 12. The first thirteen minutes of class were devoted to Quiz 4. We then continued by establishing our basic set-up. Suppose that $T : V \to W$ is a linear transformation of finite dimensional vector spaces over $F = \mathbb{R}$ or $\mathbb{C}$. Let $\alpha := \{v_1, \ldots, v_n\}$ be a basis for V and $\beta := \{w_1, \ldots, w_m\}$ be a basis for

W. We write $[T]_\alpha^\beta$ for the matrix of T with respect to the bases α and β, as defined in the previous lecture. We also write $[v]_\alpha := \begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}$, whenever $v = \gamma v_1 + \cdots + \gamma_n v_n$.

We devoted the rest of the class to discussing the following formulas:

Three Important Formulas. 1. For $v, u \in V$ amd $c, d \in F$, $[cv + du]_\alpha = c[v]_\alpha + d[u]_\alpha$.

2. For any $v \in V$, $[T(v)]_\beta = [T]_\alpha^\beta \cdot [v]_\alpha$.

3. If $S : W \to U$ is a linear transformation, and $\gamma := \{u_1, \ldots, u_t\}$ is a basis for U, then $[ST]_\alpha^\gamma = [S]_\beta^\gamma \cdot [T]_\alpha^\beta$.

We sketched proofs of the first two formulas and did a concrete example verifying Formula 2.

Friday, February 9. We began class by defining the matrix of a linear transformation with respect bases, as follows:

Definition. Suppose that $T : V \to W$ is a linear transformation of finite dimensional vector spaces. Let $\alpha := \{v_1, \ldots, v_n\}$ be a basis for V and $\beta := \{w_1, \ldots, w_m\}$ be a basis for W. Then the *matrix of T with respect to the bases α and β* is the $m \times n$ matrix defined by the equations $T(v_j) = \sum_{i=1}^m a_{ij} w_i$, for $1 \le j \le n$. We denote this matrix by $[T]_\alpha^\beta$. In other words, $[T]_\alpha^\beta$ is the $m \times n$ matrix whose jth column is $\begin{pmatrix}a_{1j}\\ \vdots\\ a_{mj}\end{pmatrix}$, for $1 \le j \le n$.

We followed this by computing some examples

Example 1. Suppose A is the real matrix $\begin{pmatrix}a & c & e\\ b & d & f\end{pmatrix}$, and $T_A : \mathbb{R}^3 \to \mathbb{R}^2$ is given by $T_A(v) := A \cdot v$, for all column vectors $v \in \mathbb{R}$. We showed $[T_A]_E^F = A$, where E is the standard basis for $\mathbb{R}^3$ and F is the standard basis for $\mathbb{R}^2$.

Example 2. We then considered the special case in the previous example, where $A = \begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 2\end{pmatrix}$, so that, as above $[T_A]_E^F = A$. We then calculated the matrix of T_A with respect to the basis $C := \{\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\}$ for $\mathbb{R}^3$ and the basis $D := \{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\}$ for $\mathbb{R}^2$ and obtained $[T]_C^D = \begin{pmatrix}1 & 3 & 2\\ 1 & -1 & 1\end{pmatrix}$.

Example 3. Letting α denote the standard basis for $\mathbb{R}^2$ and $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(e_1) = \begin{pmatrix}8\\3\end{pmatrix}$ and $T(e_2) = \begin{pmatrix}-18\\-7\end{pmatrix}$, $[T]_\alpha^\alpha = \begin{pmatrix}8 & -18\\ 3 & -7\end{pmatrix}$, while if $\beta := \{\begin{pmatrix}3\\1\end{pmatrix}, \begin{pmatrix}2\\2\end{pmatrix}\}$, $[T]_\beta^\beta = \begin{pmatrix}2 & 0\\ 0 & -1\end{pmatrix}$, showing that T is **diagonalizable**.

We ended class with the following definition.

Definition. Let V be a vector space with basis $\alpha := \{v_1, \ldots, v_n\}$. Given $v \in V$, we have a unique expression $v = a_1 v_1 + \cdots + a_n v_n$, for $a_i \in F$. We set $[v]_\alpha := \begin{pmatrix}a_1\\ \vdots\\ v_n\end{pmatrix} \in F^n$.

We noted that this sets up a nice correspondence between vectors in an arbitrary vectors space V and column vectors in F^n. This notation also lead to the following property: For $v, u \in V$, $c, d \in F$ and α a basis for V, show that $[cv + dc]_\alpha = c[v]_\alpha + d[u]_\beta$.

Wednesday, February 7. We began by reviewing the definition of linear transformation, and gave four examples: (i) Multiplication by an $m \times n$ matrix as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$; (ii) The derivative map, as a function from $P(n)$ to $P(n)$; (iii) The trace function from $\mathrm{M}_{2\times}(\mathbb{R})$ to $\mathbb{R}$; and (iv) The rotation map in $\mathbb{R}^2$, given by the matrix $\begin{pmatrix} \cos(\theta & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$. We then presented the following theorem and its proof (modulo the uniqueness in part (ii)).

Proposition. Let V and W be vector spaces over F and suppose that $\alpha := \{v_1, \ldots, v_n\}$ is a basis for V.

(i) Suppose $T : V \to W$ is a linear transformation. Then T is determined by its values on the basis α.
(ii) We may define a unique linear transformation for V to W by specifying values in W for each $v_i \in \alpha$.

The point in (i) is that for any $v \in V$, we may write $v = a_1v_1 + \cdots + a_nv_n$, for $a_i \in F$, from which it follows that

$$\begin{aligned} T(v) &= T(a_1v_1 + \cdots + a_nv_n) \\ &= T(a_1v_1) + \cdots + T(a_nv_n) \\ &= a_1T(v_1) + \cdots + a_nT(v_n), \end{aligned}$$

showing that $T(v)$ is determined by the values $T(v_i)$, for $1 \leq i \leq n$. The point of (ii) is that if we specify $T(v_i) := w_i$ (say), for $w_i \in W$, then for any $v \in V$, if $v = b_1v_1 + \cdots + b_nv_n$, then function $T : V \to W$ defined by $T(v) = b_1w_1 + \cdots + b_nw_n$ is a linear transformation. In other words, to define a linear transformation $T : V \to W$, it suffices to assign to each $v_i \in W$, a vector w_i in W.

We finished class by defining the matrix of a linear transformation with respect to two bases and showed in a special case how a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ is ultimately just multiplication by a 2×2 matrix.

Monday, February 5. The first thirteen minutes of class were devoted to Quiz 3. We then began a discussion and proof of the following:

Theorem. Let V be a finite dimensional vector space.

(i) Suppose $S \subseteq V$ is a finite set of vectors satisfying $V = \text{Span}\{S\}$. Then some subset of S forms a basis for V.
(ii) Let $T \subseteq V$ be a linearly independent subset. Then T may be extended to a basis.

The proof of this theorem involved applications of the *Exchange Theorem* from the lecture of February 2. This gave rise to the following corollary:

Corollary. Suppose V is a vector space of dimension n and $S = \{v_1, \ldots, v_n\} \subseteq V$. The following are equivalent:

(i) S is a basis for V.
(ii) S is linearly independent.
(iii) $V = \text{Span}\{S\}$.

We then defined the concept of *linear transformation*: Given vector spaces V, W over F, the function $T : V \to W$ is a linear transformation if: (a) $T(v_1 + v_2) = T(v_1) + T(v_2)$ and (b) $T(\lambda v) = \lambda T(v)$, for all $v_1, v_2, v \in V$ and $\lambda \in F$.

We ended class by providing the following examples of linear transformations:

(i) $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T((\alpha, \beta)) = (3\alpha + 2\beta, -\alpha + \beta)$.
(ii) For an $m \times n$ matrix A with entries in $\mathbb{R}$, the function $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T_A(v) := A \cdot v$ is a linear transformation. Here we view the elements of $\mathbb{R}^n$ and $\mathbb{R}^m$ as column vectors.

Friday, February 2. We began class by restating, discussing, and giving a proof of a special case of the following theorem:

Exchange Theorem. Let $w_1, \ldots, w_s, u_1, \ldots, u_r$ be vectors in V and set $W := \text{Span}\{w_1, \ldots, w_s\}$. Assume that $u_1, \ldots, u_r$ are linearly independent and belong to W. Then $r \leq s$. Moreover, after re-indexing the w_i's, we have $W = \text{Span}\{u_1, \ldots, u_r, w_{r+1}, \ldots, w_s\}$. This latter property is called the *exchange property.*

We then recalled that it follows immediately from the theorem that any two bases for the finite dimensional vector space V have the same number of elements. This common number is called the *dimension of V*. We then noted the dimensions of the following spaces, in each case by exhibiting a basis for the indicated space:

(i) $\mathbb{R}^n$ is an n-dimensional vector space over $\mathbb{R}$.

(ii) The space of $n \times n$ matrices over $\mathbb{R}$ has dimension n^2.

(iii) The vector space of 2×2 matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ over $\mathbb{R}$ such that $3a + 2d = 0$ is a three-dimensional space.

(iv) The solution space to the systems of equations with reduced row echelon augmented matrix $\begin{pmatrix} 1 & 0 & 3 & 4 & \mid & 0 \\ 0 & 1 & -2 & 6 & \mid & 0 \end{pmatrix}$ is a two-dimensional subspace of $\mathbb{R}^4$.

Wednesday, January 31. After briefly recalling the notions of linear dependence and linear independence, we discussed at length the following important principle in linear algebra: *In any vector space, the number of vectors in a linearly independent set of vectors is always less than or equal to the number of vectors in a spanning set of vectors.* We noted that an immediate consequence of this fact is that any two bases for a finite dimensional vector space have the same number of elements. This enabled us to define the *dimension* of a vector space V to be the number of vectors in any basis of V.

To get a feeling for why the principle stated above is true, we considered the following: Suppose $v_1, v_2 \in \mathbb{R}^2$ span $\mathbb{R}^2$ and $u_1, u_2, u_2 \in \mathbb{R}^3$ are independent. We wanted to see that in fact, this scenario could not occur. For this, we saw that we could exchange v_1, v_2 to get (for example) u_1, u_2 to span $\mathbb{R}^3$. But then the independent vector u_3 would be in $\text{Span}\{u_1, u_2\}$, a contradiction. This helps to explain why any linearly independent set cannot have more elements than any spanning set.

We then showed that $\mathbb{R}^n$ has dimension n, $\mathbb{C}$ has dimension one as a vector space over $\mathbb{C}$, but $\mathbb{C}$ has dimension two as a vector space over $\mathbb{R}$. This was followed by demonstrating the following:

Proposition. Suppose $v_1, \ldots v_n$ are column vectors in $\mathbb{R}^n$ (or $\mathbb{C}^n$), and let A denote the $n \times n$ matrix whose columns are $v_1, \ldots, v_n$. Then the vectors form a basis for $\mathbb{R}^n$ (or $\mathbb{C}^n$) if and only $\det A) \neq 0$, or equivalently, A is an invertible matrix.

We finished class by giving examples of vectors that do, and do not, form a basis for $\mathbb{R}^2$ or $\mathbb{R}^3$ by calculating determinants and/or using Gaussian elimination.

Monday, January 29. The first thirteen minutes of class were devoted to Quiz 2. We then began class by recalling what it means for a set of vectors $v_1, \ldots, v_r$ in the vector space V to be either *linearly dependent* or *linearly independent*. In the case where V is the vector space of column vectors in $\mathbb{R}^n$ or $\mathbb{C}^n$, we noted that these conditions can be expressed in terms of the solutions to a homogeneous system of linear equations with coefficient matrix A, where A is the $n \times r$ matrix whose columns are $v_1, \ldots, v_r$. To wit, the homogeneous system $A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \vec{0}$ has a non-trivial solution if and only if $v_1, \ldots, v_r$ are linearly dependent. Equivalently, the homogeneous system $A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \vec{0}$ has a unique solution (namely $x_1 = 0, \ldots, x_r = 0$) if and only if the vectors $v_1, \ldots, v_r$ are linearly independent. We then used Gaussian elimination to show that a particular set of four vectors in $\mathbb{R}^4$ was linearly independent.

We followed this by demonstrating the:

Proposition. Suppose $W := \text{Span}\{v_1, \ldots, v_r\}$, for linearly independent vectors $v_1, \ldots, v_r \in V$. Then every vector $w \in W$ can be written *uniquely* as a linear combination of $v_1, \ldots, v_r$. In other words, if $w = \alpha_1 v_1 + \cdots + \alpha_r v_r = \beta_1 v_1 + \cdots + \beta_r v_r$, with all $\alpha_i, \beta_i \in F$, then $\alpha_i = \beta_i$, for all i.

We ended class by defining a set $S \subseteq V$ of vectors to be a *basis* for V if it spans V and is linearly independent. We then gave a few examples, that included the standard basis for $\mathbb{R}^3$ and the standard basis for $\mathrm{M}_{2\times 2}(\mathbb{R})$. We finished by mentioning the important fact that any two bases for a vector space have the same number of elements.

Friday, January 26. We began with the question: For vectors $w, v_1, \ldots, v_r$ in the vector space V over the field F, when is $w \in \text{Span}\{v_1, \ldots, v_r\}$? We noted that when $V = \mathbb{R}^n$ or $\mathbb{C}^n$, and the vectors $w, v_1, \ldots, v_r$

are column vectors, then $w \in \text{Span}\{v_1, \ldots, v_r\}$ if and only if the system of equations given by the matrix equation $A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = w$ has a solution. We also noted that any solution $\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix}$ to the system of equations gives rise to the relation $w = \alpha_1 v_1 + \cdots + \alpha_r v_r$. This was then illustrated by using Gaussian elimination on two specific examples. We then defined the vectors $\{v_1, \ldots, v_r\}$ to be *linearly dependent* if there exists a linear combination $\alpha_1 v_1 + \cdots + \alpha_r v_r = \vec{0}$, with at least one $\alpha_i \neq 0$. The set of vectors $\{v_1, \ldots, v_r\}$ is *linearly independent* if it is not linearly dependent. We continued by discussing and proving the following **very important** proposition - which has the consequence that we may discard a redundant vector from a set of vectors spanning a subspace W and still span W with one less vector.

Proposition. Vectors $v_1, \ldots, v_r \in V$ are linearly dependent if and only if for some $1 \leq i \leq r$, v_i belongs to $\text{Span}\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_r\}$. If these conditions hold, then

$$\text{Span}\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_r\} = \text{Span}\{v_1, \ldots, v_r\}.$$

We ended class with the observation that if $W := \text{Span}\{v_1, \ldots, v_r\}$, by using the concept of linear dependence, we may remove redundant vectors from our set $\{v_1, \ldots, v_r\}$ to arrive at a linearly independent generating set $\{v_{i_1}, \ldots, v_{i_s}\}$ for W. In other words, $v_{i_1}, \ldots, v_{i_s}$ generate W *efficiently*.

Wednesday, January 24. The first thirteen minutes of class were devoted to Quiz 1. We then began the lecture by reviewing the definition of subspace and establishing the following facts about subspaces:

Basic Facts about subspaces. Let V be a vector space over F and $W_1, W_2 \subseteq V$ be subspaces. Let $S := \{v_1, \ldots, v_r\}$ and $T := \{u_1, \ldots, u_s\}$ be subsets of V.

(i) $W_1 + W_2$ is a subspace of V, where $W_1 + W_2 := \{w_1 + w_2 \mid w_1 \in W_1 \text{ and } w_2 \in W_2\}$.
(ii) $W_1 \cap W_2$ is a subspace of V.
(iii) If $W_1 = \text{Span}\{S\}$ and $W_2 = \text{Span}\{T\}$, then $W_1 + W_2 = \text{Span}\{S \cup T\}$.

We then worked the following examples.

Examples. (i) Let W_1 be the x-axis in $\mathbb{R}^2$ and W_2 be the y-axis in $\mathbb{R}^2$, then $W_1 + W_2 = \mathbb{R}^2$.

(ii) Let W_1 be the xy-plane in $\mathbb{R}^3$ and W_2 be the z-axis in $\mathbb{R}^3$. Then $W_1 + W_2 = \mathbb{R}^3$.

(iii) Let W_1 be the line through (0,0,0) containing the vector $v_1 = (1,1,1)$ and W_2 be the line through (0,0,0) containing the vector $v_2 = (-1,0,1)$. Then $W_1 + W_2$ is the plane in $\mathbb{R}^3$ spanned by v_1 and v_2, which is given parametrically by $\{((s-t, s, s+t) \mid s, t \in \mathbb{R}\}$, or algebraically by the equation $x - 2y + z = 0$.

We ended class with a discussion of the *direct sum* of two subspaces. We said that V is the direct sum of W_1 and W_2 if: (i) $V = W_1 + W_2$ and $W_1 \cap W_2 = \vec{0}$. In this case we write $V = W_1 \oplus W_2$. We finished with the:

Proposition. Suppose $V = W_1 \oplus W_2$, for subspaces $W_1, W_2 \subseteq V$.

(i) If $w_1 \in W_1$, $w_2 \in W_2$ and $w_1 + w_2 = \vec{0}$, then $w_1 = \vec{0} = w_2$.
(ii) If $w_1, u_1 \in W_1$ and $w_2, u_2 \in W_2$, and $w_1 + w_2 = u_1 + u_2$, then $w_1 = u_1$ and $w_2 = u_2$.

Friday, January 19. We began class by reviewing the eight axioms that define a vector space V over $F = \mathbb{R}$ or $F = \mathbb{C}$. We then gave proofs of the following vector space properties, noting along the way how they either follow from the vector space axioms, or a previously established property.

Proposition. Let V be a vector space over F. The following properties hold:

(i) Cancellation holds: For all $u, v, w \in V$, if $v + w = v + u$, then $w = u$.
(ii) The additive identity $\vec{0}$ is unique.
(iii) $0 \cdot v = \vec{0}$, for all $v \in V$.
(iv) For any $v \in V$, its additive inverse $-v$ is unique.
(v) For all $\lambda \in F$ and $v \in V$, $-\lambda \cdot v = -(\lambda v)$. In particular, $-1 \cdot v = -v$, for all $v \in V$.

We then defined the concept of a *subspace.*

Definition. A subset W of the vector space V is a *subspace* if it satisfies the following conditions:

(i) $w_1 + w_2 \in W$, for all $w_1, w_2 \in W$.
(ii) $\lambda w \in W$, for all $\lambda \in F$ and $w \in W$.

After demonstrating that $\vec{0} \in W$ and $-w \in W$, for all $w \in W$, we noted that all remaining vector space axioms hold for W by virtue of them holding for V, so that W is a vector space in its own right, under the operations associated with V - which is the standard definition of subspace. We then noted that: lines through the origin in $\mathbb{R}^2$ are subspaces of $\mathbb{R}^2$; lines and planes through the origin in $\mathbb{R}^3$ are subspaces of $\mathbb{R}^3$; The solution set (as elements of $\mathbb{R}^n$) to a homogeneous system of m linear equations in n unknown is a subspace of $\mathbb{R}^n$; the set of all linear combinations of a finite set of vectors forms a subspace of the ambient vector space. This led to the following definition.

Definition. Let V be a vector space over F and $v_1, \ldots, v_r \in V$ be finitely many vectors. Then the *subspace spanned by* $v_1, \ldots, v_r$, denoted $\text{Span}\{v_1, \ldots, v_r\}$ or $\langle v_1, \ldots, v_r \rangle$, is the set of all vectors of the form

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r, \qquad (*)$$

where $\alpha_1, \alpha_2, \ldots, \alpha_r \in F$. Any expression of the form (*) above is called a *linear combination* of $v_1, \ldots, v_r$. We then noted that the set of all expressions (*) does indeed forms a subspace of V.

We ended class by observing the that unit vectors $e_1, e_2, e_3 \in \mathbb{R}^3$ span $\mathbb{R}^3$ and that adding any other vector to this set gives a set that still spans $\mathbb{R}^3$, but does so inefficiently. We also noted that if A_1, A_2, A_3, A_4 are the distinct 2×2 matrices having one entry 1 and the remaining entries 0, then the space of 2×2 real matrices is spanned by A_1, A_2, A_3, A_4.

Wednesday, January 17. We began class by looking at examples of vector spaces, initially, the vector space $\mathbb{R}^3$ of column vectors defined over the real numbers. Beginning with the basic properties of vector addition, where for $v_1 = \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \end{pmatrix}$, $v_1 + v_2 := \begin{pmatrix} \alpha_1 + \alpha_2 \\ \beta_1 + \beta_2 \\ \gamma_1 + \gamma_2 \end{pmatrix}$, and scalar multiplication, $\lambda v_1 := \begin{pmatrix} \lambda\alpha_1 \\ \lambda\beta_1 \\ \lambda\gamma_1 \end{pmatrix}$, we discussed the following properties (and verified a few of them), all which follow from similar familiar properties of $\mathbb{R}$:

(i) The zero vector $\vec{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ has the property that $\vec{0} + v = v$, for all $v \in \mathbb{R}^3$. (Existence of additive identity).
(ii) For $v = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$, $-v + v = \vec{0}$, where $-v := \begin{pmatrix} -\alpha \\ -\beta \\ -\gamma \end{pmatrix}$. (Existence of additive inverses)
(iii) $v_1 + v_2 = v_2 + v_1$, for all $v_1, v_2 \in \mathbb{R}^3$. (Commutativity of addition)
(iv) $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$, for all $v_i \in \mathbb{R}^3$. (Associativity of addition).
(v) $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$, for all $\lambda \in \mathbb{R}$ and $v_i \in \mathbb{R}^3$. (First distributive property)
(vi) $(\lambda + \gamma)v = \lambda v + \gamma v$, for all $\lambda, \gamma \in \mathbb{R}$ and $v \in \mathbb{R}^3$. (Second distributive property)
(vii) $(\lambda\gamma)v = \lambda(\gamma v)$, for all $\lambda, \gamma \in \mathbb{R}$ and $v \in \mathbb{R}^3$. (Associativity of scalar multiplication)
(viii) $1 \cdot v = v$, for all $v \in \mathbb{R}^3$.

We then looked at the vector space $P(2)$ of polynomials of degree two or less over $\mathbb{R}$ and noted that since a typical element in $P(2)$ has the form $\alpha + \beta x + \gamma x^2$, when we add two expressions of this form, or multiply them by a scalar, the resulting expressions look very similar to what we get when we add or scalar multiply vectors in $\mathbb{R}^3$. Something similar happens, if, for example, we take three vectors $u, v, w \in \mathbb{R}^{17}$ and consider all expressions of the form $\alpha u + \beta v + \gamma w$. This gives a vector space that looks very similar to $\mathbb{R}^3$ and $P(2)$. These examples show the advantage of defining vector spaces in an abstract setting in a way that captures all of the properties of particular vector spaces we might encounter in different contexts. This lead to the following:

Definition. Let F denote either $\mathbb{R}$ or $\mathbb{C}$. A **vector space over** F is a set V together with two operations, addition of elements of V and multiplication of elements from F times elements in V, satisfying the eight properties above:

(i) There exists a zero vector $\vec{0} \in V$ satisfying $v + \vec{0} = v$, for all $v \in V$. (Existence of additive identity).
(ii) For each $v \in V$, there exists $-v \in V$ such that $v + -v = \vec{0}$. (Existence of additive inverses)
(iii) $v_1 + v_2 = v_2 + v_1$, for all $v_1, v_2 \in V$. (Commutativity of addition)
(iv) $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$, for all $v_i \in V$. (Associativity of addition).
(v) $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$, for all $\lambda \in F$ and $v_i \in V$. (First distributive property)
(vi) $(\lambda + \gamma)v = \lambda v + \gamma v$, for all $\lambda, \gamma \in F$ and $v \in V$. (Second distributive property)
(vii) $(\lambda\gamma)v = \lambda(\gamma v)$, for all $\lambda, \gamma \in F$ and $v \in V$. (Associativity of scalar multiplication)
(viii) $1 \cdot v = v$, for all $v \in \mathbb{R}^3$.

We also noted that $\mathbb{R}^n$ and $\mathrm{M}_2(\mathbb{R})$, the set of 2×2 matrices over $\mathbb{R}$, form vector spaces over $\mathbb{R}$ and $\mathbb{C}^n$, with coordinate-wise addition and scalar multiplication, is a vector space over $\mathbb{C}$. We ended class by noting that in an abstract vector space, additive identities and additive inverses are unique.