SPRING 2025: MATH 590 DAILY UPDATE

Thursday, May 8. The class worked in groups on the practice problems for the final exam.

Tuesday, May 6. The class worked in groups on Quiz 12 and practice problems for the final exam.

Thursday, May 1. We continued our discussion of linear transformations with the goal of stating linear transformation versions of the main theorems we have considered during the semester. We began by recalling the definition. of $[T]_\alpha^\beta$, where $T : V \to W$ is a linear transformation and $\alpha \subseteq V, \beta \subseteq W$ are basis. For $v \in V$, we then defined $[v]_\alpha := \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in F^n$, if $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$, for the basis $\alpha = \{v_1, \ldots, v_n\}$. We then discussed the following properties, illustrating the **very important** property (iii) with an example.

Important Properties. In the notation above

(i) If $v, u \in V$ and $a, b \in F$, then $[au + bv]_\alpha = a[u]_\alpha + b[v]_\alpha$.
(ii) $[T(v)]_\beta = [T]_\alpha^\beta \cdot [v]_\alpha$.
(iii) If $S : W \to U$ is another linear transformation and $\gamma \subseteq U$ is a basis, then $[ST]_\alpha^\gamma = [S]_\beta^\gamma \cdot [T]_\alpha^\beta$.

Property (ii) then led to the

Change of Basis Formula. Suppose $T : V \to V$ is a linear transformation and $\alpha, \beta \subseteq V$ are bases. Then

$$[T]_\beta^\beta = [I]_\alpha^\beta \cdot [T]_\alpha^\alpha \cdot [I]_\beta^\alpha.$$

Moreover, $[I]_\alpha^\beta = ([I]_\beta^\alpha)^{-1}$, so that if we set $A := [T]_\alpha^\alpha$, $B := [T]_\beta^\beta$, and $P := [I]_\beta^\alpha$, we have $B = P^{-1}AP$.

We noted how this formula enables us to transcribe any of our major theorems about matrices to corresponding theorem about linear operators $T : V \to V$. For example, suppose we want a basis of V so that T has some desired form, based upon a certain property T might have. We take a basis (often an orthonormal basis) $\alpha \subseteq V$ and form the matrix $A = [T]_\alpha^\alpha$. We then invoke our theorem for matrices to find P such that $B = P^{-1}AP$ has the desired form. Let $C_1, \ldots, C_n$ be the columns of P, and choose v_i such that $[v_i]_\alpha = C_i$. Then if $\beta = \{v_1, \ldots, v_n\}$, $[T]_\beta^\beta = B$.

After defining the characteristic polynomial $p_T(x)$ for $T : V \to V$ and eigenvalues and eigenvectors for T, we stated

Diagonalizability Theorem. Suppose $T : V \to V$ is a linear operator. Then T is diagonalizable if and only if $p_T(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$ and $\dim E_{\lambda_i} = e_i$, for all $1 \leq i \leq r$, for $\lambda_i \in F$.

We then defined what it means for $T : V \to V$ to be symmetric, when V is an inner product space over $\mathbb{R}$: T is said to be symmetric if $\langle T(v), w\rangle = \langle v, T(w)\rangle$, for all $v, w \in V$. We noted (but did not prove) that this is equivalent to saying that the matrix $[T]_\alpha^\alpha$ is symmetric for all orthonormal bases α. We then stated

Real Spectral Theorem. Suppose V is an inner product space over $\mathbb{R}$ and $T : V \to V$ a linear operator. Then T is symmetric if and only if there exists an orthonormal basis $\alpha \subseteq V$ such that $[T]_\alpha^\alpha$ is diagonal. Equivalently, T is symmetric if and only if there exists an orthonormal basis of V consisting of eigenvectors for T, i.e., T is orthogonally diagonalizable.

Next, for a complex inner product space V, we then defined the adjoint $T^* : V \to V$, given $T : V \to V$, as the unique linear transformation satisfying $\langle T(v), w\rangle = \langle v, T^*(w)\rangle$, for all $v, w \in V$. We noted, but did not prove, that T^* can be defined by the equation $[T^*]_\alpha^\alpha = ([T]_\alpha^\alpha)^*$, for any choice of orthonormal basis α. Given this, $T : V \to V$ is defined to be a normal operator if $T^*T = TT^*$. Thus, we were able to state

Complex Spectral Theorem. Suppose V is an inner product space over $\mathbb{C}$ and $T : V \to V$ is a linear operator. Then T is normal if and only if there exists an orthonormal basis $\alpha \subseteq V$ such that $[T]_\alpha^\alpha$ is diagonal.

Equivalently, T is normal if and only if there exists an orthonormal basis of V consisting of eigenvectors for T, i.e., T is orthogonally (unitarily) diagonalizable.

We ended class by stating

Jordan Canonical Form Theorem. Let V be a vector space over $\mathbb{C}$ and $T : V \to V$ a linear operator. Then there exists a basis $\alpha \subseteq V$ such that $[T]_\alpha^\alpha = J$, where $J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J \end{pmatrix}$ is block diagonal and each block J_i is a Jordan block, of some size associated to some eigenvalue of T. We call J the *Jordan canonical form* of T.

Tuesday, April 29. The first fifteen minutes of class were devoted to Quiz 11. We then began a discussion of linear transformations. We defined a linear transformation from the vector space V to the vector space W to be a function $T : V \to W$ satisfying: (i) $T(v_1 + v_2) = T(v_1) + T(v_2)$, for all $v_1, v_2 \in V$ and
(ii) $T(\lambda v) = \lambda T(v)$, for all $\lambda \in F$ and $v \in V$. We then gave several examples of linear transformations, including:

(i) Multiplication by a matrix, i.e., suppose A is an $m \times n$ matrix over F. Define $T_A : F^n \to F^m$ by $T_A(v) = Av$. Then T_A is a linear transformation.
(ii) $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T(x, y) = (3x + 2y, 4x - y, 6x + 5y)$ is a linear transformation.
(iii) Differentiation from $P_n(\mathbb{R})$ to $P_{n-1}(\mathbb{R})$, is a linear transformation.
(iv) Rotation of $\mathbb{R}^2$ counterclockwise through an angle of θ is a linear transformation.

We also noted that one must have $T(\vec{0_V}) = \vec{0_W}$ if T is a linear transformation, so that translation, as a function from $\mathbb{R}^2$ to $\mathbb{R}^2$ is **not** a linear transformation. We then defined the *kernel* and *image* of a linear transformation, and after observing that these are subspaces of V and W, respectively, we stated and proved, the all important

Rank Plus Nullity Theorem. Let $T : V \to W$ be a linear transformation and suppose $\dim V = n$. Then $\dim \ker(T) + \dim \operatorname{im}(T) = n$.

The last few minutes of class were devoted to discussing the

Definition. Let $T : V \to W$ be a linear transformation. Suppose $\alpha = \{v_1, \ldots, v_n\}$ is a basis for V and $\beta = \{w_1, \ldots, w_m\}$ is a basis for W. Then the matrix of T with respect to α and β, denoted $[T]_\alpha^\beta$, is the $m \times n$ matrix (a_{ij}) whose i, j entries are determined by the equations $T(v_j) = \sum_{i=1}^m a_{ij} w_i$, for $1 \leq j \leq n$.

We finished class by considering $A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$ and the corresponding transformation $T_A : \mathbb{R}^3 \to \mathbb{R}^2$. We noted that if α, β are the standard bases for $\mathbb{R}^3$ and $\mathbb{R}^2$, then $[T]_\alpha^\beta = A$, while if $\alpha' = \{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\}$, then $[T]_{\alpha'}^\beta = \begin{pmatrix} a & a+b & a+b+c \\ d & d+e & d+e+f \end{pmatrix}$.

Thursday, April 24. We continued our discussion of the JCF, focussing on 3×3 matrices. We began with 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 $P^{-1}AP = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$.

We then turned to the case that $p_A(x) = (x-\lambda)^3$, and A is not diagonalizable. 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: 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 $P^{-1}AP = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$.

Tuesday, April 22. We began our discussion of the Jordan Canonical Form (JCF) for matrices over $\mathbb{C}$, recalling that working over $\mathbb{C}$ guarantees us access to all eigenvalue s of any square matrix A. We first considered the example $A = \begin{pmatrix} 1 & -2 \\ 2 & 5 \end{pmatrix}$, with $p_A(x) = (x-3)^2$ and $\dim(E_3) = 1$, so A is not diagonalizable. We found an invertible matrix P such that $P^{-1}AP = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$, the JCF of A. This was achieved by choosing $v_2 \notin E_3$ and setting $v_1 := (A-3I_3)v_2$, which as easily seen to be an eigenvector. Then $P = [v_1 v_2]$.

We then analyzed the 2×2 case in general, noting that the only way a 2×2 matrix A over $\mathbb{C}$ is not diagonalizable, is if $p_A(x) = (X-\lambda)^2$ and $\dim E_\lambda = 1$. We outlined the steps for finding the JCF of A: (i) Find $v_1 \in \mathbb{C}^2$ that is *not* an eigenvector for A and set $v_1 := (A-\lambda I_2)v_2$. This will automatically be an eigenvectors of A (via the 2×2 case of the Cayley-Hamilton Theorem). (ii) Set $P = [v_1 v_2]$. It followed that $P^{-1}AP = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

This was followed by a general discussion of the JCF. 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 $J(\lambda, 3) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$. We then stated the general result

Jordan Canonical Form Theorem. If A is an $n\times n$ matrix over $\mathbb{C}$ then there is an invertible matrix P, with entries in $\mathbb{C}$, such that $P^{-1}AP = J$, where $J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J \end{pmatrix}$ is block diagonal and each block J_i is a Jordan block, of some size associated to some eigenvalue of A. We call J the *Jordan canonical form* of 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 **important** facts that completely determine the JCF for 2×2 and 3×3 matrices. Let A be an $n \times n$ matrix over $\mathbb{C}$. 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 finished class by verifying (a) and (b) above for a $7{\times}7$ matrix with Jordan blocks $J(\lambda_1, 3), J(\lambda_1, 2), J(\lambda_2, 2)$.

Thursday, April 17. Exam 2. Tuesday, April 15. The class worked in groups on Quiz 10, and then on practice problems for Exam 2.

Thursday, April 10. We began class by reviewing the definition and properties of the adjoint of a complex matrix discussed at the end of the previous lecture. We also noted that symmetry is not the correct notion to discuss for complex matrices by showing that the symmetric complex matrix $A = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$ does not satisfy $\langle Av, w\rangle = \langle v, Aw\rangle$ in $\mathbb{C}^n$. When then recorded the **important facts**, that if A is a self-adjoint complex matrix then: (i) The eigenvalues of A are in $\mathbb{R}$ and (ii) Eigenvectors of A corresponding to distinct eigenvalues are orthogonal. We then noted that with these facts, and the properties of adjoints, the following theorem can be deduced in a similar manner to what we did for the real symmetric matrices.

First Complex Spectral Theorem. Let A be an $n \times n$ self-adjoint complex matrix. Then there exists a unitary matrix Q such that $Q^{-1}AQ = D$, diagonal matrix. Alternatively, if A is self-adjoint, then there is an orthonormal basis for $\mathbb{C}^n$ consisting of eigenvectors of A.

We then showed how to unitarily diagonalize the self-adjoint matrix $A = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$. This was followed by noting that unlike the case for real matrices, a complex matrix that is unitarily diagonalizable is not necessarily self-adjoint. We illustrated this by showing that the non-self-adjoint matrix $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is unitarily diagonalizable over $\mathbb{C}$ even though it is not even diagonalizable over $\mathbb{R}$. After noting that $B^*B = BB^*$, we stated the crucial

Definition. A complex $n \times n$ matrix A is said to be *normal* of $AA^* = A^*A$.

We noted, but did not prove, that if A is a normal matrix, then:

(i) $||Av|| = ||A^*v||$, for all $v \in \mathbb{C}^n$.
(ii) For $v \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$, $Av = \lambda v$ if and only if $A^*v = \overline{\lambda}v$.
(iii) Eigenvectors corresponding to distinct eigenvalues are orthogonal.

We ended class by informally discussing, but not proving

Second Complex Spectral Theorem. Let A be an $n \times n$ complex matrix. Then A is normal if and only if A is unitarily diagonalizable. Equivalently, A is normal if and only if there is an orthonormal basis for $\mathbb{C}^n$ consisting of eigenvectors of A.

Tuesday, April 8. The first twenty minutes of class were devoted to Quiz 9. Then, in preparation for stating a version of 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) $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$ 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 $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_n$. 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 then noted that the crucial property $\langle Av, w\rangle = \langle v, Aw\rangle$ for A a real symmetric matrix and $v, w \in \mathbb{R}^n$ fails over the complex numbers, as seen by taking $A = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$ and $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $w = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. However, we saw that by taking $A = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$, $v, w \in \mathbb{C}^n$ as before, then over $\mathbb{C}$ we have $\langle Av, w\rangle = v, Aw\rangle$. We noted that the new A has the property that it is equal to its *conjugate transpose*, i.e., $A = (\overline{A})^t$. This lead to the

Definition. Let A be a complex $n \times n$ matrix. The *adjoint* of A, denoted A^*, is the conjugate transpose of A, i.e., $A^* = (\overline{A})^t = \overline{A^t}$.

We then 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.

Thursday, April 3. We continued our discussion of the Singular Value Theorem (SVT) and the Singular Value Decomposition (SVD) by first stating

Singular Value Theorem. 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 $Q^tAP = \sum$ 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. Equivalently, $A = Q\sum P^t$, for $P, Q, \sum$ as before.

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 worked to find the singular value decomposition of the matrix $A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. We then discussed the following steps to finding the SVD of a matrix A, presenting proofs of some of the facts contained in these steps. he following steps

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 $v_i := \frac{1}{\sigma_i}Au_i$, where $u_1, \ldots, u_n \in \mathbb{R}^n$ 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 $v_1, \ldots, v_m$ 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 $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_r > 0$. The verification of $AP = Q\sum$ required the **key fact** that the null space of A^tA equals the null space of A.

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

Tuesday, April 1. The first twenty minutes of class were devoted to Quiz 8. We began class by reviewing how the key facts from the previous lectures give a strategy for proving the spectral theorem for real symmetric matrices. We then gave a proof of the key fact that if A is a real symmetric matrix, then $p_A(x)$ has all of its roots in $\mathbb{R}$. This relied on the Fundamental Theorem of Algebra, which guarantees that $p_A(x)$ has all of its roots in $\mathbb{C}$. We then took a root $\lambda = a + bi$ of $p_A(x)$ and an eigenvector $v \in \mathbb{C}^n$, so that $(A - \lambda I_n)v = \vec{0}$. We then applied $(A - \overline{\lambda}I_i)$ to both sides of this equation, and after completing a square, we argued that $b = 0$, so $\lambda \in \mathbb{R}$. For this, we needed to use that fact that there is a natural inner product on $\mathbb{C}^n$ satisfying $\langle v, v \rangle \neq 0$. We then noted:

Converse of the Spectral Theorem. If A is a real, $n \times n$ matrix that is orthogonally diagonalizable, then A is symmetric.

The rest of the lecture was devoted to illustrating the Singular Value Theorem for the matrix $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$. The key point was that A^tA is a symmetric matrix, and therefore orthogonally diagonalizable. We then found orthogonal matrices Q, P such that $Q^{-1}AP = \Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$, which can be thought of as a *pseudo-diagonalization* of A. Along the way, we saw that 3, 1 are the non-zero eigenvalues of A^tA and $\sqrt{3}, 1 = \sqrt{1}$ are called the *singular values* of A. We also noted that $A = Q\Sigma P^{-1} = Q\Sigma P^t$, which is the *singular value decomposition* of A.

Thursday, March 27. The first fifteen minutes of class were devoted to Quiz 7. We then resumed our discussion of the spectral theorem for symmetric matrices over $\mathbb{R}$, first by recalling the following key facts. If A is a real $n \times n$ symmetric matrix over $\mathbb{R}$, then: (i) Eigenvectors corresponding to distinct eigenvalues are pair-wise orthogonal, a fact that we proved in the previous lecture and (ii) A has all of its eigenvalues in $\mathbb{R}$, something we have yet to verify in general.

We then spent most of the rest of the lecture giving a careful proof of the 3×3 case of the spectral theorem. The key fact we used here was that the mean value theorem from calculus guarantees the existence of a real eigenvalue, since any degree three polynomial with coefficients in $\mathbb{R}$ has a root in $\mathbb{R}$. The proof then showed how the 3×3 case reduces to the 2×2 case. We pointed out that this is how the general case is handled. The second key fact above enables one to find a real eigenvalue, and thus a unit eigenvector u, and then the first fact allows one to show that $W^{\perp}$ is invariant under multiplication by A, thereby reducing the problem to a matrix whose dimension is one less than the original matrix A.

We then pointed out that, in practical terms, to orthogonally diagonalize a symmetric matrix A, one proceeds to diagonalize A in the usual way - it is guaranteed to be diagonalizable. One then uses Gram-Schmidt to find an orthonormal basis for each eigenspace. Then by the first key fact above, the union of these orthonormal bases is a orthonormal basis for $\mathbb{R}^n$, and hence the matrix Q whose columns are the basis elements for $\mathbb{R}^n$ orthogonally diagonalizes A. We finished class by applying this technique to orthogonally diagonalize the matrix $A = \begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{pmatrix}$.

Tuesday, March 25. We began class by reviewing several of the salient features regarding an inner product space V over $\mathbb{R}$. We then recalled the definition of an orthonormal basis for a subspace $W \subseteq V$ and how the Gram-Schmidt process gives rise to an orthonormal basis. This was followed by defining the *orthogonal complement* $W^{\perp}$ of W as the set $W^{\perp} := \{u \in V \mid \langle w, u\rangle = 0, \text{ for all } w \in W\}$. After observing that $W^{\perp}$ is a subspace of V, we then showed

Proposition. If V is a finite dimensional inner product space over $\mathbb{R}$ and $W \subseteq V$ is a subspace, then $V = W \bigoplus W^{\perp}$.

We then defined an $n\times n$ matrix Q over $\mathbb{R}$ to be *orthogonal* if and only if its columns form an orthonormal basis for $\mathbb{R}^n$. This was easily seen to be equivalent to $Q^t = Q^{-1}$. We noted that the rotation matrices $\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}$ are orthogonal. We were then able to state one of the main theorems of this course.

Spectral Theorem for Real Symmetric Matrices. Suppose A is an $n\times n$ symmetric matrix with entries in $\mathbb{R}$. The A is *orthogonally diagonalizable*, i.e., there exists an $n\times n$ orthogonal matrix Q such that $Q^{-1}AQ = D$, with D a diagonal matrix.

We discussed two key properties of symmetric matrices that make this theorem work:

(i) If $\lambda_1 \neq \lambda_2$ are distinct (real) eigenvalues of A, with eigenvectors v_1, v_2, then $\langle v_1, v_2\rangle = 0$. This followed from the fact that if A is symmetric, then for all $v, w \in \mathbb{R}^n$, $\langle Av, w\rangle = \langle v, Aw\rangle$, which followed from the general principle that for an $n\times n$ matrix B, $\langle Bv, w\rangle = \langle v, B^t w\rangle$.

(ii) A has all of its eigenvalues in $\mathbb{R}$. A proof of this statement was deferred to the next lecture.

We ended class by giving a careful proof of the Spectral Theorem for 2×2 matrices, and indicated informally the strategy for 3×3 matrices and matrices of larger dimension.

Thursday, March 13. We began class by reviewing the basic properties of inner product discussed in the previous lecture. We also finished an example begun at the end of the previous lecture demonstrating Basic Fact (ii) concerning an orthonormal basis from the previous lecture.

We then began a lengthy discussion of how to find an orthonormal basis for a subspace W of a vector space V, starting with a given basis. We did this in stages. We first took a vector space V with subspace W having basis $\{v_1, v_2\}$. Taking $w_1 = v_1$, we wrote $w_2 = v_2 - \alpha w_1$ and first noted that $\{w_1, w_2\}$ span W. We then set $\langle w_1, w_2\rangle = 0$ and found $\alpha = \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle}$, so that $w_2 := v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle}w_1$ is orthogonal to w_1.

We then analyzed the case W has basis $\{v_1, v_2, v_r\}$. We set $w_1 := v_1$. We also set $w_2 := v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle}w_1$, which by the previous case gives orthogonal vectors w_1, w_2. We then set $w_3 := v_3 - \alpha w_1 - \beta w_2$, and solved for α, β in the equations $\langle w_3, w_1\rangle = 0$ and $\langle w_3, w_2\rangle = 0$. This yielded $w_3 = v_3 - \frac{\langle v_3, w_1\rangle}{\langle w_1, w_1\rangle}w_1 - \frac{\langle v_3, w_2\rangle}{\langle w_2, w_2\rangle}w_2$, and thus an orthogonal basis $\{w_1, w_2, w_3\}$ for W. We then stated the

Gram-Schmidt Process. Let V be a vector space with inner product $\langle\ ,\ \rangle$ and suppose $\{v_1, \ldots, v_r\}$ is a basis for the subspace $W \subseteq V$. Then there exists an orthogonal set of vectors $\{w_1, \ldots, w_r\}$ which forms a basis for W. Moreover, the vectors $w_1, \ldots, w_r$ can be constructed inductively as follows:

(i) $w_1 := v_1$.
(ii) If $w_1, \dots, w_i$ have been constructed so that $\text{Span}\{w_1, \dots, w_i\} = \text{Span}\{v_1, \dots, v_i\}$ and $w_1, \dots, w_i$ 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 $\text{Span}\{w_1, \dots, w_{i+1}\} = \text{Span}\{v_1, \dots, v_{i+1}\}$ and $\{w_1, \dots, w_{i+1}\}$ is an orthogonal set of vectors. When $i + 1 = r$, the process is complete.

We then noted the immediate

Corollary. Let V be a vector space with an inner product and $W \subseteq V$ be a subspace. Then W has an orthonormal basis.

We ended class by starting with the basis $\{1, x, x^2\}$ for $P_2(\mathbb{R})$ with inner product $\langle f, g\rangle := \int_{-1}^{1} fg\ dx$, and using the Gram-Schmidt process to find an orthogonal basis. We also noted that the G-S process is inner product specific. In other words, if instead of defining $\langle f, g\rangle = \int_{-1}^{1} fg\ dx$ on $P_2(\mathbb{R})$, we defined an different inner product, $\{f, g\} := \int_0^1 fg\ dx$, then the process will lead to a different orthogonal basis for $P_2(\mathbb{R})$.

Tuesday, March 11. The first fifteen minutes of class were devoted to Quiz 6. We then once again stated the Diagonalizability Theorem and illustrated how the assumption that A is diagonalizable implies the conditions stated in the theorem for the case of a 7×7 matrix A satisfying $P^{-1}AP = D(\lambda_1, \lambda_1, \lambda_2, \lambda_2, \lambda_2, \lambda_3, \lambda_3)$.

We then began a discussion of inner products, first by reviewing the dot product of vectors in $\mathbb{R}^3$ and listing the various properties satisfied by the dot product. We noted that if $v, w \in \mathbb{R}^3$ are column vectors, then the dot product can be expressed as a matrix product $v^t \cdot w$. We then gave the following general definition.

Definition. Let V be a vector space over $\mathbb{R}$. An *inner product* on V is a function $\langle -, -\rangle : V \times V \longrightarrow \mathbb{R}$ satisfying the following properties for all $v_i, w_i \in V$ 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) $\langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_1\rangle$.
(iv) $\langle \lambda v, w\rangle = \lambda\langle v, w\rangle = \langle v, \lambda w\rangle$.
(v) $\langle v, v\rangle \geq 0$ and $\langle v, w\rangle = 0$ if and only if $v = \vec{0}$.

We pointed out that with these axioms, one can then define the notions of length and orthogonality, just like one does using the dot product. Namely, the length of v, denoted $||v||$, is $\sqrt{\langle, v, v\rangle}$ and the angle between v, w is $\cos^{-1}(\frac{\langle v,w\rangle}{||v||\cdot||w||})$. Thus, v *is orthogonal to* w if and only if $\langle v, w\rangle = 0$.

We then gave the following examples of inner products:
(i) $V = \mathbb{R}^n$, where for column vectors $v, w \in \mathbb{R}^n$, we define $\langle v, w\rangle = v^t \cdot w$ (matrix multiplication).
(iii) Let $V = P_n(\mathbb{R})$, be the space of real polynomials of degree less than or equal to n and define $\langle f(x), g(x)\rangle := \int_{-1}^{1} f(x)g(x)\ dx$.
(iii) Let $V = \text{M}_n(\mathbb{R})$, and set $\langle A, B\rangle := \text{tr}(A^tB)$.

This was followed by a discussion and proof of

Basic Fact. Suppose V has an inner product $\langle -, -\rangle$.
(i) If $v_1, \dots, v_r \in V$ are non-zero, mutually orthogonal vectors, then $v_1, \dots, v_r$ are linearly independent.
(ii) Suppose $\{u_1, \dots, u_r\}$ is a basis for V such that: (a) $u_1, \dots, u_r$ are mutually orthogonal and (b) Each u_i has length one. Then, for any $v \in V$,

$$v = \langle v, u_1\rangle u_1 + \langle v, u_2\rangle u_2 + \cdots + \langle v, u_r\rangle u_r.$$

We noted that the conditions in (ii) are equivalent to saying that $\langle u_i, u_j\rangle = 0$ if $i \neq j$ and $\langle u_i, , u_j\rangle = 1$, if $i = j$. A basis with this property is called an *orthonormal basis*, which we will abbreviate to ONB.

We ended class by showing that $u_1 := \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, u_2 := \frac{1}{\sqrt{3}} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, u_3 := \frac{1}{\sqrt{6}} \cdot \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ is an ONB for $\mathbb{R}^3$ and writing $v = \begin{pmatrix} 2 \\ 7 \\ 13 \end{pmatrix}$ in terms of this basis using (iib) above.

Thursday, March 6. We continued our discussion of diagonalizability by recalling the theorem presented at the end of the previous lecture. For an $n \times n$ matrix over F, with eigenvalue λ, we then defined the *algebraic multiplicity* of λ to be e if $p_A(x) = (x - \lambda)^e g(x)$, where $g(\lambda) \neq 0$ and, the *geometric multiplicity* of λ to be $\dim(E_\lambda)$. We then calculated each of these multiplicities for the matrices $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ and $\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$, noting that, in each case, the geometric multiplicity did not exceed the algebraic multiplicity, and that they were equal in the second case, which was diagonalizable. This lead to the following

Fundamental Relation Between Algebraic and Geometric Multiplicity. Suppose A is an $n \times n$ matrix with eigenvalue λ. Then the geometric multiplicity of λ is less than or equal to the algebraic multiplicity of λ.

We gave a proof of this theorem in class. **Note to students.** There was a typo in what I wrote in class today for the proof. In the step where we compared the first column of AP with the first column of PB, I should have written,

$$\lambda v_1 = b_{11} v_1 + b_{21} v_2 + \cdots + b_{n1} u_r,$$

instead of the last term being $b_{n1} A u_r$. The correct equation above is two linear combinations of the basis elements $v_1, v_2, \ldots, v_{e+1}, u_1, \ldots u_r$, which shows $\lambda = b_{11}$ and $0 = b_{j1}$, for $2 \leq j \leq n$. The rest of the proof continues as shown in class.

We then stated the all important

Diagonalizability Theorem. Let A be an $n \times n$ matrix. Then A is diagonalizable if and only if we can write $p_A(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$ and $\dim(E_{\lambda_1}) = e_i$, for all $1 \leq i \leq r$. Here we are assuming the λ_i are distinct and each $e_i \geq 1$.

Rather than giving a proof in the general case, we gave an in depth analysis of what happens when $n = 2$ or $n = 3$. We concluded this discussion by noting that if λ is an eigenvalue of A and $\dim(E_\lambda) = n$, then $A = D(\lambda, \lambda, \ldots, \lambda)$ was already a diagonal - in fact scalar - matrix.

Tuesday, March 4. We began a discussion of the diagonalizability of matrices. Starting with A, and $n \times n$ matrix over F, we defined A to be *diagonalizable* if there exists an invertible $n \times n$ matrix over F such that $P^{-1}AP = D$, where D is a diagonal matrix. We use the notation $D(\lambda_1, \ldots, \lambda_n)$ to denote the $n \times n$ diagonal matrix whose diagonal entries are $\lambda_1, \ldots, \lambda_n$. This was followed by a discussion of the following

Observations. In the notation above, we have:

(i) If $D = D(\lambda_1, \ldots, \lambda_n)$, then $\lambda_1, \ldots, \lambda_n$ are eigenvalues of A.
(ii) For $\lambda_1, \ldots, \lambda_n$ as in (i), these are the only eigenvalues of A.
(iii) A is diagonalizable if and only if F^n has a basis consisting of eigenvectors of A.
(iv) Suppose A is diagonalizable, and the distinct values on the diagonal of $P^{-1}AP$ are $\lambda_1, \ldots, \lambda_r$. Then $p_A(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_r)^{e_r}$.
(v) Suppose $Av = \lambda v$, for $v \in F^n$ and $\gamma \in F$. Then $(A - \gamma \cdot I_n)v = \vec{0}$, if $\lambda = \gamma$ or $(A - \gamma \cdot I_n)v = (\lambda - \gamma)v \neq \vec{0}$, if $\lambda \neq \gamma$.

We either proved each item in the observation, or illustrated an item by showing a proof when A is a 3×3 matrix. We also pointed out that the proof of (iv) showed that if $B = Q^{-1}AQ$, for Q an invertible $n \times n$ matrix, then $p_B(x) = p_A(x)$. We then emphasized the following

Important point. If A is diagonalizable, then $p_A(x)$ can be written as a product of linear polynomials, i.e., $p_A(x)$ has all of tis roots in F. However, the converse does not hold, i.e., if $p_A(x)$ has all of its roots in F, A need not be diagonalizable. The matrix $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ illustrates this point.

We finished class by discussing the following theorem and proving it for the case $n = 3$.

Theorem. Let A be an $n \times n$ matrix with entries in F and suppose that $\lambda_1, \ldots, \lambda_r$ are distinct eigenvalues of A. Take $v_1, \ldots, v_r$ such that v_i is an eigenvector associated to λ_i. Then, $v_1, \ldots, v_r$ are linearly independent. In particular, if A has n distinct eigenvalues in F, then A is diagonalizable.

We noted that the second statement follows from the first since if A has n distinct eigenvalues, it has n linearly independent eigenvectors. Since F^n has dimension n, these vectors form a basis for F^n, and hence A is diagonalizable, by Observation (iii) above.

Thursday, February 27. Exam 1,

Tuesday, February 25. The first fifteen minutes of class were devoted to Quiz 5. The remaining class time was spent working in groups on the practice problems for Exam 1.

Thursday, February 20. The first fifteen minutes of class were devoted to Quiz 4. We then stated and discussed the following theorem.

Theorem. Let A be an $n]timesn$ matrix with coefficients in $\mathbb{R}$ or $\mathbb{C}$.

(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) Any $n \times n$ system of linear equation with coefficient matrix A has a unique solution.

We then discussed the product rule for determinants: If A, B are $n \times n$ matrices, then $|AB| = |A| \cdot |B|$. We analyzed the 2×2 case by looking at elementary matrices, where an elementary matrix E is one obtained from the identity matrix by employing an elementary row operation. We then observed: (i) If E is obtained from I_2 by applying the row operation R, then EA is obtained by applying the row operation R to A and (ii) If E is an elementary matrix, $|EA| = |E| \cdot |A|$. The proof of the formula in the $n = 2$ case then followed from the facts that if $|B| \neq 0$, B is a product of elementary matrices and if $|B| = 0$, B row reduces to a matrix with one row consisting of 0s.

We ended a class with a discussion of eigenvalues and eigenvectors. If A is an $n \times n$ matrix over F, we defined the *characteristic polynomial* $p_A(x)$ by the equation $p_A(x) = |x \cdot I_n - A|$. An *eigenvalue* of A is $\lambda \in F$ such that $p_A(\lambda) = 0$. The vector $v \in F^n$ is an *eigenvector* associated to λ if $v \neq \vec{0}$ and $Av = \lambda v$. We then defined the *eigenspace* of λ to be the set of all vectors $v \in F^n$ such that $Av = \lambda v$, which was easily seen to be the null space of the matrix $A - \lambda \cdot I_n$. We then calculated the eigenvalues, eigenvectors and eigenspaces for the matrices $A = \begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Tuesday, February 18. Snow Day.

Thursday, February 13. 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. In particular, we willsutrated how this implies that the determinant id a multilinear function of its rows and columns. 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 and only if $|A| \neq 0$. We then derived Cramer's rule and illustrated it by solving a $2 \times x$ system of linear equations.

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 $1 \leq i \leq n$, $x_i = \frac{|B_i|}{|A|}$.

Tuesday, February 11. The first twenty minutes of class were devoted to Quiz 3. Following the quiz, we began a 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 or columns. We verifed these properties for 2×2 matrices. Letting A denote an $n \times n$ matrix over F:

(i) If A' is obtained form A by multiplying a row (or column) of A times $\lambda \in F$, then $|A'| = \lambda \cdot |A|$.
(ii) If A' is obtained from A by interchanging two rows (or two columns), then $|A'| = -|A|$.
(iii) If a row (or column) of A consists entirely of 0s, then $|A| = 0$.
(iv) If two 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.

Thursday, February 6. We began class by recalling that any two bases for the finite dimensional vector space V have the same number of elements. We recalled that this was an immediate consequence of the Exchange Theorem given in the previous lecture. This common number of elements in a basis 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 $\begin{pmatrix} 1 & 0 & 3 & 4 & | & 0 \\ 0 & 1 & -2 & 6 & | & 0 \end{pmatrix}$ is a two-dimensional subspace of $\mathbb{R}^4$.

We then noted that the dimension of V depends upon the scalars over which the space is defined, by showing that $\mathbb{C}^2$ has dimension two over $\mathbb{C}$, but as a vector space over $\mathbb{R}$ it has dimension four. This was followed by a discussion of:

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. This theorem gave rise to the following corollary:

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

(i) S is a basis for V.

(ii) S is linearly independent.

(iii) $V = \text{Span}\{S\}$.

We ended class by recalling that if $v_1, \dots, v_n$ are column vectors in F^n, then they form a basis for F^n if and only if the $n \times n$ matrix $A = [v_1\ v_2\ \cdots\ v_n]$ has an inverse and observing that if in $P(2)$ we wanted to show that $1 + x, 1 + x + x^2, 3x$ form a basis ofor$P(2)$ the space of polynomials of degree two or less, it suffices to show that $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$ form a basis for $\mathbb{R}^3$.

Tuesday, February 4. We began class with Quiz 2. After the quiz, we defined a subset $S \subseteq V$ to be a *basis* for V if: (i) $\text{Span}\{S\} = V$ and (ii) S is linearly independent. Thus, S is an *efficient* spanning set in that the vectors in S span V and upon deleting any vector from S, the resulting set does not span V. We gave several examples of bases, including the standard basis for $\mathbb{R}^n$. We then noted that if $S = \{v_1, \dots, v_n\}$ is a basis for V (or any subspace of V), then every vector in V can be written *uniquely* as a linear combination of $v_1, \dots, v_r$.

After looking at the case for column vectors in $\mathbb{R}^3$, we noted that n column vectors in $\mathbb{R}^n$ or $\mathbb{C}^n$ form a basis if and only the matrix whose columns are the given vectors is invertible. We then verified this in the particular case $v_1 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. This was followed by a discussion of the

Fundamental property. The number of elements in any linearly independent subset of vectors in the subspace $W \subseteq V$ is less than or equal to the number of vectors in any spanning set for W.

This then led to the

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

We illustrated the Exchange Theorem by taking a space spanned by two vectors and showing directly there could not be three linearly independent vectors in that space. The exchange property in this case was a consequence of this calculation.

We ended class by observing that it follows immediately from the Exchange Theorem that any two bases for a (finite dimensional) vector space have the same number of elements. We defined the number of elements in a basis to be the **dimension** of the vector space. After giving a few easy examples of the dimension of some familiar vectors spaces, we noted that the dimension of the space V depends on which scalars we are using, by observing that $\mathbb{C}$ is a two dimensional vector space over $\mathbb{R}$, but $\mathbb{C}$ is a one dimensional vector space over $\mathbb{C}$.

Thursday, January 30. We began class by recalling what it means for a set of vectors $v_1, \dots, v_r$ 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 $v_1, \dots, v_r$. 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 $v_1, \dots, v_r$ 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 $x_1 = 0, \dots, x_r = 0$) if and only if the vectors $v_1, \dots, v_r$ are linearly independent. We then used Gaussian elimination to show that a particular set of three vectors in $\mathbb{R}^4$ was linearly independent. We also showed how to use Gaussian elimination to determine if a column vector u belongs to the span of $v_1, \dots, vr$ and had the class work out a concrete example of this.

We then discussed the following

Basic Principle. For a subspace $W \subseteq V$ with $W = \text{Span}\{v_1, \dots, v_r\}$, suppose the vectors $v_1, \dots, v_r$ are linearly dependent. Then there exists v_i such that $W = \text{Span}\{v_1, \dots, \hat{v_i}, \dots, v_r\}$.

The point behind this basic principle is that the linear dependence assumption not only means that one of the given vectors is in the span of the remaining vectors, but that the remaining vectors span the same space as the original set of vectors. We then saw how to use Gaussian elimination to determine the redundant vector and also noted that the process of deleting redundant vectors in the spanning set can be repeated until one ultimately arrives at a spanning set of W that is linearly independent. In other words: *Any spanning set for W can be shortened to a spanning set that is linearly independent.*

We ended class by defining the set of vectors $S := \{w_1, \dots, w_r\} \subseteq W$ to be a *basis* for W if: (i) S spans W and (ii) S is linearly independent. We noted that the vectors $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ forms a basis for $\mathbb{R}^3$.

Tuesday, January 28. We began class with Quiz 1. After that, we showed that the intersection of two subspaces of a vector space is again a subs[ace. This was followed by noting that $\mathbb{R}^3$ is the direct sum of the spaces $W_1 :=^{\{} (1,1,1), -1,0,1)\}$ and $W_2 :=^{\{} (1,-2,1)\}$. We also noted the following:

Important Property. Suppose $V = W_1 + W_2$. Then $V = W_1 \bigoplus W_2$ if and only if every vector in V can be written *uniquely* as a sum of vectors from W_1 and W_2.

We then considered the question: For vectors $w, v_1, \dots, v_r$ in the vector space V over the field F, when is $w \in \text{Span}\{v_1, \dots, v_r\}$? We noted that when $V = \mathbb{R}^n$ or $\mathbb{C}^n$, and the vectors $w, v_1, \dots, v_r$ are column vectors, then $w \in \text{Span}\{v_1, \dots, v_r\}$ 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 where A is the $r \times n$ matrix whose columns are $v_1, \dots, v_r$. 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. We then defined the vectors $v_1, \dots, v_r \in V$ to be *linearly dependent* if some $v_i \in \text{Span}\{v_1, \dots \hat{v_i} \dots, v_r\}$ and noted that this was equivalent to having a *non-trivial dependence relation* on the v_i, i.e., there exists $a_1, \dots, a_r \in F$, not all zero, such that $a_1 v_1 + \cdots + a_r v_r = \vec{0}$. In other words, $v_1, \dots, v_r \in F^n$ are linearly dependent if and only if the system of equations $A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = \vec{0}$ has a non-trivial solution.

We finished class by defining the set of vectors $\{v_1, \dots, v_r\}$ to be *linearly independent* if they are not linearly dependent. From the previous discussion, it followed that the following conditions are equivalent:

(i) $v_1, \dots, v_r$ are linearly independent
(ii) No v_i belongs to $\text{Span}\{v_1, \dots, \hat{v_i} \dots, v_r\}$
(iii) If $a_1v_1 + \cdots + a_rv_r = \vec{0}$, for $a_j \in F$, then all $a_j = 0$

This led to the:

Important Consequence. If $W = \text{Span}\{v_1, \dots, v_r\}$ and $v_1, \dots, v_r$ are linearly independent, then $v_1, \dots, v_r$ span W *efficiently*. In other words, if we delete a vector v_i from the spanning set, the remaining vectors **do not** span W.

Thursday, January 23. We continued our discussion of subspaces of a vector space, including the following examples"

(i) The set of solutions to an $m \times n$ system of homogenous linear equations is a subspace of $\mathbb{R}^n$.
(ii) Given the vector space V, and $v_1, \dots, v_t \in V$, we defined $\text{Span}\{v_1, \dots, v_t\}$ to be the set of all linear combinations of $v_1, \dots, v_t$. This is a subspace of V.
(iii) We then noted that the space V of 2×2real matrices is spanned by the four matrices having one non-zero entry equal to 1 and all other entries equal to 0.
(iv) We showed that the set of 2×2 matrices is a subspace of V in (iii) and is spanned by the matrices $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

In order to motivate the background needed to study the spectral theorems, we worked through the details showing that the real symmetric matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ is *orthogonally diagonalizable*. In other words, we found a 2×2 orthogonal matrix Q such that $Q^{-1}AQ = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, where, by definition, a matrix is orthogonal if its columns are mutually orthogonal and have length one. Two key facts were observed that play a crucial role in the spectral theorem for real symmetric matrices: (i) The eigenvalues of A are in $\mathbb{R}$ and (ii) The eigenvectors associated to 2 are orthogonal to those associated to 0. We then briefly considered the symmetric matrix $B = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. We noted that its eigenvalues are 1, 4, with 1 occurring with multiplicity 2. We also noted that eigenvectors associated to distinct eigenvalues of B are orthogonal, but two independent vectors associated to 1 need not be. To achieve orthogonality among the eigenvectors of 1, we noted that we will ultimately need an orthogonalization process: Gram-Schmidt orthogonalization.

We finished class by defining the sum $W_1 + W_2$ of two subspaces contained in the vector space V and defined the sum to be *direct* if $W_1 \cap W_2 = \vec{0}$. We noted that $\mathbb{R}^2$ is the direct sum of any two lines through the origin; $\mathbb{R}^3$ is the direct sum of the xy-plane together with the z-axis; the space of 2×2 matrices over $\mathbb{R}$ is the direct sum of the matrices with trace zero together with the space spanned by the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

Tuesday, January 21. 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 $u, v, w \in \mathbb{R}^{19}$ 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.

We then discussed gave proofs of (some 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: $\{(0,0)\}$, $\mathbb{R}^2$, and lines through the origin in $\mathbb{R}^2$ are the subspaces of $\mathbb{R}^2$; $\{(0,0,0)\}$, $\mathbb{R}^3$, lines and planes through the origin in $\mathbb{R}^3$ are the subspaces of $\mathbb{R}^3$.