Definition 1. The Trace of a matrix $M$ is the sum of the diagonal entries.

Theorem 2. $\text{Tr}(AB) = \text{Tr}(BA)$. You have to keep them in order though, i.e. $\text{Tr}(ABC) \neq \text{Tr}(ACB)$

Proof. Write out the matrix product in terms of the entries to see they have the same diagonal sum. □

Theorem 3. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A$. Then $\text{Tr}(A) = \sum_{i=1}^{n} \lambda_i$. In fact, for any polynomial $p(x)$, we have $\text{Tr}(p(A)) = \sum_{i=1}^{n} p(\lambda_i)$.

Proof. Bring matrix to Jordan form, $A = SJS^{-1}$ Then $\text{Tr}(A) = \text{Tr}(SJS^{-1}) = \text{Tr}(JS^{-1}S) = \text{Tr}(J)$. Same trick works for powers of $A$, and since $\text{Tr}$ is linear, will work for polynomials. □

Definition 4. The Determinant of a matrix $A$ is defined to be “the volume of the parallelepiped spanned by the columns of $A$”

Theorem 5. $\det([v_1 \ v_2 \ \ldots \ v_n])$ is the unique function that is multilinear in the columns $v_1, \ldots, v_n$ with $\det(I_n) = 1$, which is antisymmetric in any of the columns $\det([v_1 \ v_2 \ \ldots \ v_n]) = -\det([v_2 \ v_1 \ \ldots \ v_n])$.

Theorem 6. $\det([v_1 \ v_2 \ \ldots \ v_n]) = 0$ if and only if $\{v_1, v_2, \ldots, v_n\}$ is dependent.

Remark 7. Multilinear means that if you fix all but one column, its linear in the last column e.g. $\det([\alpha v + w \ v_2 \ \ldots \ v_n]) = \alpha \det([v \ v_2 \ \ldots \ v_n]) + \det([w \ v_2 \ \ldots \ v_n])$. Below are two other useful formulas for $\det$:

Theorem 8. Permutation expansion $\det(M) = \det([M_1, \ldots, M_n]) = \sum_{\pi \in S_n} \text{sgn}(\pi) M_{\pi_1} M_{\pi_2} \cdots M_{\pi_n}$.

Theorem 9. Laplace expansion:

$$\det(M) = \det(M_{1, \ldots, M_{1j}e_1 + \cdots + M_{nj}e_n, \ldots, M_n})$$

$$= \sum_{i=1}^{n} M_{ij} \det(M_{1, \ldots, e_i, \ldots, M_n})$$

$$= \sum_{i=1}^{n} (-1)^{i+j} M_{ij} \det(M - \{i-th \ row \ and \ j-th \ column\})$$

Example 10. $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$

Definition 11. Inner product space: $<\cdot, \cdot>$ is sesquilinear. (That is, bilinear, except complex numbers come out of the second entry with a conjugate.)

Remark 12. You should think of this like a dot product. One thing you can do is see how projections onto a vector $v$ works. Draw a picture:

$\text{Proj}_v(w) = \frac{v}{||v||} (w \cdot \frac{v}{||v||})$

Theorem 13. [Graham-Schmidt Process] Given $\{a_1, \ldots, a_n\}$ there is an orthonormal set $\{v_1, \ldots, v_n\}$ so that they have the same span.

Proof. First set $v_1 = \frac{a_1}{||a_1||}$. Then draw a picture to figure out that we should set $u_2 = a_2 - \langle a_2, v_1 \rangle v_1$, to make $\langle u_2, v_1 \rangle = 0$. Normalize $u_2$ to get $v_2$ i.e. $v_2 = \frac{u_2}{||u_2||}$. Keep doing this! i.e. $u_3 = a_3 - \langle a_3, v_1 \rangle v_1 - \langle a_3, v_2 \rangle v_2$, $v_3 = \frac{u_3}{||u_3||}$. You can check your work at every step e.g. verify that $\langle u_3, v_2 \rangle = \langle u_3, v_1 \rangle = 0$. □

Remark 14. Notice that at every step, $\text{span}\{v_1, \ldots, v_k\} = \text{span}\{a_1, \ldots, a_k\}$. In particular, $\langle v_k, a_j \rangle = 0$ for $j < k$ since $a_j \in \text{span}\{v_1, v_2, \ldots, v_j\}$ which is all orthogonal to $v_k$. If we keep track of the Gram-Schmidt process in a matrix, we will get:

$\begin{bmatrix} a_1 & a_2 & \ldots & a_n \\ 1 & 0 & \ldots & 0 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \langle v_1, a_1 \rangle & \langle v_1, a_2 \rangle & \cdots & \langle v_1, a_n \rangle \\ \langle v_2, a_2 \rangle & \langle v_2, a_3 \rangle & \cdots & \langle v_2, a_n \rangle \\ 0 & 0 & \cdots & \langle v_n, a_n \rangle \end{bmatrix}$

This is known as the $QR$ DECOMPOSITION, every matrix can be written as a product of a unitary and upper triangular matrix. (Unitary just means the columns are orthonormal). To see this is true, you can multiply both sides by $\text{Proj}_v$: 
Theorem 15. (Schur Decomposition) Every matrix is orthogonally upper triangularizable: There is an orthonormal basis \( \{v_1, \ldots, v_n\} \) so that in this basis \( A \) is upper triangular.

Proof. (By induction of the size of the matrix)

Base Case: It’s obvious for a 1x1 matrix.

Induction Step: Suppose it works for every \((n-1) \times (n-1)\) matrix. For a \(n \times n\) matrix \( A \) choose any eigenvector \( v \) of \( A \).

WOLOG \( ||v|| = 1 \). (Exists since \( \det(A-\lambda I) \) has \( n \) roots!). Extend \( v \) to a basis for all of \( \mathbb{R}^n \) e.g. \( \{v, a_1, a_2, \ldots, a_{n-1}\} \) is a basis. Now apply the Graham to this basis to get an orthogonal basis \( \{v, u_1, \ldots, u_{n-1}\} \). In this basis the matrix for \( A \) is:

\[
[A]_{\{v, u_1, \ldots, u_{n-1}\}} = \begin{bmatrix}
\lambda & * & * & * \\
0 & & & \\
& \ddots & \ddots & \ddots \\
& & \dot{\ddots} & \ddots & \\
& & & \ddots & \ddots \\
0 & & & & \\
\end{bmatrix}
\]

Here \( \tilde{A} \) is the \((n-1) \times (n-1)\) matrix that you get, \( A : \text{span}\{u_1, \ldots, u_{n-1}\} \to \text{span}\{u_1, \ldots, u_{n-1}\} \) by \( \tilde{A}u_i = Au_i - v \langle Au_i, v \rangle \).

(This is the projection into \( \text{span}\{u_1, \ldots, u_{n-1}\} \)). By the induction hypothesis, there is an orthogonal basis \( \{w_1, \ldots, w_{n-1}\} \), so that \( \text{span}\{w_1, \ldots, w_{n-1}\} = \text{span}\{u_1, \ldots, u_n\} \) so that in this basis \( \tilde{A} \) is upper triangular. Notice now that \( \{v, w_1, \ldots, w_{n-1}\} \) is an orthonormal basis for \( \mathbb{R}^n \) and that in this basis, we have:

\[
[A] = \begin{bmatrix}
\lambda & * & * & * \\
0 & & & \\
& \ddots & \ddots & \ddots \\
& & \dot{\ddots} & \ddots & \\
& & & \ddots & \ddots \\
0 & & & & \\
\end{bmatrix}
\]

Which is upper triangular since \( [\tilde{A}]_{\{w_1, \ldots, w_n\}} \) is upper triangular! \( \square \)

Definition 16. The adjoint of \( L \) is the matrix \( L^* \) so that \( \langle v, Lw \rangle = \langle L^* v, w \rangle \) for every \( v, w \).

Remark 17. You can figure out \( L^* \) by choosing an orthonormal basis \( \{e_1, \ldots, e_n\} \) and notice that \( \langle L^* e_i, e_j \rangle = \langle e_i, Le_j \rangle \). Hence \( L^* e_i = \sum_{j=1}^n e_j \langle e_i, Le_j \rangle = \sum_{j=1}^n \langle Le_j, e_i \rangle e_j \) which characterizes \( L^* \). Notice that in this basis, because of this, the matrix \( L^* \) is the conjugate transpose of the matrix for \( L \).

Definition 18. A matrix \( A \) is called \textbf{normal} if \( AA^* = A^*A \).

Theorem 19. \( A \) is normal if and only if \( A \) is orthogonally diagonalizable. That is, there is an orthogonal basis where \( A \) is diagonal.

Proof. If \( A \) is orthogonally diagonalizable, write \( A = UDU^* \) and do a computation to see \( AA^* = A^*A \).

If \( A \) is normal, use the Schur decomposition to write \( A = UTU^* \) where \( T \) is upper triangular and \( U \) is the change of basis matrix to the orthogonal basis. Notice \( AA^* = A^*A \) implies that \( T^* T = TT^* \). Now, since \( T \) is upper triangular, we have a nice expression for the diagonal terms:

\[
(T^* T)_{ii} = \sum_{k=1}^n |t_{ki}|^2 = \sum_{k=1}^n |t_{ki}|^2 \\
(TT^*)_{ii} = \sum_{k=1}^n |t_{ik}|^2 = \sum_{k=1}^n |t_{ki}|^2
\]

Since these are equal, we get a series of equations:

\[
|t_{11}|^2 = |t_{11}|^2 + |t_{12}|^2 + \ldots + |t_{1n}|^2 \\
|t_{12}|^2 + |t_{22}|^2 = |t_{22}|^2 + \ldots + |t_{n2}|^2 \\
\vdots
\]

The first line implies that \( |t_{ii}| = 0 \) for \( i = 2, \ldots, n \). Using that in the second line, the second line says that \( |t_{i2}| = 0 \) for \( i = 3 \ldots n \). Doing this for all the equations we see that \( T \) must be diagonal! So the Schur decomposition gives us exactly an orthogonal diagonalization. \( \square \)

Definition 20. A matrix \( A \) is called \textbf{Hermitian} if \( A^* = A \). Notice that every Hermitian matrix is normal, and you check all the eigenvalues are real since \( \lambda = \langle Av, v \rangle = \langle v, A^* v \rangle = \lambda \).