October 30, 2013

SPECTRAL PROPERTIES OF SELF-ADJOINT MATRICES

RODICA D. COSTIN

Contents

1. Review

1.1. **The spectrum of a matrix.** If L is a linear transformation *on a finite dimensional vector space the set of its eigenvalues* $\sigma(L)$ *is called the* **spectrum** *of* L.

Note that: 1. the spectrum $\sigma(L)$ contains no information on the multiplicity of each eigenvalue;

2. $\lambda \in \sigma(L)$ if and only if $L - \lambda I$ is not invertible.

Remark: It will be seen that for linear transformations (*linear operators*) in infinite dimensional vector spaces the spectrum of L is defined using property 2. above, and it may contain more numbers than just eigenvalues of L.

1.2. **Brief overview of previous results.**

Let F denote the scalar field $\mathbb{R}$ or $\mathbb{C}$.

1. A matrix M is called *diagonalizable* if it is similar to a diagonal matrix: exists an invertible matrix S so that $S^{-1}MS = \Lambda$=diagonal. The diagonal entries of Λ are precisely the eigenvalues of M and the columns of S are eigenvectors of M.

2. *An* $n \times n$ *matrix is diagonalizable if and only if it has* n *linearly independent eigenvectors.*

3. If $T : F^n \to F^n$ is the linear transformation given by $T\mathbf{x} = M\mathbf{x}$ then M is the matrix associated to T in the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of F^n, while $S^{-1}MS$ is the matrix associated to T in the basis $S\mathbf{e}_1, \ldots, S\mathbf{e}_n$ of F^n (recall that $S\mathbf{e}_j$ is the column j of S, eigenvector of M).

4. Assume that the n-dimensional matrix M is diagonalizable, and let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be linearly independent eigenvectors. Let $S = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$. Then $S^{-1}MS = \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$.

$S^{-1}MS$ *is the matrix of the linear transformation* $F^n \to F^n$ *given by* $\mathbf{x} \mapsto M\mathbf{x}$ *in the basis of* F^n *consisting of the eigenvectors* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *of* M.

5. *Eigenvectors corresponding to different eigenvalues are linearly independent.*

As a consequence, an $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

6. More generally, a matrix M is diagonalizable if and only if for every eigenvalue λ the eigenspace $V_\lambda = \{\mathbf{v} \,|\, M\mathbf{v} = \lambda\mathbf{v}\}$ has dimension equal to the multiplicity of λ.

7. If the matrix M is not diagonalizable, then there exists an invertible matrix S (whose columns are eigenvectors or generalized eigenvectors of M) so that $S^{-1}MS = J$=Jordan normal form: a block diagonal matrix, consisting of Jordan blocks which have a repeated eigenvalue on the diagonal and 1 above the diagonal.

8. If $J_p(\lambda)$ is a Jordan $p\times p$ block, with λ on the diagonal, then any power $J_p(\lambda)^k$ is an upper triangular matrix, with λ^k on the diagonal.

9. Let $q(t)$ be a polynomial.

If M is diagonalizable by S: $S^{-1}MS = \Lambda = \text{diag}(\lambda_1,\ldots,\lambda_n)$ then $q(M)$ is also diagonalizable by S and $S^{-1}q(M)S = \Lambda = \text{diag}(q(\lambda_1),\ldots,q(\lambda_n))$.

If M is brought to a Jordan normal form by S: $S^{-1}MS = J$ then $q(M)$ is brought to an upper triangular form by S, having $q(\lambda_j)$ on the diagonal.

As a consequence:

Theorem 1. The spectral mapping theorem. *The eigenvalues of $q(M)$ are precisely $q(\lambda_1),\ldots,q(\lambda_n)$, where $\lambda_1,\ldots,\lambda_n$ are the eigenvalues of M.*

Similar results hold for more general functions, when $q(t) = \sum_{k=0}^{\infty} c_k t^k$ and the series has radius of convergence strictly greater than $\max_j |\lambda_j|$, for example, for $q(t) = \exp(t)$.

1.3. **More on similar matrices.**

Recall that similar matrices have the same eigenvalues.

Here is an additional result[1], to complete the picture (it is not proved here):

Theorem 2. *Two matrices are similar: $S^{-1}MS = N$ if and only if M and N have the same eigenvalues, and the dimensions of their corresponding eigenspaces are equal: $\dim V_{\lambda_j}^{[M]} = \dim V_{\lambda_j}^{[N]}$ for all j.*

2. Self-adjoint matrices

2.1. **Definitions.**

Definition 3. *Let $(V,\langle,\rangle)$ be an inner product space. The linear transformation $L : V \to V$ is called* **self-adjoint** *if $L^* = L$, that is, if*

$$\langle L\mathbf{x},\mathbf{y}\rangle = \langle \mathbf{x}, L^*\mathbf{y}\rangle \quad \textit{for all } \mathbf{x},\mathbf{y}\in V$$

Recall that the matrix M of a linear transformation L with respect to orthonormal bases is related to the matrix M^* of L^* by $M^* = \overline{M^T}$ $(=\overline{M}^T)$. Note that

$$(M^*)^* = M, \qquad (MN)^* = N^*M^*$$

[1]See P.D. Lax, *Linear Algebra and Its Applications*, Wiley, 2007.

Recall that if $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is an orthonormal basis of V then the inner product is the usual dot product of coordinates (also called the Euclidian inner product):

$$\text{if } \mathbf{x} = \sum_{k=0}^{n} x_k \mathbf{u}_k,\ \mathbf{y} = \sum_{k=0}^{n} y_k \mathbf{u}_k \ \text{ then } \ \langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=0}^{n} \overline{x_k} \mathbf{y}_k$$

So it suffices (for a while) to assume $V = \mathbb{R}^n$ or $V = \mathbb{C}^n$ equipped with the Euclidian inner product:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} x_j y_j \quad \text{on } \mathbb{R}^n \tag{1}$$

and respectively[2]

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} \overline{x_j} y_j \quad \text{on } \mathbb{C}^n \tag{2}$$

For a unitary treatment we write $V = F^n$ and use the inner product (2). Of course for $F = \mathbb{R}$ the inner product (2) is just (1).

We will often use interchangeably the expressions "the matrix M" and "the linear transformation $\mathbf{x} \mapsto M\mathbf{x}$".

Definition 4. *A matrix A is called* **self-adjoint** *if $A = A^*$.*

Note that only square matrices can be self-adjoint, and that $A = A^*$ means, entrywise, that $A_{ij} = \overline{A_{ji}}$ (elements which are positioned symmetrically with respect to the diagonal are complex conjugates of each other).

When it is needed (or just desired) to distinguish between matrices with real coefficients and those with complex coefficients (perhaps not all real), the following terms are used:

Definition 5. *A self-adjoint matrix with real entries is called* **symmetric***.*
A self-adjoint matrix with complex entries is called **Hermitian***.*

Note that a symmetric matrix A satisfies $A^T = A$, hence its entries are symmetric with respect to the diagonal.

Notations.
$\mathcal{M}_n$ denotes the set of $n \times n$ matrices with entries in $\mathbb{C}$.
$\mathcal{M}_n(\mathbb{R})$ denotes the set of $n \times n$ matrices with entries in $\mathbb{R}$.

2.2. **Self-adjoint matrices are diagonalizable I.** We start with a few special properties of self-adjoint matrices.

Proposition 6. *If $A \in \mathcal{M}_n$ is a self-adjoint matrix: $A = A^*$, then*

$$\langle \mathbf{x}, A\mathbf{x} \rangle \in \mathbb{R} \quad \textit{for all } \mathbf{x} \in \mathbb{C}^n \tag{3}$$

[2]Some texts use conjugation in the second argument, rather than in the first one. Make sure you know the convention used in the text you are reading.

Proof:

$$\langle \mathbf{x}, A\mathbf{x}\rangle = \langle A^*\mathbf{x}, \mathbf{x}\rangle = \langle A\mathbf{x}, \mathbf{x}\rangle = \overline{\langle \mathbf{x}, A\mathbf{x}\rangle}$$

hence (3). $\square$

Proposition 7. *If $A \in \mathcal{M}_n$ is a self-adjoint matrix: $A = A^*$, then all its eigenvalues are real: $\sigma(A) \subset \mathbb{R}$.*

Proof:
Let $\lambda \in \sigma(A)$. Then there is $\mathbf{v} \in \mathbb{C}^n$, $\mathbf{v} \neq \mathbf{0}$ so that $A\mathbf{v} = \lambda\mathbf{v}$. Then on one hand

$$\langle \mathbf{v}, A\mathbf{v}\rangle = \langle \mathbf{v}, \lambda\mathbf{v}\rangle = \lambda\langle \mathbf{v}, \mathbf{v}\rangle = \lambda\|\mathbf{v}\|^2$$

and on the other hand

$$\langle \mathbf{v}, A\mathbf{v}\rangle = \langle A^*\mathbf{v}, \mathbf{v}\rangle = \langle A\mathbf{v}, \mathbf{v}\rangle = \langle \lambda\mathbf{v}, \mathbf{v}\rangle = \overline{\lambda}\langle \mathbf{v}, \mathbf{v}\rangle = \overline{\lambda}\|\mathbf{v}\|^2$$

therefore $\lambda\|\mathbf{v}\|^2 = \overline{\lambda}\|\mathbf{v}\|^2$ and since $\mathbf{v} \neq \mathbf{0}$ then $\lambda = \overline{\lambda}$ hence $\lambda \in \mathbb{R}$. $\square$

If a matrix is symmetric, not only its eigenvalues are real, but its eigenvectors as well:

Proposition 8. *If A is a symmetric matrix then all its eigenvectors are real.*

Indeed, the eigenvalues are real by Proposition 7. Then the eigenvectors are real since they are solutions linear systems with real coefficients (which can be solved using $+, -, \times, \div$, operations that performed with real numbers do yield real numbers (as opposed to solving polynomials equations, which may have nonreal solutions). $\square$

For self-adjoint matrices, eigenvectors corresponding to distinct eigenvalues are not only linearly independent, they are even orthogonal:

Proposition 9. *If $A \in \mathcal{M}_n$ is a self-adjoint matrix: $A = A^*$, then eigenvectors corresponding to distinct eigenvalues are orthogonal: if $\lambda_{1,2} \in \sigma(A)$ and $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$, $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$ ($\mathbf{v}_{1,2} \neq \mathbf{0}$) then*

$$\lambda_1 \neq \lambda_2 \Longrightarrow \mathbf{v}_1 \perp \mathbf{v}_2$$

Proof:
On one hand

$$\langle \mathbf{v}_1, A\mathbf{v}_2\rangle = \langle \mathbf{v}_1, \lambda_2\mathbf{v}_2\rangle = \lambda_2\langle \mathbf{v}_1, \mathbf{v}_2\rangle$$

and on the other hand

$$\langle \mathbf{v}_1, A\mathbf{v}_2\rangle = \langle A^*\mathbf{v}_1, \mathbf{v}_2\rangle = \langle A\mathbf{v}_1, \mathbf{v}_2\rangle = \langle \lambda_1\mathbf{v}_1, \mathbf{v}_2\rangle = \lambda_1\langle \mathbf{v}_1, \mathbf{v}_2\rangle$$

(since the eigenvalues are real, by Proposition 7). Therefore $\lambda_2\langle \mathbf{v}_1, \mathbf{v}_2\rangle = \lambda_1\langle \mathbf{v}_1, \mathbf{v}_2\rangle$ and since $\lambda_1 \neq \lambda_2$ then $\langle \mathbf{v}_1, \mathbf{v}_2\rangle = 0$. $\square$

As a consequence of Proposition 9: *if $A = A^*$ and all the eigenvalues of A are distinct, then the n independent eigenvectors form an orthogonal set.* We can normalize the eigenvectors, to be unit vectors, and then the eigenvectors form an orthonormal set, hence the matrix S which by conjugation diagonalizes A is a unitary matrix:*there is U unitary so that U^*AU =diagonal.*

In fact this is true for general self-adjoint matrices, as stead in Theorem 16. Its proof is included in §2.6 and requires establishing additional results, which are important in themselves.

2.3. **Further properties of unitary matrices.**

Proposition 10. *Every eigenvalue of a unitary matrix U has absolute value 1: $\sigma(U) \subset S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$.*

Proof:
Let λ be an eigenvalue of U: $U\mathbf{v} = \lambda\mathbf{v}$ for some $\mathbf{v} \neq \mathbf{0}$. Then $\|U\mathbf{v}\| = \|\lambda\mathbf{v}\|$ and since U is an isometry then $\|\mathbf{v}\| = \|\lambda\mathbf{v}\|$ which implies $\|\mathbf{v}\| = |\lambda|\,\|\mathbf{v}\|$ which implies $|\lambda| = 1$ since $\mathbf{v} \neq \mathbf{0}$. □

Exercises.
1. Show that the product of two unitary matrices is also a unitary matrix.
2. Show that the determinant of a unitary matrix has absolute value 1. What is the determinant of an orthogonal matrix?

Proposition 11. *Eigenvectors of a unitary matrix U corresponding to different eigenvalues are orthogonal.*

Proof:
Let $U\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $U\mathbf{v}_2 = \lambda_2\mathbf{v}_2$ ($\mathbf{v}_j \neq \mathbf{0}$). Since U preserves angles, $\langle U\mathbf{v}_1, U\mathbf{v}_2\rangle = \langle \mathbf{v}_1, \mathbf{v}_2\rangle$ which implies $\langle \lambda_1\mathbf{v}_1, \lambda_2\mathbf{v}_2\rangle = \langle \mathbf{v}_1, \mathbf{v}_2\rangle$ therefore $\overline{\lambda_1}\lambda_2\langle \mathbf{v}_1, \mathbf{v}_2\rangle = \langle \mathbf{v}_1, \mathbf{v}_2\rangle$. For $\lambda_1 \neq \lambda_2$ we have $\overline{\lambda_1}\lambda_2 \neq 1$ (using Proposition 10) therefore $\langle \mathbf{v}_1, \mathbf{v}_2\rangle = 0$. □

We will see a bit later that *unitary matrices are diagonalizable.*

Example. Consider the matrix

$$Q = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

(its action on $\mathbb{R}^3$ is a renumbering of the axes: $\mathbf{e}_j \mapsto \mathbf{e}_{j-1}$ cyclically). The columns of Q form an orthonormal set, therefore Q is a unitary matrix. Since its entries are real numbers, then Q is an orthogonal matrix.

The characteristic polynomial of Q is easily found to be $1 - \lambda^3$, therefore its eigenvalues are the three cubic roots of 1, namely $1, (-1 \pm i\sqrt{3})/2$. The eigenvector corresponding to 1 is $(1, 1, 1)^T$. In the plane orthogonal to this eigenvector is spanned by the other two eiegnevctors (rather, the real and imaginary parts, if we choose to work in $\mathbb{R}^3$) the action of Q is a rotation seen geometrically, and algebraically by the presence of the two complex eigenvalues.

2.4. **An alternative look at the proof of Theorem 16.** *(Optional)* In this section we give an alternative argument, with the scope of enhancing our intuition on self-adjoint matrices.

If we accept the result that the direct sum of the generalized eigenspaces is the whole vector space (which was not proved here), here is a short argument which shows that for a *self-adjoint matrix* A the dimension of any eigenspace V_λ equals the multiplicity of λ. By choosing an orthonormal basis for each eigenspace we then find a unitary matrix that diagonalizes A, thus establishing Proposition **??** for *any* self-adjoint matrix.

Claim:

If $A = A^*$ then every generalized eigenvector is a true eigenvector: $E_\lambda = V_\lambda$.

Indeed, let $\mathbf{x}$ be a generalized eigenvector of A: $(A - \lambda I)^k\mathbf{x} = \mathbf{0}$, where k is the smallest such integer, that is, $(A - \lambda I)^{k-1}\mathbf{x} \neq \mathbf{0}$. Assume, to get a contradiction, that $k \geq 2$. Then, denoting $(A - \lambda I)^{k-1}\mathbf{x} = \mathbf{v}$ it follows that $(A - \lambda I)\mathbf{v} = \mathbf{0}$, which means that $\mathbf{v}$ is an eigenvector (it is not $\mathbf{0}$ since k is the smallest). Note that equation $(A - \lambda I)\mathbf{x}_2 = \mathbf{v}$ is satisfied for $\mathbf{x}_2 = (A - \lambda I)^{k-2}\mathbf{x}$. Then

$$\langle (A - \lambda I)\mathbf{x}_2, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle$$

and since A is self-adjoint, then λ is real, so this implies

$$\langle \mathbf{x}_2, (A - \lambda I)\mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle$$

hence $0 = \langle \mathbf{v}, \mathbf{v}\rangle$, therefore $\mathbf{v} = \mathbf{0}$ which is a contradiction. $\square$

Combining the *Claim* above, Propositions 8 and **??** we find that in the special case when A is symmetric then the matrix U of Propositions **??** is real, therefore an orthogonal matrix, so:

Proposition 12. *If $A \in \mathcal{M}_n(\mathbb{R})$ is a symmetric matrix, $A^T = A$, then A is diagonalizable by conjugation using an orthogonal matrix: $Q^TAQ =$diagonal for some Q orthogonal matrix.*

2.5. **Triangularization by conjugation using a unitary matrix.** Diagonalization of matrices is not always possible, and even when it is, it is computationally expensive. However, matrices can be easily brought to a triangular form, which suffices for many applications:

Theorem 13. *Given any square matrix $M \in \mathcal{M}_n$ there is an n dimensional unitary matrix U so that $U^*MU = T$=upper triangular.*

Of course, the diagonal elements of T are the eigenvalues of M.

Proof:

The matrix U is constructed in successive steps.

1^o. Choose an eigenvalue λ_1 of M and a corresponding *unit* eigenvector $\mathbf{u}_1$. Then complete $\mathbf{u}_1$ to an orthonormal basis of $\mathbb{C}^n$: $\mathbf{u}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ (recall that this is always possible!). Let U_1 be the unitary matrix

$$U_1 = [\mathbf{u}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] = [\mathbf{u}_1 \,\big|\, X_1]$$

The goal is to simplify M by replacing it with $U_1^* M U_1$. Note that

(4)

$$U_1^* M U_1 = \begin{bmatrix} \mathbf{u}_1^* \\ - \\ X_1^* \end{bmatrix} M[\mathbf{u}_1 \,\big|\, X_1] = \begin{bmatrix} \mathbf{u}_1^* \\ - \\ X_1^* \end{bmatrix} [M\mathbf{u}_1 \,\big|\, MX_1] = \begin{bmatrix} \mathbf{u}_1^* \\ - \\ X_1^* \end{bmatrix} [\lambda_1 \mathbf{u}_1 \,\big|\, *]$$

$$= \begin{bmatrix} \lambda_1 \mathbf{u}_1^* \mathbf{u}_1 & | & * \\ - & & - \\ \lambda_1 X_1^* \mathbf{u}_1 & | & * \end{bmatrix} = \begin{bmatrix} \lambda_1 & | & * \\ - & & - \\ 0 & | & * \end{bmatrix} \equiv \tilde{M}_1$$

where we used the fact that the vector $\mathbf{u}_1$ is a unit vector, and that the columns of X_1 are orthogonal to $\mathbf{u}_1$.

Denote by M_1 the $(n-1) \times (n-1)$ lower right submatrix of $\tilde{M}_1$:

$$\tilde{M}_1 = \begin{bmatrix} \lambda_1 & | & * \\ - & & - \\ 0 & | & M_1 \end{bmatrix}$$

Note that $Sp(\mathbf{e}_2, \ldots, \mathbf{e}_n) \equiv \mathbb{C}^{n-1}$ is an invariant space for $U_1^* M U_1$ and $U_1^* M U_1$ acts on $\mathbb{C}^{n-1}$as multiplication by M_1.

Note also that $\sigma(M) = \sigma(\tilde{M}_1)$ (conjugate matrices have the same eigenvalues) and that $\sigma(\tilde{M}_1) = \{\lambda_1\} \cup \sigma(\tilde{M}_1)$.

2^o. We repeat the first step for the $n-1$ dimensional matrix M_1: let λ_2 be an eigenvalue of M_1, and $\mathbf{u}_2 \in \mathbb{C}^{n-1}$ a unit eigenvector, and complete it to an orthonormal basis $\mathbf{u}_2, \mathbf{y}_2, \ldots, \mathbf{y}_n$ of $\mathbb{C}^{n-1}$. If $U_2 = [\mathbf{u}_2, \mathbf{y}_2, \ldots, \mathbf{y}_n]$ then

(5)
$$U_2^* M_1 U_2 = \begin{bmatrix} \lambda_2 & | & * \\ - & & - \\ 0 & | & M_2 \end{bmatrix}$$

where $M_2 \in \mathcal{M}_{n-2}$.

We can extend U_2 to an n dimensional unitary matrix by

$$\tilde{U_2} = \begin{bmatrix} 1 & | & 0 \\ - & & - \\ 0 & | & U_2 \end{bmatrix}$$

and it is easy to check that

$$\tilde{U_2}^* \tilde{M}_1 \tilde{U_2} = \begin{bmatrix} \lambda_1 & * & | & * \\ 0 & \lambda_2 & | & * \\ - & - & & - \\ 0 & 0 & | & M_3 \end{bmatrix} \equiv \tilde{M}_2$$

Note that the matrix $U_1 \tilde{U_2}$ (which is unitary) conjugates M to $\tilde{M}_2$.

$3^o \dots n^o$ Continuing this procedure we obtain the unitary $U = U_1\tilde{U_2}\dots\tilde{U_n}$ which conjugates M to an upper triangular matrix. $\Box$

2.6. **All self-adjoint matrices are diagonalizable II.** Let A be a self-adjoint matrix: $A = A^*$. By Theorem 13 there is a unitary matrix U so that $U^*AU = T$ = upper triangular. The matrix U^*AU is self-adjoint, since $(U^*AU)^* = U^*A^*(U^*)^* = U^*AU$ and a triangular matrix which is self-adjoint must be diagonal! We proved Theorem 16:

Any self-adjoint matrix A is diagonalizable and there is U unitary so that U^*AU =diagonal.

2.7. **Normal matrices.** We saw that any self-adjoint matrix is diagonalizable, has a complete set of orthonormal eigenvectors, and its diagonal form is real (by Proposition 7). It is natural to inquire: what are the matrices which are diagonalizable also having a complete set of orthonormal eigenvectors, but having possible nonreal eigenvalues?

It is easy to see that such matrices have special properties. For example, if $N = U\Lambda U^*$ for some unitary U and diagonal Λ, then, by taking the adjoint, $N^* = U\overline{\Lambda}U^*$ and it is easy to see that N commutes with its adjoint:

$$NN^* = N^*N \tag{6}$$

It turns out that condition (6) suffices to ensure that a matrix is diagonalizable by a unitary. Indeed, we know that any matrix can be conjugated to an upper triangular form by a unitary: $U^*MU = T$ as in Theorem 2.5; therefore also $U^*M^*U = T^*$. If M satisfies $MM^* = M^*M$, then also $TT^* = T^*T$; a simple calculation shows that this can only happen if T is, in fact, diagonal.

For example, in the 2-dimensional case:

$$\text{for } T = \begin{bmatrix} \lambda_1 & \alpha \\ 0 & \lambda_2 \end{bmatrix} \ \text{ then } T^* = \begin{bmatrix} \overline{\lambda_1} & 0 \\ \overline{\alpha} & \overline{\lambda_2} \end{bmatrix}$$

and therefore

$$TT^* = \begin{bmatrix} |\lambda_1|^2 + |\alpha|^2 & \alpha\overline{\lambda_2} \\ \overline{\alpha}\lambda_2 & |\lambda_2|^2 \end{bmatrix} \ \text{ and } \ T^*T = \begin{bmatrix} |\lambda_1|^2 & \alpha\overline{\lambda_1} \\ \overline{\alpha}\lambda_1 & |\lambda_2|^2 + |\alpha|^2 \end{bmatrix}$$

and $TT^* = T^*T$ if and only if $\alpha = 0$.

Exercise. Show that an upper triangular matrix T commutes with its adjoint if and only if T is diagonal.

Definition 14. *A matrix N which commutes with its adjoint, $NN^* = N^*N$, is called* **normal**.

We proved:

Theorem 15. The spectral theorem for normal matrices

A square matrix N can be diagonalized by a unitary matrix: U^*NU =diagonal *for some unitary U, if and only if N is normal: $N^*N = NN^*$.*

In particular:

Theorem 16. The spectral theorem for self-adjoint matrices
A self-adjoint matrix $A = A^*$ *can be diagonalized by a unitary matrix:* $U^*AU =$real diagonal, *for some unitary* U.

And in the more particular case:

Theorem 17. The spectral theorem for symmetric matrices
A symmetric matrix $A = A^T \in \mathcal{M}_n(\mathbb{R})$ *can be diagonalized by an orthogonal matrix:* $Q^T AQ =$real diagonal, *for some orthogonal matrix* Q.

Exercise. True or False? "A normal matrix with real eigenvalues is self-adjoint."

Note. *Unitary matrices are normal, hence are diagonalizable by a unitary.*

2.8. **Generic matrices (or: "beware of roundoff errors").** Choosing the entries of a square matrix M at random, *it is almost certain* that M has distinct eigenvalues and a complete set of eigenvectors which is not orthogonal. In other words, M is almost certainly diagonalizable, but not by a unitary conjugation, rather by conjugation with an S which requires a deformation of the space: modifications of angles, stretching of lengths.

Why is that?

I. Generic matrices have no repeated eigenvalues, hence are diagonalizable.

Indeed, of all matrices in $\mathcal{M}_n$ (a vector space of dimension n^2) the set of matrices with repeated eigenvalues form a surface of lower dimension since their entries satisfy the condition that the characteristic polynomial and its derivative have a common zero (two polynomial equations with $n^2 + 1$ unknowns).

To illustrate, consider 2 dimensional real matrices. A matrix

$$M \in \mathcal{M}_2(\mathbb{R}), \quad M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{7}$$

has its characteristic polynomial $p(\lambda) = \lambda^2 - (a + d)\,\lambda + (ad - bc)$. Then some λ is not a simple eigenvalue if and only if λ satisfies

$$p(\lambda) = 0 \quad \text{and} \quad p'(\lambda) = 0$$

Equation $p'(\lambda) = 0$ implies $\lambda = (a + d)/2$ which substituted in $p(\lambda) = 0$ gives

$$(a - d)^2 + 4\,bc = 0 \tag{8}$$

which is the condition that a two dimensional matrix has multiple eigenvalues: one equation in the four dimensional space of the entries (a, b, c, d); its solution is a three dimensional surface.

II. Among diagonalizable matrices, those diagonalizable by a unitary matrix form a set of lower dimension, due to the conditions that eigenvectors be orthogonal.

To illustrate, consider 2 dimensional matrices (7) with distinct eigenvalues. The eigenvalues are

$$\lambda_{\pm} = \frac{a+d}{2} \pm \frac{1}{2}\sqrt{(a-d)^2 + 4\,bc}$$

and the eigenvectors are

$$\mathbf{v}_{\pm} = \begin{bmatrix} b \\ \lambda_{\pm} - a \end{bmatrix} \text{ if } b \neq 0, \text{ and } \mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathbf{v}_2 = \begin{bmatrix} a-d \\ c \end{bmatrix} \text{ if } b = 0$$

which are orthogonal if and only if $b = c$, a 3-dimensional subspace in the four dimensional space of the parameters (a, b, c, d).

Why study normal and self-adjoint transformations? Problems which come from applications often have symmetries (coming from conservation laws which systems have, or are approximated to have) and normal or self-adjoint matrices often appear. We will see that in infinite dimensions, there are important linear transformations which are, or are reducible to, self-adjoint ones; differentiation is one of them.