SPECTRAL PROPERTIES OF SELF-ADJOINT MATRICES

RODICA D. COSTIN

CONTENTS

1. Review 3
1.1. The spectrum of a matrix 3
1.2. Brief overview of previous results 3
1.3. More on similar matrices 4
2. Self-adjoint matrices 4
2.1. Definitions 4
2.2. Self-adjoint matrices are diagonalizable I 5
2.3. Further properties of unitary matrices 7
2.4. One look at the proof of Theorem 11 8
2.5. Triangularization by conjugation using a unitary matrix 8
2.6. All self-adjoint matrices are diagonalizable II 10
2.7. Normal matrices 10
2.8. Generic matrices (or: "beware of roundoff errors") 11
3. The spectral projection theorem, functional calculus 13
3.1. The spectral projection theorem for self-adjoint matrices 13
3.2. The spectral projection theorem and functional calculus for normal matrices 16
3.3. Anti-self-adjoint (skew-symmetric and skew-Hermitian) matrices 17
3.4. Applications to linear differential equations 17
3.5. More analytic functional calculus 19
3.6. The spectral projection theorem and functional calculus for normal matrices 20
3.7. The resolvent matrix 20
3.8. Even more general analytic functional calculus 21
3.9. Diagonalization of unitary matrices 25
4. Quadratic forms and Positive definite matrices 26
4.1. Quadratic forms 26
4.2. Critical points of functions of several variables 29
4.3. Positive definite matrices 30
4.4. Negative definite, semidefinite and indefinite matrices 31
4.5. Applications to critical points of several variables 32
4.6. Application to differential equations: Lyapunov functions 33
4.7. Solving linear systems by minimization 36
4.8. Generalized eigenvalue problems 37
5. The Rayleigh’s principle and the minimax principle for the
eigenvalues of a self-adjoint matrix  39
5.1. The Rayleigh’s quotient  39
5.2. Extrema of the Rayleigh’s quotient  39
5.3. The minimax principle  41
5.4. The minimax principle for the generalized eigenvalue problem.  44
1. Review

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

Note that: 1. the spectrum $\sigma(T)$ contains no information on the multiplicity of each eigenvalue; 2. $\lambda \in \sigma(T)$ if and only if $T - \lambda I$ is not invertible.

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

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 $S^{-1}MS = \Lambda = \text{diagonal}$. The diagonal entries of $\Lambda$ 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 $Tx = Mx$ then $M$ is the matrix associated to $T$ in the standard basis $e_1, \ldots, e_n$ of $F^n$, while $S^{-1}MS$ is the matrix associated to $T$ in the basis $Se_1, \ldots, Se_n$ of $F^n$ (recall that $Se_j$ is the column $j$ of $S$, eigenvector of $M$).

4. Assume that the $n$-dimensional matrix $M$ is diagonalizable, and let $v_1, \ldots, v_n$ be linearly independent eigenvectors. Let $S = [v_1, \ldots, v_n]$. 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 $x \mapsto Mx$ in the basis of $F^n$ consisting of the eigenvectors $v_1, \ldots, v_n$ 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 $\lambda$ the eigenspace $V_\lambda = \{ v \mid Mv = \lambda v \}$ has dimension equal to the multiplicity of $\lambda$. 
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 $\lambda$ on the diagonal, then any power $J_p(\lambda)^k$ is an upper triangular matrix, with $\lambda^k$ 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\(^\dagger\), 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 $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. The linear transformation $T : V \to V$ is called **self-adjoint** if $T^* = T$, that is, if

   \[
   \langle Tx, y \rangle = \langle x, T^*y \rangle \quad \text{for all } x, y \in V
   \]

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

   Note that

   \[
   (M^*)^* = M, \quad (MN)^* = N^*M^*
   \]

Recall that if \( u_1, \ldots, u_n \) is an orthonormal basis of \( V \) then the inner product is the usual dot product of coordinates (also called the Euclidian inner product):

\[
\text{if } x = \sum_{k=0}^{n} x_k u_k, \ y = \sum_{k=0}^{n} y_k u_k \text{ then } \langle x, y \rangle = \sum_{k=0}^{n} \overline{x_k} y_k
\]

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

\[(1) \quad \langle x, y \rangle = \sum_{j=1}^{n} x_j y_j \text{ on } \mathbb{R}^n \]

and respectively\(^2\)

\[(2) \quad \langle x, y \rangle = \sum_{j=1}^{n} \overline{x_j} y_j \text{ on } \mathbb{C}^n \]

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 \( x \mapsto Mx \).

**Definition 4.** A matrix \( M \) is called self-adjoint if \( M = M^* \).

Note that only square matrices can be self-adjoint, and that \( M = M^* \) means, entrywise, that \( M_{ij} = \overline{M_{ji}} \) (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

\[(3) \quad \langle x, Ax \rangle \in \mathbb{R} \text{ for all } x \in \mathbb{C}^n\]

\(^2\)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:
\[ \langle x, Ax \rangle = \langle A^*x, x \rangle = \langle Ax, x \rangle = \langle x, Ax \rangle \]
hence (3). □

**Proposition 7.** If \( A \in M_n \) 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 \( v \in \mathbb{C}^n \), \( v \neq 0 \) so that \( Av = \lambda v \). Then on one hand
\[ \langle v, Av \rangle = \langle v, \lambda v \rangle = \lambda \langle v, v \rangle = \lambda \|v\|^2 \]
and on the other hand
\[ \langle v, Av \rangle = \langle A^*v, v \rangle = \langle Av, v \rangle = \langle \lambda v, v \rangle = \overline{\lambda} \langle v, v \rangle = \overline{\lambda} \|v\|^2 \]
therefore \( \lambda \|v\|^2 = \overline{\lambda} \|v\|^2 \) and since \( v \neq 0 \) then \( \lambda = \overline{\lambda} \) hence \( \lambda \in \mathbb{R} \). □

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

**Proposition 8.** If \( A \) is a real, symmetric matrix, \( A = A^T \), 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). □

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

**Proposition 9.** If \( A \in M_n \) is a self-adjoint matrix: \( A = A^* \), then eigenvectors corresponding to distinct eigenvalues are orthogonal: if \( \lambda_{1,2} \in \sigma(A) \) and \( Av_1 = \lambda_1 v_1 \), \( Av_2 = \lambda_2 v_2 \) (\( v_{1,2} \neq 0 \)) then
\[ \lambda_1 \neq \lambda_2 \implies v_1 \perp v_2 \]

**Proof:**
On one hand
\[ \langle v_1, Av_2 \rangle = \langle v_1, \lambda_2 v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle \]
and on the other hand
\[ \langle v_1, Av_2 \rangle = \langle A^*v_1, v_2 \rangle = \langle Av_1, v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle = \lambda_1 \langle v_1, v_2 \rangle \]
(since the eigenvalues are real, by Proposition 7). Therefore \( \lambda_2 \langle v_1, v_2 \rangle = \lambda_1 \langle v_1, v_2 \rangle \) and since \( \lambda_1 \neq \lambda_2 \) then \( \langle v_1, v_2 \rangle = 0 \). □