MAS 5145 Matrix Theory
Final and Comp Prep

Philip Hurst

December 12, 2017

Contents

1 Definitions

1.1 Eigenvalues and Eigenvectors

Let A be $n \times n$.
The scalar λ is said to be an eigenvalue of A if there is a non-zero vector $\underline{x}$ such that $A\underline{x} = \lambda\underline{x}$.
Such a vector $\underline{x}$ is called an eigenvector of A belonging to the eigenvalue λ.

1.2 Characteristic Polynomial

Let A be $n \times n$.
The polynomial calculuated by $\det(A - \lambda I)$ is called the characteristic polynomial.
Its roots are the eigenvalues of A.

1.3 Algebraic Multiplicity

Let A be $n \times n$ and λ be an eigenvalue of A.
The algebraic multiplicity of λ is the number of times λ appears as a root of the char. poly. of A.

1.4 Geometric Multiplicity

Let A be $n \times n$ and λ be an eigenvalue of A.
The geometric multiplicity of λ is the dimension of the eigenspace of λ.
This is the same as the number of linearly independent eigenvectors belonging to λ.

1.5 Similar and Unitarily Similar

Let A and B be $n \times n$.
B is said to be similar to A if there exists an invertible matrix S so that $S^{-1}AS = B$.
A and B are unitarily similar if P is a unitary matrix (i.e. $PP^* = P^*P = I$).

1.6 Diagonalizable Matrix

Let A be $n \times n$.
A is said to be diagonalizable if A is similar to a diagonal matrix.
There is an invertible matrix S such that $S^{-1}AS = D$ is diagonal.

1.7 Orthogonal and Unitary

A real matrix Q is an orthogonal matrix if $QQ^T = Q^TQ = I$.
A complex matrix U is a unitary matrix if $UU^* = U^*U = I$.

1.8 Minimal Polynomial

Let A be $n \times n$.
The minimal polynomial of A is the monic (i.e. coefficient of highest powered x in $p(x)$ is 1) polynomial of least degree that annihilates the matrix A.

1.9 Symmetric/Hermitian/Skew-symmetric/Skew-hermitian/Normal

$$\begin{aligned}
\text{symmetric} &\Leftrightarrow A = A^T \\
\text{hermitian} &\Leftrightarrow A = A^* \\
\text{skew-symmetric} &\Leftrightarrow A = -A^T \\
\text{skew-hermitian} &\Leftrightarrow A = -A^* \\
\text{normal} &\Leftrightarrow AA^* = A^*A \Leftrightarrow \text{exists unitary } U \text{ s.t. } UAU^{-1} \text{ is diagonal}
\end{aligned}$$

1.10 Positive Definite and Semi-Definite

Let A be hermitian.
If $\underline{x}^* A \underline{x} > 0$ for all $\underline{x} \neq 0$ then A is positive definite.
If $\underline{x}^* A \underline{x} \geq 0$ for all $\underline{x} \neq 0$ then A is positive semi-definite.

1.11 Singular Values

Let A be $n \times n$.
The singular values of A are the square roots of the eigenvalues of $A^* A$.
The singular value decomposition of A is given as $A = UDV^*$, where U, V are unitary and D is a diagonal matrix whose elements are the singular values.

1.12 Trace

The trace of a matrix is the sum of its diagonal elements and the sum of its eigenvalues.

2 Proofs of Key Results

2.1 Hermitian/Skew-hermitian matrix eigenvalues are real/purely imaginary

If A is a hermitian matrix, then every eigenvalue of A is real.

Let A be a hermitian matrix (i.e. $A = A^*$), $\underline{x}$ be an eigenvector belonging to λ, and $\underline{x} \neq 0$.

$$\begin{aligned}
\text{Then } A\underline{x} &= \lambda\underline{x} \\
\underline{x}^* A\underline{x} &= \underline{x}^*(\lambda\underline{x}) \\
(\underline{x}^* A)\underline{x} &= \lambda(\underline{x}^*\underline{x}) \\
(\underline{x}^* A)\underline{x} &= \lambda(\underline{x}^*\underline{x}) \\
(A^*\underline{x})^*\underline{x} &= \lambda(\underline{x}^*\underline{x}) \\
(A\underline{x})^*\underline{x} &= \lambda(\underline{x}^*\underline{x}) \quad \text{because } A = A^* \\
(\lambda\underline{x})^*\underline{x} &= \lambda(\underline{x}^*\underline{x}) \\
\overline{\lambda}(\underline{x}^*\underline{x}) &= \lambda(\underline{x}^*\underline{x}) \\
(\overline{\lambda} - \lambda)\underline{x}^*\underline{x} &= 0
\end{aligned}$$

We can see that $\underline{x}^*\underline{x} = ||\underline{x}||^2 > 0$ since $\underline{x} \neq 0$. Hence $(\overline{\lambda} - \lambda) = 0$.

Let $\lambda = a + ib$ and $\overline{\lambda} = a - ib$ for $a, b \in \mathbb{R}$. So $(\overline{\lambda} - \lambda) = a - ib - (a + ib) = -2ib = 0$. This implies that $b = 0$ and therefore λ is a real number.

If A is a skew-hermitian matrix, then every eigenvalue of A is purely imaginary.

Let A be a skew-hermitian matrix (i.e. $A = -A^*$), $\underline{x}$ be an eigenvector belonging to λ, and $\underline{x} \neq 0$.

$$\begin{aligned}
\text{Then } A\underline{x} &= \lambda\underline{x} \\
\underline{x}^* A\underline{x} &= \underline{x}^*(\lambda\underline{x}) \\
(\underline{x}^* A)\underline{x} &= \lambda(\underline{x}^*\underline{x}) \\
(\underline{x}^* A)\underline{x} &= \lambda(\underline{x}^*\underline{x}) \\
(A^*\underline{x})^*\underline{x} &= \lambda(\underline{x}^*\underline{x}) \\
(-A\underline{x})^*\underline{x} &= \lambda(\underline{x}^*\underline{x}) \quad \text{because } A = -A^* \\
(-\lambda\underline{x})^*\underline{x} &= \lambda(\underline{x}^*\underline{x}) \\
-\overline{\lambda}(\underline{x}^*\underline{x}) &= \lambda(\underline{x}^*\underline{x}) \\
(\overline{\lambda} + \lambda)\underline{x}^*\underline{x} &= 0
\end{aligned}$$

We can see that $\underline{x}^*\underline{x} = ||\underline{x}||^2 > 0$ since $\underline{x} \neq 0$. Hence $(\overline{\lambda} + \lambda) = 0$.

Let $\lambda = a + ib$ and $\overline{\lambda} = a - ib$ for $a, b \in \mathbb{R}$. So $(\overline{\lambda} + \lambda) = a - ib + (a + ib) = 2a = 0$. This implies that $a = 0$ and therefore λ is purely imaginary.

2.2 Unitary matrix eigenvalues have absolute value 1

Let A be a unitary matrix (so $AA^* = A^*A = I$) with eigenvalue λ.

Then for any vector $\underline{x} \neq 0$, we have

$$\begin{aligned} A\underline{x} &= \lambda\underline{x} \text{ and} \\ \underline{x}^*A^* &= \lambda^*\underline{x}^* \Rightarrow \\ \\ \underline{x}^*A^*A\underline{x} &= \lambda^*\underline{x}^*\lambda\underline{x} \Rightarrow \\ \underline{x}^*\underline{x} &= \lambda^*\lambda\underline{x}^*\underline{x} \Rightarrow \\ ||\underline{x}||^2 &= |\lambda|^2\,||\underline{x}||^2 \Rightarrow \\ |\lambda| &= 1 \quad \square \end{aligned}$$

2.3 Eigenvectors belonging to distinct eigenvalues are linearly independent

Suppose $c_1\underline{x}_1 + c_2\underline{x}_2 = 0$, where one of the coefficients (say c_1) is not zero. Then $\underline{x}_1 = \alpha\underline{x}_2$ for some $\alpha \neq 0$. Left multiplying both sides by A gives

$$A\underline{x}_1 = \lambda_1\underline{x}_1 = \alpha A\underline{x}_2 = \alpha\lambda_2\underline{x}_2$$

But multiplying $\underline{x}_1 = \alpha\underline{x}_2$ by λ_1 also gives

$$A\underline{x}_1 = \lambda_1\underline{x}_1 = \alpha\lambda_1\underline{x}_2$$

Subtracting these equations gives

$$\alpha\lambda_2\underline{x}_2 - \alpha\lambda_1\underline{x}_2 = \alpha(\lambda_2 - \lambda_1)\underline{x}_2 = \underline{0}$$

But we know that $\alpha \neq 0$ and $\lambda_2 - \lambda_1 \neq 0$ and $\underline{x}_2 \neq \underline{0}$.
Thus our assumption that the coefficients c_1 and c_2 are not zero is incorrect and $\underline{x}_1$ and $\underline{x}_2$ are linearly independent.

The can easily be extended for the $\underline{x}_n$ case. $\square$

2.4 Eigenvectors for distinct eigenvalues of a hermitian matrix are orthogonal

Let A be an $n \times n$ Hermitian matrix, λ_1, λ_2 distinct eigenvalues of A, and $\underline{x}_1, \underline{x}_2$ eigenvectors belonging to respectively λ_1, λ_2. We will show that $\underline{x}_1$ and $\underline{x}_2$ are orthogonal.

Since A is hermitian, we know these eigenvalues are real.

We can calculate $< \underline{x}_1\,,\, A\underline{x}_2 >$ in two ways:

$$< \underline{x}_1, A\underline{x}_2 >= \underline{x}_1^*A\underline{x}_2 = (\underline{x}_1^*A)\underline{x}_2 = (A^*\underline{x}_1)^*\underline{x}_2 = (A\underline{x}_1)^*\underline{x}_2 = (\lambda_1\underline{x})^*\underline{x}_2 = \overline{\lambda_1}\underline{x}_1^*\underline{x}_2 = \overline{\lambda_1} < \underline{x}_1, \underline{x}_2 >$$

$$< \underline{x}_1, A\underline{x}_2 >= \underline{x}_1^*A\underline{x}_2 = \underline{x}_1^*\,\lambda_2\,\underline{x}_2 = \lambda_2\,\underline{x}_1^*\,\underline{x}_2 = \lambda_2\; < \underline{x}_1\,,\,\underline{x}_2 >$$

We then set these two results equal.

$$\overline{\lambda_1} < \underline{x}_1\,,\,\underline{x}_2 >= \lambda_2\; < \underline{x}_1\,,\,\underline{x}_2 > \quad \Rightarrow \quad \left(\overline{\lambda_1} - \lambda_2\right) < \underline{x}_1\,,\,\underline{x}_2 >= 0$$

Since λ_1 and λ_2 are real we know that $\overline{\lambda_1} = \lambda_1 \neq \lambda_2$, so $< \underline{x}_1\,,\,\underline{x}_2 >$ must be zero. Therefore $\underline{x}_1$ and $\underline{x}_2$ are orthogonal.

This can easily be extended to the $\underline{x}_n$ case. $\square$

2.5 Trace of a matrix = Sum of its eigenvalues

For every square matrix A there is a nonsingular P such that $A = PJP^{-1}$ where J is upper tringular with its eigenvalues on the diagonals.

We know that $trace(A) = trace(PJP^{-1}) = trace(PP^{-1}J) = trace(J)$. $\square$

2.6 Gershgorin disc theorem part (a)

Let A is an $n \times n$ matrix.

Define the Gershgorin Radii as $R'_i = \sum\limits_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$.

and the Gershgorin Disc as $D_i = \{z | z \in \mathbb{C}, |z - a_{ii}| \leq R'_i\}$

and the Gershgorin Region as $G = \bigcup\limits_{i=1}^{n} D_i$.

(a) Every eigenvalue of A is in the Gershgorin disc

(b) If the union of k of the discs is disjoint from the remaining $n-k$ discs, the there are exactly k eigenvalues (counting multiplicity) in the union of the k discs.

Proof of (a): Let λ be an eigenvalue of A and $\underline{x} \neq \underline{0}$ be an eigenvector belonging to λ.

So $A\underline{x} = \lambda\underline{x}$ for $\underline{x} \in \mathbb{C}^n$. Let $|\underline{x}_p| = \max(|\underline{x}_1|, \ldots, |\underline{x}_n|)$ and we look at the p-th entries of $A\underline{x}$ and $\lambda\underline{x}$.

$$
\begin{aligned}
(A\underline{x})_p &= \lambda\underline{x}_p \Rightarrow \\
\sum_{j=1}^{n} a_{pj}x_j &= \lambda\underline{x}_p \Rightarrow \\
a_{pp}\underline{x}_p + \sum_{\substack{j=1 \\ j\neq p}}^{n} a_{pj}x_j &= \lambda\underline{x}_p \Rightarrow \\
(\lambda - a_{pp})\underline{x}_p &= \sum_{\substack{j=1 \\ j\neq p}}^{n} a_{pj}x_j \Rightarrow \\
|(\lambda - a_{pp})\underline{x}_p| &= \left|\sum_{\substack{j=1 \\ j\neq p}}^{n} a_{pj}x_j\right| \Rightarrow \\
|(\lambda - a_{pp})|\,|\underline{x}_p| &= \left|\sum_{\substack{j=1 \\ j\neq p}}^{n} a_{pj}x_j\right| \Rightarrow \\
|(\lambda - a_{pp})|\,|\underline{x}_p| &\leq \sum_{\substack{j=1 \\ j\neq p}}^{n} |a_{pj}|\,|x_j| \Rightarrow \\
|(\lambda - a_{pp})|\,|\underline{x}_p| &\leq \sum_{\substack{j=1 \\ j\neq p}}^{n} |a_{pj}|\,|x_p| \quad \text{since } |x_i| \leq |x_p|\,\forall i \\
|(\lambda - a_{pp})| &\leq \sum_{\substack{j=1 \\ j\neq p}}^{n} |a_{pj}| = R'_p \quad \text{since } |x_p| > 0
\end{aligned}
$$

$\therefore \lambda$ lies in $D_p \leq G \quad \square$

2.7 Results pertaining to positive definite matrices

2.8 Min poly divides every annihilating polynomial

Let A be $n \times n$ and let $p(x)$ annihilate A. Then the minimal polynomial, $m(x)$ is a factor of $p(x)$.

From the Euclidean algorithm we know that $p(x) = q(x)\,m(x) + r(x)$.

Substituting $x = A$ gives us

$$\begin{aligned} p(A) &= q(A)\,m(A) + r(A) \\ 0 &= q(A)\,0 + r(A) \end{aligned}$$

Therefore $r(A) = 0$ and we conclude that $m(x)$ divides $p(x)$.

2.9 Roots of the min poly are precisely the eigenvalues of the matrix

We know that $C_A(x)$ annihilates A and therefore $m_A(x)$ is a factor of $C_A(x)$.

We also know that the only roots of $C_A(x)$ are the eigenvalues of A. Therefore, every root of $m_A(x)$ is also an eigenvalue of A.

2.10 Hermitian matrix is Positive Definite if and only if all its eigenvalues are positive

($\Rightarrow$) Let A be positive definite and λ be an eigenvalue of A.

$$\begin{aligned} A\underline{x} &= \lambda \underline{x} \Rightarrow \\ \underline{x}^* A \underline{x} &= \lambda \underline{x}^* \underline{x} = \lambda ||\underline{x}||^2 \Rightarrow \\ \lambda &= \frac{\underline{x}^* A \underline{x}}{||\underline{x}||^2} > 0 \quad \square \end{aligned}$$

($\Leftarrow$) Let A be hermitian and let all its eigenvalues be positive.

Since A is hermitian there exists unitary U such that $U^* A U = D$ where D is a diagonal matrix composed of the eigenvalues of A.

We want to show that $\underline{x}^* A \underline{x} = \underline{x}^* U D U^* \underline{x}$ for all $\underline{x} \neq \underline{0}$.

Let $U^* \underline{x} = \underline{y} = [y_1 \dots y_n]^T$. Since $\underline{x} \neq \underline{0}$, we know $U^* \underline{x} \neq \underline{0}$.

So $\underline{x}^* U D U^* \underline{x} = \underline{y}^* D \underline{y} = [\overline{y}_1 \dots \overline{y}_n] \begin{bmatrix} \lambda_1 y_1 \\ \vdots \\ \lambda_n y_n \end{bmatrix} = \lambda_1 |y_1|^2 + \dots + \lambda_n |y_n|^2 > 0$ because $\lambda_i > 0$ and at least one $y_i \neq 0$.

Thus $\underline{x}^* A \underline{x} > 0$ and A is positive definite. $\square$

2.11 Trace(A^*A) = sum of square of moduli of eigenvalues of A when A is normal

Since A is normal, there exists unitary diagonalizable matrix U such that $U^* A U = D$ is diagonal with the eigenvalues of A.

We can write $A^* A = U A U^* \, U A^* U^* = U D^* D U^*$.

So $trace(A^* A) = trace(U D^* D U^*) = trace(U U^* D^* D) = trace(I D^* D) = trace(D^* D)$.

$D^* D = diag(\lambda_1 \overline{\lambda_1}, \dots, \lambda_n \overline{\lambda_n})$ which means the trace is the sum of square modulii of the eigenvalues of A.

3 Statements of Theorems

3.1 Schur's Upper Triangularization Theorem

Let A be $n \times n$ over $\mathbb{F}$. Then there exists a unitary matrix U such that $U^*AU = T$ is upper triangular.

The eigenvalues of A are the diagonal entries of T.

If A is real and all eigenvalues of A are real, then U can be chosen to be real orthogonal.

3.2 Cayley-Hamilton Theorem

The characteristic polynomial of an $n \times n$ matrix A annihilates A.

3.3 Gershgorin Disc Theorem (include radii, discs, and region)

Let A is an $n \times n$ matrix.

Define the Gershgorin Radii as $R_i' = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$.

and the Gershgorin Disc as $D_i = \{z | z \in \mathbb{C}, |z - a_{ii}| \leq R_i'\}$

and the Gershgorin Region as $G = \bigcup_{i=1}^{n} D_i$.

(a) Every eigenvalue of A is in the Gershgorin disc

(b) If the union of k of the discs is disjoint from the remaining $n-k$ discs, the there are exactly k eigenvalues (counting multiplicity) in the union of the k discs.

3.4 Necessary and sufficient conditions for the diagonalizability of a matrix

3.5 Significant of a matrix being normal (A is normal iff A is unitarily diagonalizable)

4 Problems - LOOK AT CLASS NOTES TOO

4.1 Proof Type Problems

Similar to those on problem sets, homework assignments, class tests, and those done in class.

4.2 Computational Problems

1. Finding eivenvalues/bases for eigenspaces
2. Illustrating Schur's Theorem
3. Applying the Cayley-Hamilton Theorem
4. Eigenvalues of polynomials of a matrix
5. Computing singular values and related computations
6. Jordan Canonical Form and related problems (e.g. J form ofthe power of a J-block)
7. Gershgorin discs
8. Computing singular values and related matters
9. Eigenvalues of a polynomials of a matrix