Lecture Notes for Math 623
Matrix Analysis

Michael E. O'Sullivan
mosulliv@math.sdsu.edu
www-rohan.sdsu.edu/ ˜ mosulliv

April 18, 2013

1 Normal Matrices

All matrices are in $\mathcal{M}_n(\mathbb{C})$. A^* is the conjugate transpose.

We recall several definitions and a new term "normal."

Definition 1.1. $A \in \mathcal{M}_n(\mathbb{C})$ is

- *normal* when $AA^* = A^*A$.
- *Hermitian* when $A^* = A$.
- *skew- Hermitian* when $A^* = -A$.
- *unitary* when $A^* = A^{-1}$.

A real unitary matrix is called *orthogonal*. A real hermitian matrix is called *symmetric*.

Exercises 1.2.

(a) Hermitian, skew-hermitian, unitary and diagonal matrices are all normal.

(b) Check that the following two matrices are normal, but they are not unitary, nor Hermitian nor skew-Hermitian.

$$\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix} \qquad \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

(c) Show that a normal 2×2 real matrix is either hermitian, skew-hermitian or a constant multiple of a rotation matrix (i.e).

(d) The set of normal matrices is closed under scalar multiplication, but is not closed under addition and is not closed under multiplication.

(e) The set of normal matrices is closed under unitary conjugation.

(f) $A \in \mathcal{M}_m\mathbb{C}$ and $B \in \mathcal{M}_n\mathbb{C}$ are normal iff $A \oplus B$ is.

(g) Suppose A and B are real normal matrices satisfying $AB^* = B^*A$. Show that $A + B$ and AB are both normal.

Lemma 1.3. *Suppose that T is upper (or lower) triangular and normal. Then T is a diagonal matrix.*

Proof. We proceed by induction, the case $n = 1$ is immediate. Let T be $n \times n$ normal and upper triangular.

$$T = \begin{bmatrix} t_{1,1} & t_{1,2} & t_{1,3} & \dots & t_{1,n-1} & t_{1,n} \\ 0 & t_{2,2} & t_{2,3} & \dots & t_{2,n-1} & t_{2,n} \\ 0 & 0 & t_{3,3} & \dots & t_{3,n-1} & t_{3,n} \\ \dots & \dots & \dots & \dots & & \\ 0 & 0 & 0 & \dots & 0 & t_{n,n} \end{bmatrix} \qquad T^* = \begin{bmatrix} \overline{t_{1,1}} & 0 & 0 & \dots & 0 & 0 \\ \overline{t_{1,2}} & \overline{t_{2,2}} & 0 & \dots & 0 & 0 \\ \overline{t_{1,3}} & \overline{t_{2,3}} & \overline{t_{3,3}} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & & \\ \overline{t_{1,n}} & \overline{t_{2,n}} & \overline{t_{3,n}} & \dots & \overline{t_{n-1,n}} & \overline{t_{n,n}} \end{bmatrix}$$

Then $(T^*T)_{1,1} = \overline{t_{1,1}}t_{1,1}$, which is real and nonnegative. On the other hand,

$$(TT^*)_{1,1} = t_{1,1}\overline{t_{1,1}} + t_{1,2}\overline{t_{1,2}} + \cdots + t_{1,n}\overline{t_{1,n}} = \sum_{i=1}^{n} |t_{1,i}|^2$$

This is a sum of nonnegative reals. Equating these two expressions we get $|t_{1,1}|^2 = \sum_{i=1}^{n} |t_{1,i}|^2$, so $|t_{1,i}| = 0$ for $i > 1$. Thus T has the form

$$T = \begin{bmatrix} t_{1,1} & 0 \\ 0 & T' \end{bmatrix}$$

Where T' is upper triangular and normal (by an exercise, $A \oplus B$ is normal implies A and B are normal). Applying the induction hypothesis, T' is diagonal, which gives the result. □

Exercises 1.4.

(a) Show that for any matrix A, $\mathrm{Tr}(A^*A) = \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{i,j}|^2$. We will write this more concisely as $\sum_{i,j=1}^{n} |a_{i,j}|^2$.

Lemma 1.5. *If A, B are unitarily equivalent then $\sum_{i,j=1}^{n} |a_{i,j}|^2 = \sum_{i,j=1}^{n} |b_{i,j}|^2$.*

Proof. Let $A = U^*BU$.

$$\sum_{i,j=1}^{n} |a_{i,j}|^2 = \operatorname{Tr}(A^*A) = \operatorname{Tr}(U^*B^*UU^*BU)$$

$$= \operatorname{Tr}(U^*B^*BU) = \operatorname{Tr}(B^*B) = \sum_{i,j=1}^{n} |b_{i,j}|^2$$

Note that $U^* = U^{-1}$ so that B^*B is similar to U^*B^*BU and therefore these matrices have the same trace. □

Theorem 1.6 (Spectral)**.** *The following are equivalent for $A \in \mathcal{M}_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$.*

(1) A is unitarily diagonalizable.

(2) A is normal.

(3) $\sum_{i,j=1}^{n} |a_{i,j}|^2 = \sum_{i=1}^{n} |\lambda_i|^2$.

Suppose A is unitarily diagonalizable, $A = U^*\Lambda U$. Then

$$A^*A = U^*\Lambda^*UU^*\Lambda U = U^*\Lambda^*\Lambda U = U^*\Lambda\Lambda^*U = AA^*$$

Thus A is normal. This computation also showed A^*A is similar to $\Lambda^*\Lambda$ and therefore

$$\sum_{i,j=1}^{n} |a_{i,j}|^2 = \operatorname{Tr}(A^*A) = \operatorname{Tr}\Lambda^*\Lambda = \sum_{i=1}^{n} |\lambda_i|^2$$

Now we show that either (2) or (3) imply that A is unitarily diagonalizable. Any matrix is unitarily triangularizable, so let $A = U^*TU$ with U unitary and T upper triangular.

Suppose A is normal. Since T is similar to A it is also normal. By the Lemma 1.3, it is diagonal. Thus A is unitarily diagonalizable.

Now suppose $\sum_{i,j=1}^{n} |a_{i,j}|^2 = \sum_{i=1}^{n} |\lambda_i|^2$. By the Lemma 1.5

$$\sum_{i=1}^{n} |\lambda_i|^2 = \sum_{i,j=1}^{n} |a_{i,j}|^2$$

$$= \sum_{i,j=1}^{n} |t_{i,j}|^2$$

and since the diagonal entries of T are the eigenvalues of A

$$= \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i=1}^{n} \sum_{j=i+1}^{n} |t_{i,j}|^2$$

From the first and last lines and the nonnegativity of $|t_{i,j}|^2$ we can conclude that $t_{i,j} = 0$. So T is diagonal.

Proof. □

The following results are immediate, but worth explicitly recognizing. Some are properties that we showed for Hermitian matrices, so we see they also apply to normal matrices.

Corollary 1.7. *Let A be normal.*

- *A normal matrix is nondefective. For each eigenvalue, algebraic multiplicity = geometric multiplicity. that is, the eigenspace equals the whole generalized eigenspace.*
- *The minimal polynomial of A equals the characteristic polynomial of A.*
- *For any eigenvalue λ, the eigenspace of λ is equal to the generalized eigenspace of λ.*
- *If A is normal and x is a right eigenvector associated to λ then x^* is a left eigenvector for λ^**
- *If x and y are eigenvectors associated to distinct eigenvalues then x and y are orthogonal.*

We can now say more about unitary matrices.

Theorem 1.8. *Let $A \in \mathcal{M}_n$. The following are equivalent.*

(1) A is unitary.

(2) A is normal and has eigenvalues of modulus 1.

(3) $\|Av\| = \|v\|$ for all $v \in \mathbb{C}^n$.

Proof. We have already shown the equivalence of (1) and (3). If A is unitary it is normal, and is therefore unitarily diagonalizable, say $U^*AU = \Lambda$. Then

$$\Lambda^*\Lambda = U^*A^*UU^*AU = U^*A^*AU = U^*U = I$$

So the eigenvalues have modulus 1. Conversely, A is normal and the eigenvalues have modulus 1, then it is unitarily diagonalizable $U * AU = \Lambda$. Since Λ is unitary, and A is unitarily similar to it A is also unitary. Explicity,

$$A^*A = U^*\Lambda^*UU^*\Lambda U = U^*\Lambda^*\Lambda U = U^*U = I$$

□

We can also say more about Hermitian matrices.

Theorem 1.9. *Let $A \in \mathcal{M}_n$. The following are equivalent.*

(1) A is Hermitian.

(2) A is normal and has real eigenvalues.

(3) v^*Av is real for all $v \in \mathbb{C}^n$.

Proof. We have already proven that A Hermitian implies (2) and (3). If A is normal and has real eigenvalues then it is unitarily diagonalizable, $A = U\Lambda U^*$ with $\Lambda \in \mathcal{M}_n(\mathbb{R})$. Then A is unitarily similar to a Hermitian matrix, so it is also Hermitian.

Suppose that v^*Av is real for all $v \in \mathbb{C}^n$. Then $(u+v)^A(u+v) - u^*Au - v^*Av$ is real for any $u, v \in \mathbb{C}^n$. This is equal to $u^*Av + v^*Au$. Set $u = e_i$ and $v = e_j$ to obtain $a_{i,j} + a_{j,i}$ is real. This shows $\operatorname{Im}(a_{i,j}) = -\operatorname{Im}(a_{j,i})$. Set $u = ie_i$ and $v = e_j$ to obtain $-ia_{i,j} + ia_{j,i}$ is real. This shows $\operatorname{Re}(a_{i,j}) = \operatorname{Re}(a_{j,i})$. Thus $a_{j,i} = \overline{a_{i,j}}$ and A is Hermitian. □

For $z \in \mathbb{C}$, writing $z > 0$ means that z is real and positive (and similarly $z \geq 0$ means nonnegative).

Corollary 1.10. *Let A be Hermitian.*

(1) $v^*Av \geq 0$ for all $v \in \mathbb{C}^n$, iff the eigenvalues of A are nonnegative.

(2) $v^*Av > 0$ for all nonzero $v \in \mathbb{C}^n$, iff the eigenvalues of A are positive.

Proof. If $v^*Av \geq 0$ for all $v \in \mathbb{C}^n$, then in particular for an eigenpair λ, u we have $0 \leq u^*Au = \lambda u^*u$. Since u^*u is positive for nonzero u, we must have $\lambda \geq 0$. Furthermore if the inequality is strict, as assumed in (2), then $\lambda > 0$.

If A is Hermitian, then by the spectral theorem it is unitarily diagonalizable, so $A = U * \Lambda U$. Suppose the eigenvalues of A are all nonnegative. Then

$$v^*Av = v^*U^*\Lambda Uv = (Uv)^*\Lambda(Uv) = \sum_{i=1}^{n} \lambda_i \overline{x}_i \overline{x}$$

where $x = Uv$. Since all the λ_i are nonnegative, and so is $\overline{x}_i x_i$, we must have $v^*Av \geq 0$.

If all the λ_i are actually positive, and v is a nonzero vector, then $x = Uv$ will be nonzero as well. Thus $\sum_{i=1}^{n} \lambda_i \overline{x}_i \overline{x}$ will have a strictly positive term and $v^*Av > 0$ for all nonzero x. □

2 Real normal matrices

I will leave aside the proof of this, but wanted you to see the result.

Theorem 2.1. *For $A \in \mathcal{M}_n(\mathbb{R})$, TFAE*

(1) A is symmetric.

(2) There exists an orthogonal matrix $Q \in \mathcal{M}_n(\mathbb{R})$ and a real diagonal matrix D such that $Q^TAQ = D$

(3) A is normal and all eigenvalues of A are real.

(4) There exists an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A.

3 Unitary Similarity

Definition 3.1. We say A is *unitarily similar* to B when there exists a unitary matrix U such that $A = U^*BU$. We write $A \sim_U B$.

Exercises 3.2.

(a) Unitary similarity is an equivalence relation.

(b) If $A \sim_U B$ then $A \sim B$, but matrices can be similar without being unitarily similar. (Find an example).

(c) If $A \sim_U B$ then $A^* \sim_U B^*$ and $AA^* \sim_U BB^*$.

We want to generalize the last result.

Definition 3.3. A *word* , $W(x,y)$, is an arbitrarily long product of xs and ys. For example $W(x,y) = xyxxxyyx$ is a word of degree 8. Evaluating this word $W(x,y)$ at A, A^* gives the matrix $AA^*AAAA^*A^*A$.

Proposition 3.4. *If $A \sim_U B$ then, for any word $W(x,y)$, $W(A,A^*) \sim_U W(B,B^*)$. Furthermore,* $\mathrm{Tr}(W(A,A^*)) = \mathrm{Tr}((W(B,B^*))$.

Proof. Let $A = U^*BU$ and notice $A^* = U^*B^*U$. We can prove formally that $W(A,A^*) = U^*W(B,B^*)U$ for any word $W(x,y)$ via induction on the degree of the word. The trace formula follows from similar matrices having equal traces.

The claim is obvious for degree 1. Suppose it is true for words of degree d, let $W(x,y)$ have degree $d+1$ and suppose the last letter of $W(x,y)$ is x (the proof for y is analogous). Then $W(x,y) = W'(x,y)x$ with $W'(x,y)$ having length d. By the induction hypothesis, $W'(A,A^*) = U^*W'(B,B^*)U$. So,

$$W(A,A^*) = W'(A,A^*)A = U^*W'(B,B^*)UU^*BU = U^*W(B,B^*B)U = W(B,B^*)$$

$\square$

Theorem 3.5 (Sprecht)**.** *If* $\mathrm{Tr}(W(A,A^*) = \mathrm{Tr}(W(B,B^*)$ *for all words $W(x,y)$, then $A \sim_U B$.*

[Peary] If $\mathrm{Tr}(W(A,A^*) = \mathrm{Tr}(W(B,B^*)$ *for all words $W(x,y)$ of degree at most $2n^2$, then $A \sim_U B$.*

Horn does not prove Sprecht's (nore Peary's) result. I don't know if there is a canonical for unitary similarity classes, analogous to Jordan form for similarity classes.

4 Miscellaneous

At the end of Section 2.1 Horn brings up a generalization of unitary matrices: instead of requiring that $U^{-1} = U^*$, consider A *similar* to A^*. I'm not sure if there is some reason for interest in this idea, or it is just a curiosity. I found Theorem 2.1.9 a bit hard to follow, so I broke it into pieces.

Lemma 4.1. *Let $A \in \mathcal{M}_n$. Let α be such that $-\alpha/\overline{\alpha}$ is not an eigenvalue of A. Then $(\alpha + \overline{\alpha}A)$ is invertible.*

Proof. Let x be a nonzero vector in $\mathbb{C}^n$.

$$\begin{aligned}(\alpha + \overline{\alpha}A)x = 0 &\Longleftrightarrow Ax = -\frac{\alpha}{\overline{\alpha}}x \\ &\Longleftrightarrow -\frac{\alpha}{\overline{\alpha}} \text{ is an eigenvalue of } A\end{aligned}$$

Choosing α as in the statement of the theorem guarantees $\alpha I + \overline{\alpha}A$ is trivial nullspace and is therefore invertible. $\square$

Corollary 4.2. *Let S be invertible. There is a $\theta \in [0, 2\pi)$ such that $(S_\theta + S_\theta)*$ is invertible. Here $S_\theta = e^{i\theta}S$.*

Proof. Apply the previous Lemma, let $\alpha = e^{i\theta}$ be such that $-\alpha/\overline{\alpha} = -e^{-2i\theta}$ is *not* be an eigenvalue of $S^{-1}S^*$. Then $e^{i\theta}I + e^{-i\theta}S^{-1}S^*$ is invertible. Multiply by S to get $S_\theta + S_\theta^*$ is invertible. $\square$

Theorem 4.3. *Let A be invertible. $A^{-1} \sim A^*$ if and only if $A = B^{-1}B^*$ for some invertible matrix B.*

Proof. Suppose $A = B^{-1}B^*$. Then

$$B^{-1}(B^{-1}B^*)B = B^{-1}(BB^{-1})^*B = (B^{-1})^*B = (B^{-1}B^*)^{-1}$$

So $B^-1A^*B = A^{-1}$, showing A^* is similar to A^{-1}.

Conversely, suppose that A^* is similar to A^{-1}. We will find an invertible Hermitian matrix H such that $H = A^*HA$. Then choose α such that $\alpha I + \overline{\alpha}A^*$ is invertible. Set $B = (\alpha I + \overline{\alpha}A^*)H$. Then B is invertible since it is the product of two invertible matrices.

$$\begin{aligned}BA &= (\alpha I + \overline{\alpha}H)A = \alpha HA + \overline{\alpha}H \\ &= H(\alpha A + \overline{\alpha}I) = B^*\end{aligned}$$

Therefore $A = B^{-1}B^*$.

To get H we know $A^* = SA^{-1}S^{-1}$ for some S. Choose θ so that $S_\theta + S_\theta^*$ is invertible. Then $A^* = S_\theta A^{-1}S_\theta^{-1}$ so

$$\begin{aligned}S_\theta &= A^*S_\theta A \\ S_\theta^* &= A^*S_\theta^*A\end{aligned}$$

Adding these two equations

$$S_\theta + S_\theta^* = A^*(S_\theta + S_\theta^*)A$$

Thus $H = S_\theta + S_\theta^*$ is Hermitian and invertible and satisfies $H = A^*HA$. $\square$

5 Positive Definite Matrices

Definition 5.1. A Hermitian matrix $A \in \mathcal{M}_n$ is *positive definite* when $v^*Av > 0$ for all nonzero $v \in \mathbb{C}^n$. It is *positive semi-definite* when $v^*Av \geq 0$ for all nonzero $v \in \mathbb{C}^n$.

Here we follow the convention that $z > 0$ for $z \in \mathbb{C}$ means that z is real and positive.

Exercises 5.2.

- **(a)** Any principal sub matrix of a positive definite matrix is positive definite (and similarly for positive semi-definite).
- **(b)** The trace and determinant of all principal minors are positive.
- **(c)** For any $i \neq j$, $a_{ii}a_{jj} > |a_{ij}|^2$.
- **(d)** The set of positive definite matrices is closed under addition and scalar multiplication by *positive* real numbers. It doesn't form a vector space, but is a union of rays from the origin (the 0 matrix) and is called a *cone*.
- **(e)** Applying our results on normal matrices, a positive definite matrix A has positive eigenvalues and is unitarily diagonalizable, while a semi-definite matrix may have 0 eigenvalues.
- **(f)** The definition could be rephrased to say A is normal and $v^*Av > 0$, since these imply that A is Hermitian.
- **(g)** Let $A \in \mathcal{M}_n$ be positive definite. For any $C \in \mathcal{M}_{n,m}$, C^*AC is positive definite and $\mathrm{rk}(C^*AC) = \mathrm{rk}(C)$.

For a nonsquare matrix $D \in \mathcal{M}_{n,m}$ I will say that D is diagonal when for $i \neq j$, $d_{i,j} = 0$.

Theorem 5.3 (Singular Value Decomposition[H2)**.** *2.6.3] Let $A \in \mathcal{M}_{n,m}$ with $m \geq n$ and $r = \mathrm{rk}(A)$. There is a unique diagonal matrix $\Sigma \in \mathcal{M}_{n,m}$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$, and there are unitary matrices $U \in \mathcal{M}_n$ and $W \in \mathcal{M}_m$ such that*

$$A = U\Sigma W$$

The parameters $\sigma_1, \ldots, \sigma_r$ are the nonzero eigenvalues of A^*A, arranged in decreasing order. They are also the nonzero eigenvalues of AA^*.

The σ_i are called the singular values of A. Exactly $r = \mathrm{rk}(A)$ of them are nonzero. They are uniquely defined, since they are the nonzero eigenvalues of AA^*. It is easy to check that AA^* is positive semi-definite.