Complex Numbers, Normal Matrices, Positive Definite Matrices (cont.), Matrix Ellipses

Today is mainly a review session, though there might still be some new perspectives on matrices—you may find them helpful later after the course when you actually use linear algebra.

1. **Extending some concepts with complex numbers**

   (a) **Euler's Identity**
   The "Eu" pronounces /o/.

   $$e^{i\theta} = \cos\theta + i\sin\theta$$

   where θ denotes a real number. This formula says that the real and imaginary parts of a complex exponential function $f(\theta) = e^{i\theta}$ are trigonometric functions of θ. This formula comes from the Maclaurin series representation for the exponential function

   $$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

   Replace z with $i\theta$. Every even power of i will be a real number (± 1) and every odd power of i will be imaginary ($\pm i$). So split the series into two pieces:

   $$e^{i\theta} = \sum_{k=0}^{\infty} \frac{(i\theta)^k}{k!} = (1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \ldots) + i(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \ldots) \tag{1}$$

   $$= \cos\theta + i\sin\theta \tag{2}$$

   (b) **Hermitian Matrix** (symmetric matrix in the real case)

   > When you transpose a complex vector or a matrix, take the complex conjugate too.

   This action of transposing and also "conjugating" on a matrix A can be denoted as A^H or A^*. It reads "A Hermitian".

   $$\text{If } A = \begin{bmatrix} 1 & i \\ 0 & 1+i \end{bmatrix}, \text{then } A^H = \begin{bmatrix} 1 & 0 \\ -i & 1-i \end{bmatrix}$$

   The Hermitian matrix is such that $A = A^H$. You will see that it has orthogonal eigenvectors and real eigenvalues, just like the real symmetric matrix. Although, in the Hermitian case, the eigenvectors are complex. If they are normalized in length, the eigenvector matrix is not Q, it is U, which comes next.

   (c) **Unitary Matrix** (orthogonal matrix in the real case)
   > A unitary matrix is a (complex) square matrix that has orthonormal columns.

   Similar to $Q^T Q = I$, here we have $U^H U = I$.
   Also similar to $Q^{-1} = Q^T$, we have $U^{-1} = U^H$.
   Finally, both Q and U have $|\lambda| = 1$.

2. **When are the eigenvectors orthogonal?**
   > iif $AA^H = A^H A$, A has an orthogonal basis of eigenvectors.

   The matrices that satisfied the condition $AA^H = A^H A$ are called **normal** matrices—you might

have seen it while reading the textbook by Strang.

I have attached the proof [1] of this theorem—fairly straightforward. Please take a look at it if you are curious! You may read and skip the proof of the Schur decomposition when you first read it. So, given the following classes of matrices we have learned so far, which of these have orthogonal eigenvectors?

(a) Real symmetric

(b) Hermitian

(c) Orthonormal (meaning that it is not necessarily square)

(d) Orthogonal

(e) Unitary

(f) Skew symmetric

(g) Skew Hermitian ($A^* = -A$)

What about their eigenvalues? Real or purely imaginary or just complex in general?

3. **Ellipses of $A\boldsymbol{x}$ when $\boldsymbol{x}$ moves on the unit circle**
Before we see $A\boldsymbol{x}$ with the constraint on $\boldsymbol{x}$, let's see the ellipses behind $\boldsymbol{x}^T A\boldsymbol{x} = 1$.

(a) **The geometry (change of basis) behind $\boldsymbol{x}^T A\boldsymbol{x} = 1$** (Strang 6.5)
Suppose $A = Q\Lambda Q^T$ is positive definite, so $\lambda_i > 0$. The graph of $\boldsymbol{x}^T A\boldsymbol{x} = 1$ is an ellipse:

$$\begin{bmatrix} x & y \end{bmatrix} Q\Lambda Q^T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} X & Y \end{bmatrix} \Lambda \begin{bmatrix} X \\ Y \end{bmatrix} = \lambda_1 X^2 + \lambda_2 Y^2 = 1 \quad [\text{let } \begin{bmatrix} X \\ Y \end{bmatrix} = Q^T \begin{bmatrix} x \\ y \end{bmatrix}]$$

The axes point along eigenvectors. The half-lengths are $1/\sqrt{\lambda_1}$ and $1/\sqrt{\lambda_2}$.

(b) **Ellipses of $A\boldsymbol{x}$ when $\boldsymbol{x}$ moves on the unit circle** (Strang 7.4-14)
For vectors in the unit circle $\|\boldsymbol{x}\| = 1$ and for invertible A, the vectors $\boldsymbol{y} = A\boldsymbol{x}$ in the ellipse will have $\|A^{-1}\boldsymbol{y}\| = 1$. This then gives $\boldsymbol{y}^T (A^{-1})^T A^{-1} \boldsymbol{y} = 1 \Rightarrow \boldsymbol{y}^T (AA^T)^{-1} \boldsymbol{y} = 1$. This tells you that this ellipse has axes along the singular vectors with lengths equal to $\sigma_1, ..., \sigma_r$. For singular A, can you use the pseudoinverse? Why not? For a 2×2 case, type `eigshow` and enter the `return` key into MATLAB if you have downloaded the function. It shows $A\boldsymbol{x}$ is a line. You know why it is a line.

4. **Practice of proofs—positive (semi-)definite matrices**

(a) **The Laplacian matrix**
The main tools for spectral clustering are graph Laplacian matrices. Before we introduce Laplacian matrices as an exercise to prove positive semi-definiteness, I want to motivate you by showing you a result of clustering in the following figure.
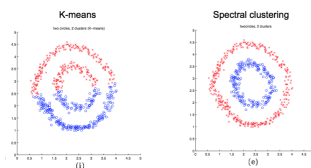
In developing spectral clustering, one key element is to show that the Laplacian matrix is positive semi-definite. This, surprisingly, can let you derive the number of clusters from

---

[1]http://www.cs.uleth.ca/ holzmann/notes/eigen.pdf

the Laplacian matrices. We would not go that far, but we want to at least prove the positive semi-definiteness. The point is that I want you to see "positive semi-definite" in real applications. Now here is the exercise.

Problem: A graph of n vertices, and their pair-wise similarities are put into **the weight matrix** $W^{n \times n}$. We also assume that the graph is undirected: $w_{ij} = w_{ji}$. We also define degree as $d_i = \sum_{j=1}^{n} w_{ij}$. We put the degrees $d_1, \ldots, d_n$ on the diagonal of the matrix $D^{n \times n}$—**the degree matrix**. The (unnormalized) Laplacian matrix L is defined as

$$L = D - W$$

Let's show that L is a positive semi-definite matrix.

Solution: First, check if this matrix is symmetric! It is, then there are 4 ways to show that a $n \times n$ symmetric matrix is positive (semi-) definite: n pivots, n determinants, n eigenvalues, $\boldsymbol{x}^T A \boldsymbol{x} > 0$ for $\boldsymbol{x} \neq \boldsymbol{0}$. For positive definite, you can also find R with independent columns such that $R^T R = A$. We will use the 4th method.

For any $\boldsymbol{x}$, $\boldsymbol{x} \neq \boldsymbol{0}$:

$$\boldsymbol{x}^T L \boldsymbol{x} = \boldsymbol{x}^T D \boldsymbol{x} - \boldsymbol{x}^T W \boldsymbol{x} \tag{3}$$

$$= \sum_{i=1}^{n} d_i x_i^2 - \boldsymbol{x}^T \sum_{j=1}^{n} w_{ij} x_j \tag{4}$$

$$= \sum_{i=1}^{n} d_i x_i^2 - \sum_{i=1}^{n} x_i \sum_{j=1}^{n} w_{ij} x_j \tag{5}$$

$$= \sum_{i=1}^{n} d_i x_i^2 - \sum_{i,j=1}^{n} x_i x_j w_{ij} \tag{6}$$

$$= \frac{1}{2} \left(\sum_{i=1}^{n} d_i x_i^2 - 2 \sum_{i,j=1}^{n} x_i x_j w_{ij} + \sum_{j=1}^{n} d_j x_j^2 \right) \tag{7}$$

$$= \frac{1}{2} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} x_i^2 - 2 \sum_{i,j=1}^{n} x_i x_j w_{ij} + \sum_{j=1}^{n} \sum_{i=1}^{n} w_{ij} x_j^2 \right) \tag{8}$$

$$= \frac{1}{2} \sum_{i,j=1}^{n} w_{ij} (x_i - x_j)^2 \geq 0 \tag{9}$$

This material is by no means required. I included this as an exercise because it gives you practice on writing matrix multiplying a vector as sums, and it reviews the concept of positive definite matrices. You also see that the seemingly unnecessary concepts such as positive definite matrices can be theoretical building blocks in real applications. If you are interested in Spectral Clustering, one visual reference is here, while a mathematical reference is here.

The next exercise was inspired by a homework exercise. I came upon the seemingly unnecessary inequality while studying information theory. I wrote it into an exercise to guide you through the proof.

(b) **Prove Hadamard inequality:** [2]

Hadamard's inequality states that if N is a matrix having columns v_i, then

$$|\det(N)| \leq \prod_{i=1}^{n} \|v_i\|. \tag{10}$$

This is the general form of Hadamard's inequality.

  i. Show that the inequality is trivial for the singular case.

  ii. **Alternative form of Hadamard inequality for positive definite matrices**
Show that in the case of a positive definite case, the Hadamard inequality can be written in the following form, which states that for any positive definite matrix K, its determinant is less than the product of its diagonal elements,

$$\det(K) \leq \prod_{i} K_{ii} \tag{11}$$

Hint: K can be written as $N^T N$ where N's columns are independent.

To prove the inequality, you would need some algebraic manipulations, which is not complicated but not important for a linear algebra review exercise. Instead of proving the general inequality,

  iii. Derive the condition that satisfies the equality for the form for positive definite matrix.

  iv. Derive the condition that satisfies the equality in its general form.

Thanks for the whole semester! I have been trying to introduce intuitions through physics meanings alongside the theoretical concepts in the course—right from our first tutorial. I hope you had fun and will recall bits of the maths and intuitions when you use linear algebra in your studies later. I enjoyed grading your homework and saw some of your sparkling ideas. Good luck on finals. Matthew is holding a review session: 8th Jan (Tuesday) 15:00-16:40, Teaching Center 401.

I will send out reviews on Quiz3 and the mid-term later. Be sure to take a look. I will also get back to you on Homework 9 and 10 via email.

---

[2]https://en.wikipedia.org/wiki/Hadamard%27s_inequality