Discrete Math 37110 - Class 19 (2016-12-01)

Instructor: László Babai
Notes taken by Jacob Burroughs
Partially revised by instructor

Review: Tuesday, December 6, 3:30-5:20 pm; Ry-276
Final: Thursday, December 8, 10:30-12:30; Ry-251

DO 19.1. IMPORTANT. Study the relevant chapters of LN (about Markov Chains) and the Linear Algebra online text.

19.1 Similar matrices

Definition 19.2. Let $A, B \in M_n(\mathbb{F})$ where $\mathbb{F}$ is a field (i.e. $\mathbb{R}, \mathbb{C}$)
A, B are similar if $\exists C \in M_n(\mathbb{F})$, $\exists C^{-1}$ such that $B = C^{-1}AC$. Notation $A \sim B$.

DO 19.3. Similarity is an equivalence relation on $M_n(\mathbb{F})$

DO 19.4. If $A \sim B$ then $\text{trace}(A) = \text{trace}(B)$. (Hint $\text{trace}(CD) = \text{trace}(DC)$)

DO 19.5. If $A \sim B$ then $\det(A) = \det(B)$. (Hint $\det(CD) = \det(C)\det(D)$)

DO 19.6. If $A \sim B$ then $f_A = f_B$ (Their characteristic polynomials are equal)

19.2 Matrix of a linear map

Definition 19.7. A *linear transformation* is a function $f : V \to V$ where $f(a + b) = f(a) + f(b)$ and $f(\lambda a) = \lambda f(a)$. Equivalently, $f(\sum \alpha_i a_i) = \sum \alpha_i f(a_i)$

Definition 19.8. A *linear map* is a function $f : V \to W$ with the same attributes.

DO 19.9. If $v_1, \dots, v_n$ are a basis of V, $w_1, \dots, w_n$ are arbitrary vectors in W, there exists a unique linear map such that $(\forall i)(f(v_i) = w_i)$

Definition 19.10. Coordinates: Let $\underline{e} = (e_1, \dots, e_n)$ be a basis of V. Then every $v \in V$ can be uniquely written as $v = \sum \alpha_i e_i$. THe α_i are the *coordinates* of v wrt (with respect to) the basis $\underline{e}$.Arranged in a column vector, we write $[v]_{\underline{e}} = (alpha_1, \dots, \alpha_n)^T$ (transpose, to make it a column vector).

19.3 Change of basis

Definition 19.11 (Change of basis matrix)**.** Take two bases: $\underline{e} = (e_1, \dots, e_n)$ (the "old" basis) and $\underline{e'} = (e'_1, \dots, e'_n)$. The change of basis matrix is $S = [[e'_1]_{\underline{e}}, \dots, [e'_n]_{\underline{e}}]$. (The i-th column lists the coordinates of e'_i wrt $\underline{e}$.)

DO 19.12 (Change of coordinates under change of basis)**.** $[v]_{\text{new}} = S^{-1}[v]_{\text{old}}$

DO 19.13. $[\underline{e}]_{\underline{e}'} = [\underline{e}']_{\underline{e}}^{-1}$

Definition 19.14 (Matrix of a linear map)**.** Let $\varphi : V \to W$ be a linear map. Let $\underline{e} = (e_1, \dots, e_n)$ be a basis of V and $\underline{f} = (f_1, \dots, f_k)$ be a basis of W. The matrix of φ wrt this pair of bases is

$$[\varphi]_{\underline{e},\underline{f}} = [[\varphi(e_1)_{\underline{f}}, \dots, \varphi(e_n)_{\underline{f}}]$$

So this is a $k \times n$ matrix.

DO 19.15 (Change of matrix under change of bases)**.** Let us have a linear map $\varphi : V \to W$. Let $\underline{e}, \underline{f}$ be "old" bases of V and W respectively, and $\underline{e}', \underline{f}'$ be new bases. Then define S, T as the change of basis matrices.

Let $A = [\varphi]_{\underline{e},\underline{f}}$ and $A' = [\varphi]_{\underline{e}',\underline{f}'}$

Then $A' = T^{-1}AS$

Corollary 19.16. *If $\varphi : V \to V$, then $[\varphi]_{new} = S^{-1}[\varphi]_{old}S$*

Corollary 19.17. *$A \sim B$ if and only if $\exists \varphi : V \to V$ and bases $\underline{e}$, $\underline{e}'$ such that $A = [\varphi]_{\underline{e}}$ and $B = [\varphi]_{\underline{e}'}$*

Corollary 19.18. *A linear transformation has a characteristic polynomial. (because similar matrices have the same characteristic polynomial)*

Definition 19.19. A is *diagonalizable* if $A \sim$ a diagonal matrix $= D = \begin{bmatrix} \lambda_1 \dots 0 \\ \vdots \ddots \vdots \\ 0 \dots \lambda_n \end{bmatrix}$

Then $f_A(t) = f_D(t) = \prod(t - \lambda_i)$

Example 19.20. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is not diagonalizable (since it is not similar to I).

DO 19.21. A matrix is diagonalizable if and only if it has an eigenbasis

Hint: $A = [\varphi]_{\underline{e}}$ and make $\underline{e}'$ the eigenbasis. Make $S = [\underline{e}']_{\underline{e}}$ Then $S^{-1}AS$ is a diagonal matrix D. You need to show that $AS = SD$

Corollary 19.22. *If f_A has n distinct roots, then A is diagonalizable.*

Caveat: I is diagonalizable, yet it has multiple eigenvalues.

DO 19.23. $A \sim B \implies \operatorname{rk} A = \operatorname{rk} B$

DO 19.24. $(\forall A)(\forall S)(\text{if } S \text{ nonsingular then } \operatorname{rk}(AS) = \operatorname{rk} A)$

DO 19.25. $\operatorname{rk}(AB) \le \operatorname{rk}(A)$ and $\operatorname{rk}(AB) \le \operatorname{rk}(B)$

19.4 Eigensubspaces. Geometric and algebraic multiplicity of eigenvalues

Definition 19.26. Algebraic multiplicity of eigenvalue λ is its multiplicity in the characteristic polynomial, i.e., it is the largest m such that $(t-\lambda)^m \mid f_A$.

Definition 19.27. Geometric multiplicity: The maximum number of linearly independent eigenvectors to λ: $\quad U_\lambda = \{x \mid Ax = \lambda x\} \leq \mathbb{F}^n$

This is the *eigensubspace* to λ.

DO 19.28. λ is an eigenvalue if and only if $\dim U_\lambda \geq 1$

The geometric multiplicity of λ is the dimension of U_λ

DO 19.29. $\dim U_\lambda = n - \mathrm{rk}(\lambda I - A)$

Hint: Rank–nullity

DO 19.30. The geometric multiplicity of λ is less than or equal to the algebraic multiplicity of λ.

DO 19.31. Over $\mathbb{C}$: $\sum_{\lambda\in\mathbb{C}}$ algebraic multiplicity of $\lambda = n$.

DO 19.32. Over $\mathbb{C}$: A is diagonalize if and only if $\sum_{\lambda\in\mathbb{C}}$ geometric multiplicity of $\lambda = n$
$\forall\lambda$, the algebraic and geometric multiplicities are equal.

DO 19.33. *

Over $\mathbb{C}$, every matrix is similar to a triangular matrix
(This is a hint for the above exercise)

19.5 Norm, orthogonality in $\mathbb{R}^n$

Over $\mathbb{R}$:

The standard dot product in $\mathbb{R}^n$: $x \cdot y = x^T y = \sum x_i y_i$
x and y are orthogonal if $x \cdot y = 0$
We define the norm $\|x\| = \sqrt{x \cdot x} = \sqrt{\sum x_i^2}$

DO 19.34. Cauchy-Schwarz inequality: $|a \cdot b| \leq \|a\|\|b\|$

DO 19.35. Triangle inequality: $\|a + b\| \leq \|a\| + \|b\|$

DO 19.36. Show Cauchy-Schwarz is equivalent to the triangle inequality.

DO 19.37. If $v_1, \ldots, v_k \in \mathbb{R}^n$ are orthogonal and non-zero, they are linearly independent.

Definition 19.38. The operator norm of $A \in \mathbb{R}^{k\times n}$:

$$\|A\| = \sup \frac{\|Ax\|}{\|x\|}$$

DO 19.39. Show this supremum is a maximum.

DO 19.40. If $A = (\alpha_{ij})$ then $\|A\| \geq |\alpha_{ij}|$

Theorem 19.41 (Spectral theorem)**.** *If $A \in M_n(\mathbb{R})$ and $A = A^T$ (A is a symmetric real matrix)then A has an orthonormal eigenbasis, i.e.,*

$$(\exists b_1, \dots b_n \in \mathbb{R}^n)(b_i \cdot b_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \ \|b_i\| = 1, \text{ and } Ab_i = \lambda_i b_i)$$

Any orthonormal system of eigenvectors can be extended to an orthonormal eigenbasis.

DO 19.42. If $A = A^T$ (A is symmetric) then $\|A\| = |\lambda|_{\max} = \max_i |\Lambda_i|$.
Hint: Spectral theorem

19.6 Elements of spectral graph theory

DO 19.43. A connected undirected graph is aperiodic if and only if it is not bipartite.

Definition 19.44. The *adjacency matrix* of a graph: $A_g = (a_{ij})$ where $a_{ij} = \begin{cases} 1 & i \sim j \\ 0 & i \not\sim j \end{cases}$

DO 19.45. The graph G is regular of degree r if and only if $\mathbf{1}$ is an eigenvector to eigenvalue r.

DO 19.46. For an r-regular graph G, let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of the adjacency matrix. (Note: these are real because the adjacency matrix is symmetric.) $\forall \lambda$, $|\lambda| \leq r$

DO 19.47. Let G be a regular graph of degree r with eigenvalues $r = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Prove: $\lambda_2 = r$ if and only if G is disconnected. (a) Show that this follows from Perron–Frobenius. (b) Prove this without using Perron–Frobenius.

DO 19.48. Let G be a connected regular graph of degree r. Then $-r$ is an eigenvalue if and only if G is bipartite.

19.7 Rate of convergence of random walk on a graph: spectral estimate

Definition 19.49 (SLEM)**.** Let A be a symmetric real matrix with eigenvalues $\lambda_1 \geq \lambda_2 \dots \geq \lambda_n$. Then $\text{SLEM}(A) = \max_i |\lambda_i| = \max |\lambda_2|, |\lambda_n|$ (second largest eigenvalue modulus)

Notation. J is the all-ones matrix.

Theorem 19.50 (Convergence rate of naive random walk on regular graph)**.** *Let G be a connected regular non-bipartite graph. Let A be the adjacency matrix of A. So $T = (1/r)A$ is the transition matrix of the naive random walk on G. Then $\lim T^t = \frac{1}{n}J$ and $\|T^t - \frac{1}{n}J\| \leq \lambda^t$ where λ is SLEM$(T) = (1/r)$SLEM(A) (so $0 < \lambda < 1$).*

Proof. We shall show that the maximum absolute value of the eigenvalues of $T^t - \frac{1}{n}J$ is λ^t.

The all-ones vector $\mathbf{1}$ is an eigenvector of T to eigenvalue 1. Let e_1 be its normalized value: $e_1 = (1/\sqrt{n})\mathbf{1}$. Let $e_1, e_2, \ldots, e_n$ be an orthonormal eigenbasis of T. We claim that $e_1, e_2, \ldots, e_n$ is also an orthonormal eigenbasis of $T^t - (1/n)J$.

For $i \geq 2$ we have $e_i \perp e_1$ and therefore $e_i \perp \mathbf{1}$. Therefore, for $i \geq 2$ we have $Je_i = 0$ and therefore $(T^t - (1/n)J)e_i = T^t e_i = \lambda_i^t e_i$. Also, $T^t e_1 = e_1$ (because T is stochastic and therefore T^t is stochastic), and the same holds for $(1/n)J$, so $(1/n)Je_1 = e_1$. Therefore $(T^t - (1/n)J)e_1 = 0$. So $e_1, e_2, \ldots, e_n$ form an orthonormal eigenbasis of $T^t - (1/n)J$ with eigenvalues (in this order) $0, \lambda_2^t, \ldots, \lambda_n^t$. The maximum absolute value among these is therefore λ^t, as claimed. □