SUMMER REU 2009
APPRENTICE COURSE
PART II

LECTURES BY LÁSZLÓ BABAI

Contents

1. Lecture 1

Definition 1.1. If $G(V, E)$ is some graph then $\overline{G} = (V, E)$ is the graph with the same vertex set V, but for which we have $\{i, j\} \in E'$ iff $\{i, j\} \notin E$.

Example 1. $\overline{K_n}$ is the graph on n vertices with no edges. $\overline{K_{r,s}}$ is the graph on $r + s$ vertices with two components which are K_r and K_s.

Definition 1.2. If G is a graph, and v, w are two vertices of G then we write $v \sim w$ (or $v \sim_G w$) to indicate that they are joined by an edge.
With this notation a graph isomorphism between $G = (V, E)$ and $H = (W, F)$ as a map f from V to W so that $v \sim_G w$ iff $f(v) \sim_H f(w)$.

Consider the problem of isomorphism: to show two graphs are isomorphic it suffices to find an isomorphism; whereas if two graphs on n vertices aren't isomorphic, a naïve proof of this fact must show that none of the $n!$ bijections on vertices are isomorphisms. Therefore it is desirable to find other methods to exhibit the fact that two graphs are non-isomorphic.

Definition 1.3. A graph G is k-colorable if we can assign colors from a choice of k to the vertices of G in such a way that no two adjacent vertices are the same color.

Consider the two graphs on six vertices pictured:
They are not isomorphic, which we can see by noting that the left graph cannot be 2-colored (it has a triangle), whereas the right-hand one can be; implicit in this is the fact

Date: July 23, 2009.
[1]Notes scribed by David Chudzicki, Aaron Marcus, Mike Miller, Jonathan Stephenson, and Rita Jimenez Rolland
[2]Notes edited by Aaron Marcus

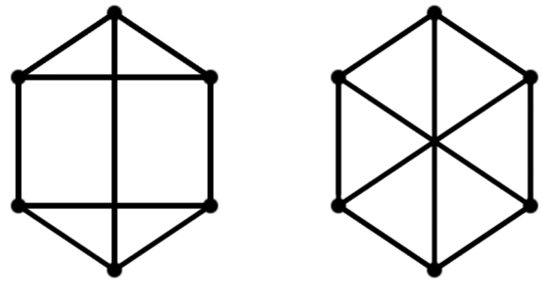

FIGURE 1. Two nonisomorphic graphs

that 2-colorability is an isomorphism *invariant*. This means that any two graphs which are isomorphic are either both 2-colorable, or neither is.

Example 2. A particularly famous graph is the Petersen graph. This graph is often used as a counterexample to conjectures in graph theory.

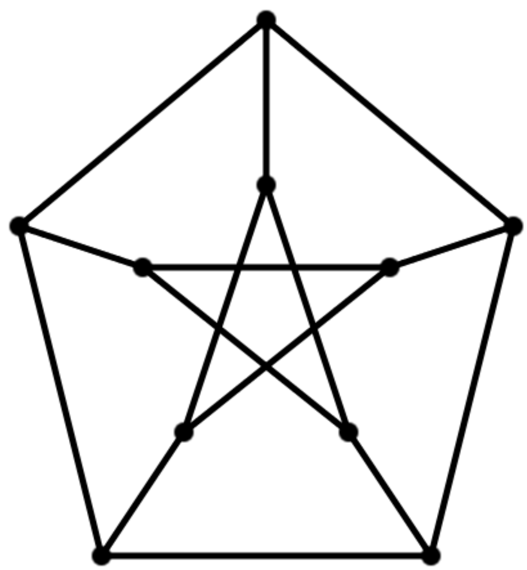

FIGURE 2. The Peterson graph

Exercise 1. Are there any graphs which are self-complementary (i.e. G isomorphic to $\overline{G}$)?

Exercise 2. If G is a graph on n vertices and is self-complementary then show that $n \equiv 0$ or $n \equiv 1$ modulo 4. $n \cong m$ modulo k means that k divides $n - m$.

Definition 1.4. If G is a graph then we define the *diameter* of $G = (V, E)$ to be $\mathrm{diam}(G) = \max_{v,w \in V} \mathrm{dist}_G(v, w)$ where $\mathrm{dist}_G(v, w)$ is the minimum length of a path joining v, w in G.

Exercise 3. If G is disconnected then $\mathrm{diam}(\overline{G}) \leq 2$.

Exercise 4. Find the minimal k so that $\min(\mathrm{diam}(G), \mathrm{diam}(\overline{G})) \leq k$ for all graphs G. Notice first that $k > 2$ by finding a counterexample.

Lemma 1.1. If G is a d-regular graph on n vertices and $\mathrm{diam}(G) \leq 2$, then $n \leq d^2 + 1$

Proof. Consider the case of one vertex; it has d neighbors, each of which have a further $d-1$ neighbors. In the case that all of these vertices are distinct, we have $1+d+d(d-1=)d^2+1$ by summing these values. $\square$

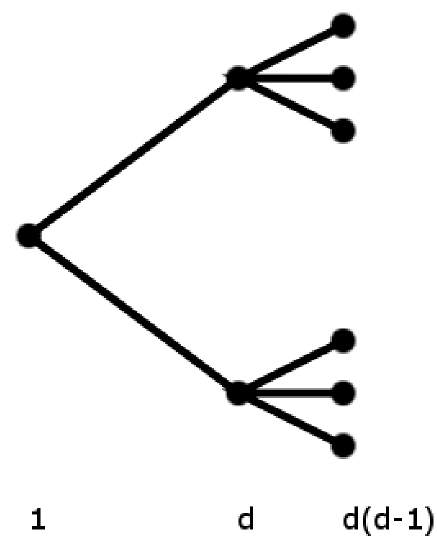

Figure 3. A vertex in a d-regular graph

Now we ask the obvious question: can we manage to have equality in this equation? If so, when?
In the case $d = 1, n = 2$ the bound is attained by a single edge joining 2 vertices. For $d = 2, n = 5$, the pentagon attains the bound. For $d = 3, n = 10$ Petersen's graph is the only example. A characteristic of all of these graphs is that they all exhibit a surprising amount of symmetry. For instance:

Exercise 5. Show that Petersen's graph has 120 automorphisms (an automorphism of a graph G is an isomorphism of G with itself).

There are no more examples which attain the bound until the case $d = 7, n = 50$; this graph is known as the Hoffman-Singleton graph. There is at most one more d for which a graph satisfies our hypotheses. We'll devote the rest of the day to proving this.
Recall that if $A, B \in M_n(K)$ then $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Exercise 6. There are matrices A, B, C for which we have $\mathrm{tr}(ABC) \neq \mathrm{tr}(ACB)$.

Definition 1.5. Two matrices $A, B \in M_n(K)$ are *similar* if there's an invertible matrix $C \in M_n(K)$ so that $B = C^{-1}AC$. Denote this by $A \sim B$.

Definition 1.6. A matrix $A \in M_n(K)$ is *diagonalizable* if it is similar to a diagonal matrix; i.e. one for which the off-diagonal entries are all 0.

Lemma 1.2. If A, B are similar then $\operatorname{tr}(A) = \operatorname{tr}(B)$ and $\det(A) = \det(B)$.

Proof.

$$\begin{aligned}\operatorname{tr}(C^{-1}AC) &= \operatorname{tr}(ACC^{-1}) \\ &= \operatorname{tr}(AI) \\ &= \operatorname{tr}(A)\end{aligned}$$

and the same proof holds for the determinant as well as trace. □

Recall that if A is n by n symmetric with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ then

$$\sum_i \lambda_i = \operatorname{tr}(A) \quad \text{and} \quad \prod_i \lambda_i = \det A$$

and we can choose an orthonormal eigebasis $u_1 \cdots u_n$ so that $Au_i = \lambda_i u_i$ for each i.

Lemma 1.3. If A is a symmetric real matrix then it is diagonalizable.

Proof. We can write

$$U = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}$$

with eigenvectors as columns. Then we have

$$AU = A\begin{bmatrix} Au_1 & Au_2 & \cdots & Au_n \end{bmatrix} = U \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

so that $A = UDU^{-1}$, for D the diagonal matrix given by the eigenvalues. □

Lemma 1.4. If $a, b, c, d, x \in \mathbb{Z}$ and $ax^4 + bx^3 + cx^2 + dx - 15 = 0$ then $x \in \{\pm 1, \pm 3, \pm 5, \pm 15\}$.

Proof.

$$15 = x(ax^3 + bx^2 + cx + d)$$

so the right side consists of factors of 15. □

Lemma 1.5. If G is d-regular with $d^2 + 1$ vertices and diameter at most 2, then each pair of vertices which are not adjacent have a unique common neighbor, and any two which are adjacent have no common neighbor.

Proof. Consider the counting method which led us to the bound $d^2 + 1$; this bound can only be attained if all of the vertices are counted exactly once. The above "forbidden occurrence" would mean that some vertices were counted twice as we built our tree of adjacent vertices. □

Theorem 1.1. (Hoffman-Singleton) If G is d-regular with $d^2 + 1$ vertices and $\operatorname{diam}(G) \leq 2$ then $d \in \{1, 2, 3, 5, 7, 57\}$.

Proof. Let $A = (a_{ij})$ be the adjacency matrix of G. Then write $A^2 = B = (b_{ij})$. The (i, j) entry of B is the number of common neighbors of vertices i, j, since we have

$$b_{ij} = \sum_{k=1}^{n} a_{ik}a_{kj}$$

Thus we have $b_{ii} = d$ for each i, and for $i \neq j$ we have $b_{ij} = 1$. Let J be the n by n matrix whose entries are all 1. We have

$$A^2 + A = (d-1)I + J$$

. Now the spectral theorem tells us that the maximal eigenvalue of our graph is $\lambda_1 = d$, and that the eigenvector corresponding to it is

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Now construct the rest of an orthogonal eigenbasis $u_2, \cdots, u_n$ so that $Au_i = \lambda_i u_i$. Then apply u_i on the right of the equation to get

$$A^2 u_i + Au_i = (d-1)u_i + Ju_i$$

and note that for $i \neq 1$ $Ju_i = 0$ because u_i is orthogonal to each row of J. Therefore we get

$$\lambda_i^2 u_i + \lambda_i u_i = (d-1)u_i$$

and hence

$$\lambda_i^2 + \lambda_i = (d-1).$$

Therefore each eigenvalue other than the first is one of

$$\lambda_+ = \frac{-1 + \sqrt{4d-3}}{2}, \lambda_- = \frac{-1 - \sqrt{4d-3}}{2}$$

and assigning multiplicity m_1 to the first possibility and m_2 to the second, we get

$$1 + m_1 + m_2 = d^2 + 1$$

by counting the eigenvalues, and

$$\operatorname{tr} A = d + m_1\lambda_+ + m_2\lambda_- = 0.$$

These equations in m_1, m_2 are independent if $\lambda_+ \neq \lambda_-$. But this cannot happen, because $4d - 3 \neq 0$. We can write $s = \sqrt{4d-3}$ for convenience. Then we have

$$-\frac{d^2}{2} + \frac{m_1 - m_2}{2}s = -d$$

so that either s is rational (and hence integer) or $m_1 = m_2$.

In the case that $m_1 = m_2$ we get $\frac{d^2}{2} = -d$ and therefore $d = 2$. This corresponds to the pentagon.

Finally, suppose that s is an integer. We get

$$d^2 - (m_1 - m_2)s = 2d$$

and then, since $s^2 + 3 = 4d$,

$$\left(\frac{s^2+3}{4^2}\right)^2 - (m_1 - m_2)s - 2\left(\frac{s^2+3}{4}\right) = 0$$

which expands to

$$s^4 - 2s^2 - 16(m_1 - m_2)s - 15 = 0$$

and using our earlier result we see that $s \in \{\pm 1, \pm 3, \pm 5, \pm 15\}$ which correspond to $d \in \{1, 3, 5, 7, 57\}$ □

Finally, some exercises:

Exercise 7. Show that if V_1, V_2 are vector subspaces of W then $V_1 + V_2 = \{v_1 + v_2 \mid v_1 \in V_1, v_2 \in V_2\}$ is also a subspace of W.

Exercise 8. Prove the modular equation: $\dim V_1 + \dim V_2 = \dim V_1 \cap \dim V_2 + \dim(V_1 + V_2)$ with V_1, V_2 as above.

Definition 1.7. A *linear map* between vector spaces V, W over K is a map f from V to W for which we have $f(v + w) = f(v) + f(w)$, and $f(av) = af(v)$ for $a \in K$. If $V = W$ we call f a linear transformation.

Exercise 9. Find a linear transformations f, g on a vector space V for which we have $fg = I$ but $gf \neq I$ for I the identity transformation.

The last exercise may be aided by considering the fact that for matrices $A, B \in M_n(\mathbb{R})$ $AB - BA = I$ cannot hold; this rules out easy finite-dimensional examples. As a further hint consider $\mathbb{R}[X]$, the vector space of polynomials with real coefficients.

2. Lecture 2

Exercise 10. Prove the triangle inequality ($d(x,z) \leq d(x,y) + d(y,z)$) for the distance function defined on $\mathbb{R}^n$ by $d(x,y) = ||x - y||$, where the norm $||\cdot||$ is defined by $||(x_1, \ldots, x_n)|| = (\sum x_i)^{\frac{1}{2}}$.

Definition 2.1. Given an $n \times n$ matrix, we define the Rayleigh quotient R (a function defined on $\mathbb{R}^n$) by

$$R(x) = \frac{x^T A x}{x^T x}.$$

Note that R attains its maximum over $\mathbb{R}^n$ on the unit ball, since $R(cx) = R(x)$.

Let A be an $n \times n$ real symmetric matrix. Then there are n eigenvalues (counting multiplicity). Let us order the eigenvalues of A from largest to smallest, $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{n-1}$. We proved the Spectral Theorem by first proving:

Theorem 2.1.

$$\lambda_0 = \max_{||x||=1} R(x) = \max_{x \neq 0} R(x).$$

We proved in class that:

Theorem 2.2.

$$\lambda_1 = \max_{x \perp u_0} R(x),$$

where u_0 is the eigenvector corresponding to λ_0.

More generally,

Theorem 2.3.

$$\lambda_i = \max_{x \perp u_0, \ldots, u_{i-1}} R(x),$$

where u_i is the eigenvector corresponding to λ_i.

By applying this to the matrix $-A$, we also obtain:

Theorem 2.4.

$$\lambda_{n-1} = \min_{x \neq 0} R(x).$$

The following theorem is important for providing a characterization of the eigenvalues that does not make reference to the eigenvectors:

Theorem 2.5. Courant-Fischer min-max theorem. With the above notation,

$$\lambda_i = \max_{\substack{W \leq \mathbb{R}^n, \\ \dim(W) = i+1}} \left(\min_{\substack{x \in W, \\ x \neq 0}} R(x) \right).$$

(The notation '$\leq$' means 'subspace'.)

Exercise 11. Let A be the adjacency matrix for a graph with n vertices, whose eigenvalues are $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{n-1}$. Remove a vertex and all of its edges, and let $\mu_0 \geq \cdots \geq \mu_{n-2}$. Then

$$\lambda_0 \geq \mu_0 \geq \lambda_1 \geq \mu_1 \geq \cdots \geq \mu_{n-2} \geq \lambda_{n-1}.$$

Exercise 12. Let λ_0 be the largest eigenvalue for the adjacency matrix of a graph G, $\deg_{\text{avg}}$ the average degree of G, and $\deg_{\max}$ the maximum degree of G. Then

$$\deg_{\text{avg}} \leq \lambda_0 \leq \deg_{\max}.$$

Claim 1. If A and B are $n \times n$ matrices and $AB = I$, then $BA = I$.

Example 3. The above claim does not hold for linear transformations in infinite dimensions: For example, in $\mathbb{R}[x]$, take the integration operator L defined by $Lf(x) = \int_x^0 f(t)dt$, and the differentiation operator D. Then $DL = I$ but $LD \neq I$.

Theorem 2.6. For an $n \times n$ matrix, the following are equivalent:

- A has a right inverse
- $\operatorname{rank}(A) = n$
- A has a left inverse
- A has a (two-sided) inverse
- $\det(A) \neq 0$.

If any (or all) are satisfied, A is 'non-singular.' Otherwise, A is 'singular.'

Exercise 13. (Dimension of the solution space of a system of homogeneous linear equations.) Let $U = \{x \mid Ax = 0\} = \ker(A)$. Then show that $\dim(U) = n - \operatorname{rank}(A)$.

More about eigenvalues: λ is an eigenvalue of A iff there is $x \neq 0$ so $Ax = \lambda x$ iff there is x so $(\lambda I - A)x = 0$ iff $\lambda I - A$ is singular iff $\det(\lambda I - A) = 0$.

Example 4. Let

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}.$$

Then $\det(A - \lambda I) = \lambda^2 - 6\lambda + 1$, which has roots $3 \pm \sqrt{8}$.

Definition 2.2. Given a (square) matrix A, the the characteristic polynomial f_A is defined by

$$f_A(t) = \det(tI - A).$$

Theorem 2.7. Given a matrix A, λ is an eigenvalue of A iff $f_A(\lambda) = 0$. (Note that in $\mathbb{C}$, there always exist solutions to this.)

Exercise 14. Let

$$R_\alpha = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}.$$

Find the characteristic polynomial, eigenvalues, and an eigenbasis.

Theorem 2.8. (Cayley-Hamilton Theorem.) For any (square) matrix A, $f_A(A) = 0$.

Exercise 15. Prove this theorem for symmetric real matrices.

Definition 2.3. Over $\mathbb{C}$, f_A will always factor as $\prod (t - \lambda_i)^{k_i}$, where the λ_i are distinct. Then k_i is the *algebraic multiplicity* of λ_i.

Definition 2.4. Given λ, let $V_\lambda = \{x \mid Ax = \lambda x\}$. Then the *geometric multiplicity* of λ is $\dim(V_\lambda) = n - \operatorname{rank}(\lambda I - A)$.

Exercise 16. The algebraic multiplicity is always greater than or equal to the geometric multiplicity. Find a 2×2 matrix where this inequality is strict.

Exercise 17. For an $n \times n$ matrix A over $\mathbb{C}$, the following are equivalent:

- The algebraic and geometric multiplicities of the eigenvalues of A are all equal.
- The matrix is A diagonalizable.

3. Lecture 3

We begin with three exercises.

Exercise 18. Let $A = \begin{pmatrix} d_1 & 1 & \dots & 1 \\ 1 & d_2 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & d_n \end{pmatrix}$. Show that if each $d_i > 1$, then A is nonsingular.

Exercise 19. Suppose that B is an $n \times n$ matrix with $B = B^T$, $b_i i = 1$ and $b_{ij} < 1$ for all $i \neq j$. Then B is not necessarily nonsingular.

Exercise 20. Prove that there exists an $n \times n$ matrix C with $c_{ii} > 0$ for all i, $c_{ij} < 0$ for all $i \neq j$ and $\operatorname{rank}(C) \leq 3$

Theorem 3.1. For any basis $e_1, \dots e_n \in V$ and n (not necessarily distinct) vectors $w_1, \dots w_n \in W$, there is a unique linear map $\varphi : V \to W$ with $\varphi(v_i) = w_i$ for all i.

Proof. Suppose that we have such a φ. If $v \in V$, then we can write v uniquely as $v = \sum a_i e_i$. By linearity, $\varphi(v) = \varphi(\sum a_i e_i) = \sum a_i \varphi(e_i) = \sum a_i w_i$. Thus, φ must be unique.

Now, define φ by the equation $\varphi(\sum a_i e_i) = \sum a_i w_i$. We must verify that φ is linear and that $\varphi(e_i) = w_i$. We leave this as an exercise. $\square$

Given a linear map, φ, let $\ker(\varphi) = \{v \in V \mid \varphi(v) = 0\} \subset V$, and let $\operatorname{im}(\varphi) = \{\varphi(v) \mid v \in V\} \subset W$. These are both subspaces.

Theorem 3.2 (Rank-Nullity Theorem)**.** $\dim(\ker \varphi) + \dim(\operatorname{im} \varphi) = \dim V$

Proof. Let $e_1, \dots e_k$ be a basis of $\ker \varphi$, and extend this to a basis of V, $e_1 \dots e_k, e_{k+1} \dots e_n$. We have that $\varphi(e_i) = 0$ if $i \leq k$. We wish to show that $\{\varphi(e_i) \mid i > k\}$ is a basis for $\operatorname{im} \varphi$. We must show that this set both spans and is linearly independent. We leave this as an exercise. $\square$

We also define $\operatorname{rank} \varphi = \dim(\operatorname{im} \varphi)$.

If A is a $k \times n$ matrix with coefficients in a field F, then A defines a linear map $\varphi_A : x \mapsto Ax \in F^k$. Then $\ker \varphi_A = \{x \in F^n \mid Ax = 0\} = \mathcal{M}_A$, the solution space of a system of homogeneous linear equations. $\operatorname{im} \varphi_A = \{Ax \mid x \in F^n\}$ is the column space of A, the span of $a_1, \dots a_n$. The rank-nullity theorem then tells us that $\dim \mathcal{M}_A + \operatorname{rank}(A) = n$.

Suppose that $e_1, \dots e_n$ is a basis for V and $f_1, \dots f_k$ is a basis for W. Then we can encode the information of φ as a matrix relative to these bases as follows. $phi(e_i)$ is a vector in W, and we can thus express it in coordinates relative to our basis for W. These coordinates form the ith column.

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

Exercise 21. If $\varphi : V \to W$, $(e_1, \dots e_n)$ is a basis of V, $(f_1, \dots, f_k)$ a basis for W. Prove that, for $v \in V$, $[\varphi(v)]_f = [\varphi]_{ef}[v]_e$.

Example 5. Let $V = \mathbb{R}_2[x] := \{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}$ and define φ by $\varphi(f) = f'$ Let us express φ as a matrix relative to $e = f = (1, x, x^2)$. Since $\varphi(1) = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2$, $\varphi(x) = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2$ and $\varphi(x^2) = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2$, the matrix is

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

We could have chosen a different basis to work with respect to, such as $(x^2, (x+1)^2, (x+2)^3)$, and the resulting matrix would have looked very different. The resulting computation of the matrix would have been drastically harder. However, there are still numerical invariants of the linear map, such as determinant and trace, which remain independent of basis. To see why, we must see how our matrix changes if we change our basis.

Suppose that we have two bases for V, $(e_1, \dots e_n)$ the old basis and $(e_1', \dots, e_n')$ the new basis. Let $\sigma : V \to V$ be the linear map sending e_i to e_i'. Note that σ is invertible because it sends a basis to a basis. Then, from $\sigma(\sum a_i e_i) = \sum a_i e_i'$, we see that $[\sigma(v)]_{\text{new}} = [v]_{\text{old}}$. If we let S and S^{-1} be the corresponding matrices, then we have $S^{-1}[v]_{\text{old}} = [v]_{\text{new}}$ and $S[v]_{\text{new}} = [v]_{\text{old}}$. If we have an old and new basis for W, we can do a similar thing there, denoting the transition matrix by T. Then, if $\varphi : V \to W$ is a linear map, we can relate the matrix form in the old basis to the matrix form in the new basis as follows.

Theorem 3.3. $[\varphi]_{\text{new}} = T^{-1}[\varphi]_{\text{old}}S$

Proof. We can prove this by applying both sides to a vector. We leave the details as an exercise. □

Corollary 3.3.1. If $\varphi : V \to V$ is a linear map, then under a change of basis, $[\varphi]_{\text{new}}$ and $[\varphi]_{\text{old}}$ are similar matrices, and given two similar matrices, they represent the same linear transformation under different bases.

Example 6. Let ρ_α be the linear transformation which rotates the plane by an angle α. If e_1 and e_2 are the standard basis vectors, then

$$[\rho_\alpha] = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

With respect to the basis $(e_1, \rho_\alpha(e_1))$, then since $\rho_\alpha(e_1)$ is the angle bisector of e_1 and $\rho_\alpha(\rho_\alpha(e_1))$, basic geometry and trigonometry allows us to show that the corresponding matrix is

$$[\rho_\alpha] = \begin{pmatrix} 0 & -1 \\ 1 & 2\cos\alpha \end{pmatrix}$$

Exercise 22. Given linear maps $V \xrightarrow{\varphi} W \xrightarrow{\psi} Z$ and corresponding bases e, f, g, then $[\psi\varphi]_{eg} = [\psi]_{fg}[\varphi]_{ef}$

Now, using that $\rho_{\alpha+\beta} = \rho_\alpha\rho_\beta$, we can recover the addition identities for sin and cos.

Exercise 23. Find the eigenvalues and eigenvectors of ρ_α.

Given a matrix A, let $f_A(t) = \det(tI - A)$ be the *characteristic polynomial* of A.

Theorem 3.4. If $A \sim B$, then $f_A(t) = f_B(t)$.

Proof. We have some matrix S such that $B = S^{-1}AS$, so $tI - B = tI - S^{-1}AS = S^{-1}(tI - A)S$, hence $\det(tI - B) = \det(S^{-1}(tI - A)S) = \det(tI - A)$ □

Note that, from $f_A(0) = \det(-A) = (-1)^n \det(A)$, we see that $(-1)^n \det(A)$ is the constant term of $f_A(t)$, and from using rook configurations to calculate the determentent, we see that the coefficient of t^n is 1 and the coefficient of t^{n-1} is $-\operatorname{tr}(A)$. Additionally, we see that if λ_i are the eigenvalues of A. counted with multiplicity, then $\sum \lambda_i = \operatorname{tr}(A)$ and $\prod \lambda_i = \det(A)$.

Exercise 24. If $e_1, \dots e_n$ are eigenvectors of an operator A, and $Ae_i = \lambda_i e_i$, with $\lambda_i \neq \lambda_j$ when $i \neq j$, show that the e_i are linearly independent.

Exercise 25. Find a 2×2 matrix which is not diagonalizable.

Exercise 26. For what p is $x^2 + 1$ irreducible over the field $\mathbb{Z}/p\mathbb{Z}$, p prime.

4. Lecture 4

Exercise 27. Prove that x^2+1 is irreducible over $\mathbb{F}_p = \mathbb{Z}/(p)$ if and only if $p \cong -1 \pmod 4$.

Exercise 28. Suppose that A is an $n \times n$ matrix (over $\mathbb{R}$ or $\mathbb{C}$). Then,

$$\det(e^A) = e^{\operatorname{tr} A}.$$

Proposition 4.1. Suppose that $\varphi\colon V \to V$ is a linear transformation and $v_1, \ldots, v_k$ are eigenvectors ($\varphi(v_i) = \lambda_i v_i$) with distinct eigenvalues ($\lambda_i \neq \lambda_j$ for $i \neq j$). Then, show that $v_1, \ldots, v_k$ are linearly independent.

Proof. Suppose that

$$\sum_{i=1}^{k} \alpha_i v_i = 0;$$

we'll show that $\alpha_1 = \cdots = \alpha_k = 0$. Applying φ to both sides of the above equation gives

$$0 = \varphi(0) = \varphi\left(\sum_{i=1}^{k} \alpha_i v_i\right) = \sum_{i=1}^{k} \alpha_i \varphi(v_i) = \sum_{i=1}^{k} \alpha_i \lambda_i v_i.$$

Applying φ again gives

$$0 = \sum_{i=1}^{k} \alpha_i \lambda_i^2 v_i.$$

By induction, we can deduce that

$$0 = \sum_{i=1}^{k} \alpha_i \lambda_i^j v_i$$

for all j. Since the determinant of the Vandermonde matrix

$$\det \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \cdots & \lambda_k \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \cdots & \lambda_k^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \lambda_3^{k-1} & \cdots & \lambda_k^{k-1} \end{pmatrix} = \prod_{i>j} (\lambda_i - \lambda_j) \neq 0$$

is non-zero (recall $\lambda_i \neq \lambda_j$ for all $i \neq j$); this implies that every $\alpha_i = 0$. (One could also do this proof by induction.) $\square$

Definition 4.1. A *Euclidean space* (over $\mathbb{R}$) is a vector space V (over $\mathbb{R}$) equipped with a positive-definite inner product

$$\begin{aligned} V \times V &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \langle x, y \rangle. \end{aligned}$$

Recall that the assignment $(x, y) \mapsto \langle x, y \rangle$ should be bilinear and symmetric. Positive definite means $\langle x, x \rangle \geq 0$ with equality if and only if $x = 0$.

Example 7. Let $V = \mathbb{R}[x]$, and given $f, g \in \mathbb{R}[x]$, define

$$\langle f, g \rangle = \int_0^1 f(t) g(t) dt.$$

With this definition, you should verify that all the conditions for a positive-definite inner product hold.

In fact, the same definition works if we take $f, g \in C[0,1]$, i.e., if we consider all continuous functions on $[0,1]$.

Also, you should check that $\mathbb{R}[x]$ with inner product given by

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(t)g(t)e^{-t^2}dt$$

is a Euclidean space.

Exercise 29. Show that if we equip $C[0, 2\pi]$ with inner product given by

$$\langle f, g \rangle = \int_0^{2\pi} fg.$$

Show that $1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots$ are orthogonal.

Proposition 4.2. Show that if $u_1, \ldots, u_k$ are non-zero pairwise orthogonal vectors, then the are linearly independent.

Proof. Suppose that

$$\sum_{i=1}^{k} \alpha_i u_i = 0.$$

Then,

$$0 = \left\langle u_j, \sum_{i=1}^{k} \alpha_i u_i \right\rangle = \sum_{i=1}^{k} \alpha_i \langle u_j, u_i \rangle = \alpha_j \langle u_j, u_j \rangle,$$

so $\alpha_j = 0$. □

Definition 4.2. If V is a Euclidean space, we define the *norm* of $x \in V$ to be

$$||x|| = \sqrt{\langle x, x \rangle}.$$

We define the distance between two vectors x and y to be

$$||x - y||\,.$$

Exercise 30. Show that with the definition above, the distance satisfies the triangle inequality.

Exercise 31. Show that in any Euclidean space, the Cauchy-Schwarz inequality holds:

$$|\langle x, y \rangle| \leq ||x||\,||y||$$

with equality if and only if x is a scalar multiple of y.

This gives an enormous number of integral inequalities. For example,

$$\left| \int_{-\infty}^{\infty} f(t)g(t)e^{-t^2}dt \right| \leq \sqrt{\int_{-\infty}^{\infty} f(t)^2 e^{-t^2}dt \int_{-\infty}^{\infty} g(t)^2 e^{-t^2}dt}$$

Gram-Schmidt Orthogonalization. The input is a sequence of vectors in a Euclidean space, say $v_1, v_2, \ldots$, and the output is another sequence of vectors $b_1, b_2, \ldots$ such that

(1) for all k, the span of $v_1, \ldots, v_k$ is equal to the span of $b_1, \ldots, b_k$
(2) for all $i \neq j$, $\langle b_i, b_j \rangle = 0$
(3) $v_k - b_k$ is in the span of $v_1, \ldots, v_{k-1}$.

In fact, (3) implies (1), and (2) and (3) uniquely determine the output. By (3), we know that $b_1 = v_1$. Now, by (3) again, we know that $v_2 - b_2$ is in span of v_1. In fact, b_2 must be on the line $v_2 - \alpha v_1$, which is the line parallel to v_1 that passes through v_2. Using (2) allows us to determine α uniquely. Uniqueness of the other vectors follows similarly.

Explicitly,

$$\begin{aligned} b_1 &= v_1 \\ b_2 &= v_2 - \frac{\langle b_1, v_2 \rangle}{||b_1||^2} b_1 \\ b_3 &= v_3 - \frac{\langle b_1, v_3 \rangle}{||b_1||^2} b_1 - \frac{\langle b_2, v_3 \rangle}{||b_2||^2} b_2 \\ &\vdots \end{aligned}$$

(The only caveat is if one of the $||b_j||^2 = 0$, but in this case, $b_j = 0$, so it doesn't matter which coefficient we take.) Note that the order of the input matters.

Exercise 32. Orthogonalize $1, x, x^2, x^3$ with respect to the inner product

$$\langle f, g \rangle = \int_0^1 fg.$$

Exercise 33. When will $b_k = 0$? $b_k = 0$ if and only if v_k is in the span of $v_1, \ldots, v_{k-1}$.

If our input is a basis, our output an orthogonal basis, since the output has the same span and same number of elements as the input basis. To turn this basis into an orthonormal basis, just normalize the vectors: divide each vector by its length. So, we've proved that every (finite-dimensional) Euclidean space has an orthonormal basis. (This is not true in infinite-dimensional spaces, but it is sort of true if we allow "infinite linear combinations".) In fact, let V be a finite-dimensional Euclidean space with ONB $e_1, \ldots, e_n$. Let $x = \sum \alpha_i e_i$, and let $y = \sum \beta_i e_i$. Then,

$$\langle x, y \rangle = \sum_i \sum_j \alpha_i \beta_j \langle e_i, e_j \rangle = \sum \alpha_i \beta_i = \vec{\alpha}^t \vec{\beta}.$$

So, this shows that V is isomorphic as a Euclidean space or (*isometric*) to $\mathbb{R}^n$ with the standard inner product.

Definition 4.3. An *orthogonal transformation* of a Euclidean space V is a linear map $\varphi \colon V \to V$ such that

$$\langle \varphi x, \varphi y \rangle = \langle x, y \rangle$$

for all x and y.

Exercise 34. All the (complex) eigenvalues of an orthogonal transformation have unit absolute value.

Exercise 35. Show that

$$\lim_{x,y\to 0^+} x^y$$

is almost always 1.

Exercise 36. What is the probability that two random positive integers are relatively prime? Does this question even make sense? In other words, first compute the probability that two random positive integers chosen from $\{1,\ldots,n\}$ are relatively prime, and then let $n\to\infty$.

Exercise 37. Find a sequence that is convergent in the Abel sense, but divergent in the Fejer sense.

Recall that to compute the Fejer sum of a series $\sum a_j$, one computes the limit of the averages

$$\sigma_N = \frac{S_1+\cdots+S_N}{N},$$

where S_j is the partial sum

$$S_j = a_1+\cdots+a_j.$$

To compute the Abel sum, one considers the power series

$$f(x) = \sum a_n x^n,$$

and computes the limit

$$\lim_{x\to 1^-} f(x).$$

Exercise 38. Find the determinant of the *circulant matrix*:

$$\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{pmatrix}$$

The way to compute the determinant is to find the eigenvalues. In fact, all circulant matrices commute, so they have a common eigenbasis.

So, find the eigenvalues of this matrix, prove that all circulant matrices commute, and find the common eigenbasis.

Exercise 39. Find a 2×2 matrix that is not diagonalizable over the complex numbers. In other words, find A such that A is not similar to a diagonal matrix. (Recall that two matrices are similar $A\sim B$ if and only if there exists an invertible matrix S such that $B = S^{-1}AS$.) Maybe

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

work. What about

$$\begin{pmatrix} 1 & 1 \\ 0 & 2. \end{pmatrix}$$

What's the characteristic polynomial of the above matrix?

$$\det\begin{pmatrix} t-1 & -1 \\ 0 & t-2 \end{pmatrix} = (t-1)(t-2).$$

So, the eigenvalues of a triangular matrix are the diagonal entries. The above matrix is diagonalizable, because we can take an eigenvector of eigenvalue 1 and an eigenvector of eigenvalue 2, and we know these are linearly independent, since they are eigenvectors with distinct eigenvalues. So, we conclude

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

We know that if A is diagonalizable,

$$A \sim \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_n \end{pmatrix},$$

where the λ_i's are the eigenvalues of A. So, we know that B cannot be diagonalizable, since if it were, it would be similar to the identity matrix, but the identity matrix is similar only to itself.

Exercise 40. Show that there cannot be more than countably many pairwise orthogonal functions in $C[0,1]$ with respect to the inner product

$$\langle f, g \rangle = \int_0^1 fg.$$

5. Lecture 5

Exercise 41. Suppose that one has 13 coins of different weights such that, if you remove any one coin, then the remaining coins can be arranged into two groups of six coins, with each group of equal total weight. Show that each coin must have the same weight. (This holds for weights in $\mathbb{Z}, \mathbb{R}, \mathbb{Q}, \mathbb{C}, \mathbb{R}[x]$. What about $\mathbb{Z}/n$?)

Exercise 42. Suppose that a rectangle is broken up into smaller rectangles such that each subrectangle has at least one integer-length side. Show that the larger rectangle has at least one integer-length side.

Exercise 43. (Generalized) Fisher Inequality. Suppose that $A_1, A_2, \dots A_m$ are subsets of $\{1, \dots, n\}$. Fix $a \neq 0$. Suppose that, for all $i \neq j$, $|A_i \cap A_j| = a$. Prove that $m \leq n$.

Recall that a Euclidean space (over $\mathbb{R}$) is a vector space V with a positive definite inner product $\langle -, - \rangle$.

Definition 5.1. $\varphi : V \to V$ is an *orthogonal transformation* if $\langle \varphi(x), \varphi(y) \rangle = \langle x, y \rangle$ for all $x, y \in V$. Such transformations are also called isometries or congruences of V.

We can define the angle between two vectors in analogy to $\mathbb{R}^2$ by saying that $\langle v, w \rangle = \|v\| \, \|w\| \cos \theta$.

Exercise 44. If $\varphi : V \to V$ is a linear transformation such that $\|\varphi x\| = \|x\|$ for all $x \in V$. Then φ is an orthogonal transformation.

Lemma 5.1. Let $O(V)$ be the set of orthogonal transformations of V. Then

(1) If $\varphi, \psi \in O(V)$, then $\varphi\psi \in O(V)$,
(2) If $\varphi \in O(V)$, then there exists $\varphi^{-1} \in O(V)$.

Thus, $O(V)$ is a group.

Proof. First, we show that $\ker \varphi = 0$. Suppose that $\varphi(x) = 0$. Then $\|x\| = \|\varphi(x)\| = \|0\| = 0$ because orthogonal transformations preserve norm. Since our inner product is positive definite, we must have $x = 0$. The other properties are straight forward. $\square$

In what follows, V will not necessarily be Euclidean. Suppose that $\varphi : V \to V$ is a linear transformation, and $U \subseteq V$ is a subspace. Then we say U is an *invariant subspace* for φ if $\varphi(u) \in U$ for all $u \in U$.

Let $e_1, \ldots e_k$ be a basis of U, and extend it to a basis $e_1, \ldots e_n$ of V. If we express φ in terms of this basis, then it is of the form

$$[\varphi]_e = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

where A is the matrix form of the restriction of φ to U. Conversely, if we can find a basis such that $[\varphi]_e$ has a 2×2 block upper triangular decomposition with the top left block a $k \times k$ square, our first k vectors span an invariant subspace.

Using this, we see that if $[\varphi]$ is upper triangular (and not just block upper triangular), then $\operatorname{span}(e_1), \operatorname{span}(e_1, e_2), \ldots \operatorname{span}(e_1, \ldots e_n)$ are all invariant subspaces.

Remark 1. Note that if $v \neq 0$, then $\operatorname{span} v$ is an invariant subspace if and only if v is an eigenvector.

We can rephrase the upper-triangular property in terms of invariant subspaces by saying that we have a maximal length chain of invariant subspaces $\{0\} \subset U_1 \subset \ldots \subset U_n = V$. Such a chain is called a *flag*.

Theorem 5.1. Over $\mathbb{C}$, any matrix is similar to an upper triangular matrix.

Proof. To begin, we find an eigenvector. This gives us an invariant subspace of dimension 1. We can then proceed by induction, the details of which is left to the reader. (Hint: Given an invariant subspace U, then the set of equivalence classes $v + U$ is a vector space which is invariant under φ.) $\square$

Exercise 45. If n is even and A is an $n \times n$ matrix, then one can reduce the rank of A no more than $n/2$ by changing no more $n^2/4$ entries.

Exercise 46. For almost all matrices, changing fewer than $n^2/4$ entries cannot yield a rank of less than or equal to $n/2$.

Remark 2. The following problem is open. Find an explicit family of matrices such that more than $n^{1.01}$ entries need to be changed to bring the rank down below $n/2$. This problem, one of matrix rigidity, was posed by Valiant around 1980.

Question 1. Given $\varphi : V \to V$ and a basis e, how can we tell from $[\varphi]_e$ whether φ is orthogonal?

We rephrase orthogonality from $\langle x, y\rangle = \langle \varphi(x), \varphi(y)\rangle$ to $[x]^T[y] = [Ax]^T[Ay] = [x]^T[A^T]Ay$. This holds for every x, y if and only if $[A]^T[A] = I$ (why?)

Definition 5.2. If $A \in M_n(\mathbb{R})$, then we say A is an *orthogonal matrix* if $A^T A = I$.

Note that if the columns of A are a_i and $A^T A = (b_{ij})$, then $b_{ij} = a_i^T a_j = \langle a_i, a_j \rangle$, so A is orthonormal if and only if its columns form an orthonormal basis. As we showed previously, this is equivalent to the rows forming an orthonormal basis because any left inverse of an $n \times n$ matrix is a right inverse.

Theorem 5.2. $A^T A = I \Leftrightarrow A^T = A^{-1} \Leftrightarrow AA^T = I$

Exercise 47. The rotation

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$$

is orthogonal, and the orthogonal matrix

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

is a reflection. What is this a reflection about?

Exercise 48. In $\mathbb{R}^3$, every orthogonal transformation has an eigenvector.

Exercise 49. In $\mathbb{R}^n$, every linear transformation has an invariant subspace of dimension at most 2.

Definition 5.3. If $S \subset V$ is a set of vectors in a Euclidean space, define $S^\perp = \{v \in V \mid \langle v, s \rangle \forall s \in S$.

Exercise 50. $S^\perp$ is a subspace.

Exercise 51. If $U \subset V$ is a subspace, then $\dim U + \dim U^\perp = \dim V$. Furthermore, every vector v can be written uniquely as $v = u + u^\perp$ where $u \in U$ and $u^\perp \in U^\perp$

Exercise 52. If $\varphi \in O(V)$, and U is an invariant subspace, then $U^\perp$ is an invariant subspace.

Exercise 53. If U is a subspace, then $U^{\perp\perp} = U$ (if V is finite dimensional).

Exercise 54. In $C[0, 2\pi]$, then $\{1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots\}^\perp = \{0\}$ where $\langle f, g \rangle = \int_0^{2\pi} fg \,\mathrm{dx}$

Definition 5.4. We say $\varphi : V \to V$ is a *symmetric transformation* if $\langle \varphi x, y \rangle = \langle x \varphi y \rangle$ for all $x, y \in V$.

Exercise 55. φ is symmetric if and only if $[\varphi]_{\text{ONB}} = [\varphi]_{\text{ONB}}^T$

Exercise 56. If φ is symmetric and U is an invariant subspace, then U^T is invariant.

Theorem 5.3. If φ is a symmetric transformation, then φ has an eigenvector.

Proof. Consider the Rayleigh quotient $R(x) = \frac{\langle x, \varphi x \rangle}{\|x\|^2}$. Then if $R(x)$ has a max at x_0, then x_0 is an eigenvector. $\square$

Exercise 57. Find a linear transformation and an invariant subspace U such that $U^\perp$ is not invariant. This can be done over $\mathbb{R}$ when the dimension is 2.

Theorem 5.4 (Spectral Theorem)**.** If $A \in M_n(\mathbb{R}), A^T = A$, then there exists $S \in O(V)$ such that

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

In the next theorem and definition, assume that V is a Euclidean space.

Theorem 5.5. For all $\varphi : V \to V$, there exists $\psi : V \to V$ such that, for all $x, y \in V$, $\langle x, \varphi y \rangle = \langle \psi x, y \rangle$.

Proof. We leave this as an exercise. Note that if e is an orthonormal basis, then $[\psi]_e = [\varphi]_e^T$. $\square$

Definition 5.5. In the theorem, ψ is call the *transpose* or *adjoint* of φ, and we write $\psi = \varphi^T$

Corollary 5.5.1. φ is symmetric $\Leftrightarrow \varphi = \varphi^T$, and φ is orthogonal $\Leftrightarrow \varphi^{-1} = \varphi^T$

Exercise 58. If $A \in O(n)$ (orthogonal matrices) and $\lambda \in \mathbb{C}$ is an eigenvalue of A, then $|\lambda| = 1$.

Hermitian (Complex Euclidean) Spaces. Let V be a vector space over $\mathbb{C}$, and $\langle -, - \rangle : V \times V \to \mathbb{C}$ such that

- $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$
- $\langle x, \lambda y \rangle = \lambda \langle x, y \rangle$
- $\langle x, y \rangle = \overline{\langle y, x \rangle}$
- $\langle x, x \rangle \geq 0$ with equality if and only if $x = 0$

Then we say that V is a *Hermitian* space.

Example 8. On $C[0, 1]$, $\langle f, g \rangle = \int_0^1 \overline{f(t)} g(t)\, \mathrm{dt}$. On $\mathbb{C}^n$, $\langle x, y \rangle = \overline{x}^T y$.

Let $A^* = \overline{A}^T$

Exercise 59. If $A = A^*$, then all eigenvalues of A are real and there exists an orthonormal basis of eigenvectors.

Exercise 60. Redo everything we've done above with the spectral theorem, Euclidean vector spaces and orthogonal transformations using Hermitian vector spaces and *unitary matrices*, those such that $B^{-1} = B^*$.

6. Lecture 6

Problem 1. Let G be a graph with n vertices and $A = (a_{ij})$ be its adjacency matrix with eigenvalues $\lambda_0 \geq \ldots \lambda_{n-1}$. Denote *average degree* by $d_{\text{average}} = \frac{d_1 + \ldots + d_n}{n}$, where d_i is the degree at the vertex i. Show that $\lambda_0 \geq d_{\text{average}}$.

Quadratic Forms. Consider a matrix $A = (a_{ij})$ and column vectors x and y, then

$$B_A(x, y) := x^T A y = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i a_{ij} y_j$$

is called *bilinear form*. Indeed B_A is linear in both x and y. In particular, when $y = x$, we get

$$Q_A(x) := B_A(x, x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

which is called a *quadratic form*. Observe that $Q_A(x)$ doesn't change if we replace the matrix A by $(A + A^T)/2$, therefore without loss of generality, we can assume that $A^T = A$ (A is symmetric). Hence there is a 1-to-1 correspondence between real quadratic forms and real symmetric matrices.

Example 9. Consider the set of points satisfying an equation of the form $Q(x, y) = c$, where $Q(x, y) = ax^2 = bxy = cy^2$ is a quadratic form, $(x, y) \in \mathbb{R}^2$ and c is a real constant. There are a number of different geometric objects of this form.

- $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ gives us an ellipse with semimajor horizontal axis of length a and semiminor vertical axis of length b.

- $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ gives us an hyperbola centered on the origin.

- $xy = 1$ give us again an hyperbola centered on the origin, but with asymptotes the Cartesian axes.

- $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$ gives two lines intersecting at $(0, 0)$

These are some examples of what we call *conic sections* on $\mathbb{R}^2$.

Example 10. *Rotated ellipse.*

The points $(x, y) \in \mathbb{R}^2$ satisfying $x^2 + 3xy + 10y^2 = 1$ correspond to a rotated ellipse $\mathcal{C}$. In terms of quadratic forms we have $Q_A(x, y) = 1$ with matrix:

$$A = \begin{pmatrix} 1 & 3/2 \\ 3/2 & 10 \end{pmatrix}$$

The characteristic polynomial of this matrix is $f_A(t) = t^2 - 11t + 31/4$ and the eigenvalues are $\lambda_{1,2} = \frac{11 \pm 3\sqrt{(10)}}{2}$. The eigenvectors v_1 and v_2 corresponding to these eigenvalues give the directions of the axes of the rotated ellipse $\mathcal{C}$. The semimajor axis has length $a = 1/\sqrt{\lambda_1}$ and the semiminor axis $b = 1/\sqrt{\lambda_2}$. To prove this last observation we go back to the general setting.

Change of basis and quadratic forms. Let us observe what happen when we change from an old basis e to a new basis e'. The linear transformation $\sigma : V \to V$ that sends e_i to e'_i gives us the matrix of change of coordinates $S := [\sigma]_e = [[e'_1]_e \cdots [e'_n]_e]$. Then the relation $[x]_{e'} = S^{-1}[x]_e$ gives the the new coordinates of x in terms of the old coordinates. We will write this as $x_{\text{new}} = S^{-1} x_{\text{old}}$ or $x_{\text{old}} = S x_{\text{new}}$.

Then

$$Q(x) = x_{\text{old}}^T A x_{\text{old}} = (S x_{\text{new}})^T A (S x_{\text{new}}) = x_{\text{new}}^T (S^T A S) x_{\text{new}},$$

and $A' = S^T A S$ is the matrix of the quadratic form in terms of the new basis.

Observe that regarding A as the matrix of a linear transformation, under a change of basis we get $A'' = S^{-1}AS$. Then $A' = A''$ just when $S^T = S^{-1}$, i.e. when S is an orthogonal matrix that sends an orthonormal basis to an orthonormal basis.

We want to switch to the orthonormal eigenbasis. Let e' denote such basis, then $Ae'_i = \lambda_i e'_i$, and the matrix of change of basis S is orthogonal. Moreover,

$$AS = A\begin{pmatrix} e'_1 & \dots & e'_n \end{pmatrix} = \begin{pmatrix} Ae'_1 & \dots & Ae'_n \end{pmatrix} = \begin{pmatrix} e'_1 & \dots & e'_n \end{pmatrix} D = SD,$$

where the matrix D is given by

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Therefore,if $x_{\text{new}}^T = (x'_1, \dots, x'_n)$ then, with respect to the eigenbasis, we have

$$Q(x) = x_{\text{new}}^T D x_{\text{new}} = \sum_{i=1}^{n} \lambda_i {x'_i}^2.$$

Back to Example 10, we get that under the change to the eigenbasis:

$$Q(x) = \lambda_1 {x'_1}^2 + \lambda_2 {x'_2}^2 = \frac{{x'_1}^2}{a^2} + \frac{{x'_2}^2}{b^2},$$

with $a = 1/\sqrt{\lambda_1}$ and $b = 1/\sqrt{\lambda_2}$ and the corresponding eigenvectors give the direction for those axes as we claimed before.

Exercise 61. Consider an equation of the form $F(x) = Q(x) + L(x) + C = 0$, where $Q(x)$ is a quadratic from, $L(x) = a_1x_1 + \dots + a_nx_n$ a linear form, and C a constant matrix. Suppose that the matrix A associated to Q is non-singular. Show that the curve $\{F(x) = 0\}$ is a translate of $\{Q(x) = \textit{const.}\}$. In particular, this means that the nature of the curve can be predicted from the quadratic form alone.

Observe that a symmetric matrix A is non-degenerate if and only if $\det A \neq 0$ if and only if none of the eigenvalues of A are zero.

Quadratic forms and graphs. Let us consider to the setting in Problem 1.Let Q the quadratic form associated to the adjacency matrix $A = (a_{ij})$ of the graph G with n vertices and m edges. Observe that the sum of the degrees of the vertices is given by

$$\sum_{i=1}^{n} d_i = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} = Q(1),$$

where $1 = (1, \dots, 1)$.

Moreover, after we prove the handshake theorem, we will see that $\sum_{i=1}^{n} d_i = Q(1) = 2m$.

Theorem 6.1 (Handshake Theorem)**.**

$$n(d_{\text{average}}) = 2m.$$

Recall that for the symmetric matrix A the largest eigenvalue $\lambda_0 = \max_{x\neq 0} R(x)$, where

$$R(x) = \frac{x^T A x}{x^T x} = \frac{Q_A(x)}{||x||^2},$$

is the Rayleigh quotient. In particular, that implies that

$$\lambda_0 \geq \frac{Q_A(1)}{||1||^2} = \frac{\sum d_i}{n} = d_{\text{average}},$$

which concludes the proof of Problem 1.

Exercise 62. Show that $\lambda_0 = d_{\text{average}}$ if and only if G is a regular graph (and therefore λ_0 is an integer).

Definition 6.1. Given a matrix A, we define

$$e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Exercise 63. Prove that the series $\sum_{k=0}^{\infty} \frac{A^k}{k!}$ converges for any matrix A. What are the eigenvalues of the matrix e^A?

Consider first the linear differential equation

$$\dot{x} = ax,$$

where $x = x(t)$, $\dot{x} = \frac{dx}{dt}$ and a is a constant. Recall that the solution of such differential equation is $x(t) = ce^{at}$, where c is any real constant. Now translate this to a more general situation: a matrix equation.

$$\dot{x}(t) = Ax(t),$$

where $x(t)^T = (x_1(t), \ldots, x_n(t))$ and A is an $n \times n$ matrix.

Exercise 64. Verify that the solution for this matrix differential equation is $x(t) = e^{At}C$, where C is a constant matrix.

This general situation actually arises:

Example 11. Consider the linear differential equation of degree 2:

$$\ddot{y} = -y.$$

Observe that if we write $x^T = (y, \dot{y})$, we get

$$\dot{x} = \begin{pmatrix} \dot{y} \\ \ddot{y} \end{pmatrix} = \begin{pmatrix} \dot{y} \\ -y \end{pmatrix} = \begin{pmatrix} \dot{y} \\ -y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} y \\ \dot{y} \end{pmatrix} = \begin{pmatrix} \dot{y} \\ -y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \dot{x}.$$

which is a matrix equation.

Exercise 65. By solving this matrix equation for x, show that you can get that expected solution for y: $y = c_1 \cos(t) + c_2 \sin(t)$, with c_1 and c_2 constants.

Example 12. The graph K_n has adjacency matrix

$$\begin{pmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 0 \end{pmatrix} = J - I,$$

where I is the $n \times n$ identity matrix and

$$J = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}.$$

Let us first get the eigenvalues of the matrix J. First observe that

$$\begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = n \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Therefore n is an eigenvalue with correspondant eigenvector $(1, \dots, 1)^T$. Any other eigenvector $x^T = (x_1, \dots, x_n)$ of J is orthogonal to $(1, \dots, 1)^T$, then it satisfies $\sum_{i=1}^n x_i = 0$ and

$$\begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Therefore, the eigenvalue for the eigenvector x is 0. Hence, J has eigenvalues $\lambda_0 = n$ and $\lambda_1 = \dots = \lambda_n = 0$, where the eigenvectors with 0 eigenvalue generate the $(n-1)$-dimensional space $\{\sum_{i=1}^n x_i = 0\}$. It follows that $J - I$, the adjacency matrix of K_n, has eigenvalues $\lambda_0 = n - 1$, $\lambda_1 = \dots = \lambda_n = -1$. Observe that $tr(J - I) = (n - 1) + (-1) + \dots + (-1) = 0$ as expected.

Claim 2. If G is connected then the largest eigenvalue is *unique* and it has an eigenvector with all entries positive.

Proof. Consider A the adjacency matrix of G.

The largest eigenvalue is given by $\lambda_0 = \max R(x)$, where $R(x) = \frac{x^T A x}{xTx}$. Suppose that $R(u) = \lambda_0$. Then $Au = \lambda_0 u$, and we may assume without loss of generality that $||u||^2 = 1$ and that the first non-zero coordinate of u is positive. If $u^T = (x_1, \dots, x_n)$ has some negative coordinate, we can take $v^T = (|x_1|, \dots, |x_n|)$ and observe that $R(v) > \lambda_0$ which contradicts the maximality of λ_0. Therefore, u has all coordinates non-negative.

Now suppose that $\lambda_1 = \lambda_0$. Then the corresponding eigenvectors must have all coordinates non-negative and be orthogonal to each other, which forces them to have disjoint support and for instance at least one zero coordinate. The proof is concluded by citing the next exercise. □

Exercise 66. If the eigenvector u associated to λ_0 has a zero coordinate, then the graph G is disconnected.

Example 13. Let G be the star graph with n vertices (the vertex v_0 is connected to each vertex v_i by an edge for $i = 1, \ldots, n-1$. Let us find its largest eigenvalue. Since G is connected, we know that λ_0 is unique and if $u^T = (x_1, \ldots, x_n)$ is such that $Au = \lambda_0 u$, then $x_i > 0$. Moreover, the symmetry of the graph implies that $x_1 = \ldots = x_n$. Otherwise we could interchange coordinates to get another eigenvector contradicting uniqueness in Claim 2. Therefore $u^T = (1, \beta \ldots \beta)$ and we know that $(\lambda_0)(1) = \beta(n-1)$ and $(\lambda_0)(\beta) = 1$. Therefore the largest eigenvalue is given by $\lambda_0 = \sqrt{n-1}$.

Theorem 6.2 (Alon-Boppana Theorem)**.** Let $\epsilon > 0$. Then for sufficiently large values of $n = |G|$, if G is d-regular, then $\lambda_1 > \sqrt{2d-1} - \epsilon$.

Exercise 67. *Weaker version of Alon-Boppana Theorem* As an application of the Spectral Theorem, prove the theorem above replacing $\sqrt{2d-1}$ by $\sqrt{d}$ in the inequality.
Hint: Use the Interlacing Theorem and the observations in the previous example.

Observe that if $Ax = \lambda x$ and f is a polynomial, then $f(A)x = f(\lambda)x$. In other words, if λ is an eigenvalue of A, then $f(\lambda)$ is an eigenvalue of the matrix $f(A)$. By a limit argument, we can also conclude that e^λ is an eigenvalue of the matrix e^A.

Example 14. Consider the matrix of the cyclic permutation $(n\ (n-1)\ \ldots\ 2\ 1)$ of the basis elements:

$$P = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & & \vdots \\ 1 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

and the polynomial $f(t) = a_0 + a_1 t + \ldots + a_{n-1} t^{n-1}$. Then

$$f(P) = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{pmatrix}$$

is the circulant matrix. If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of P, then, from the observation above $f(\lambda_1), \ldots, f(\lambda_n)$ are the eigenvalues for $f(P)$. In particular, we can compute the determinant of the circulant matrix since $\det f(P) = \prod f(\lambda_i)$. Moreover, we can find an eigenbasis for P to get an eigenbasis for the circulant matrix $f(P)$.

Exercise 68. Find an eigenbasis and the eigenvalues of the matrix P in $\mathbb{C}$.