REU APPRENTICE CLASS #11

INSTRUCTOR: LÁSZLÓ BABAI
SCRIBE: MARKUS KLIEGL

Tuesday, July 12, 2011

1. Changing Basis

Definition 1.1. Given two bases $\mathbf{b} = (b_1, ..., b_n)$ (the old basis) and $\mathbf{b}' = (b'_1, ..., b'_n)$ (the new basis), define the *basis change transformation* $\sigma : V \to V$ by

$$\sigma(b_i) := b'_i.$$

Note that this transformation is invertible (why?). Define also the corresponding matrix

$$[\sigma]_{\text{old}} := [[\sigma(b_1)]_{\text{old}}, ..., [\sigma(b_n)]_{\text{old}}] = [b'_1, ..., b'_n]_{\text{old}} = \mathbf{b'}_{\text{old}},$$

where $[v]_{\text{old}}$ denotes the coordinate vector of v with respect to the old basis. More explicitly, if

$$v = \sum_{i=1}^{n} \alpha_i b_i = \sum_{i=1}^{n} \alpha'_i b'_i,$$

then

$$[v]_{\text{old}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \qquad [v]_{\text{new}} = \begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{pmatrix}.$$

Theorem 1.2. *The coordinates of a vector v with respect to the old and new bases are related by*

$$[v]_{\text{old}} = [\sigma]_{\text{old}} \cdot [v]_{\text{new}}.$$

Recall that given a linear map $\phi : V \to W$ and bases $\mathbf{e}$ for V and $\mathbf{f}$ for W, respectively, we denote by $[\phi]_{\mathbf{e},\mathbf{f}}$ the matrix for ϕ relative to the bases $\mathbf{e}$ and $\mathbf{f}$. The j-th column of this matrix is $[\phi(e_j)]_{\mathbf{f}}$. Recall further the fundamental relation this notation establishes between matrix multiplication and the action of a linear transformation:

$$[\phi(v)]_{\mathbf{f}} = [\phi]_{\mathbf{e},\mathbf{f}} \cdot [v]_{\mathbf{e}} \qquad \text{for all } v \in V.$$

Theorem 1.3. *Let $\phi : V \to W$ be a linear function, and let $\sigma : V \to V, \mathbf{e} \mapsto \mathbf{e}'$ and $\tau : W \to W, \mathbf{f} \mapsto \mathbf{f}'$ be basis change transformations. Writing*

$$\begin{aligned} A &:= [\phi]_{\text{old}} = [\phi]_{\mathbf{e},\mathbf{f}}, \\ A' &:= [\phi]_{\text{new}} = [\phi]_{\mathbf{e}',\mathbf{f}'}, \\ S &:= [\sigma]_{\text{old}}, \\ T &:= [\tau]_{\text{old}}, \end{aligned}$$

we have

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

That is,

$$[\phi]_{\text{new}} = [\tau]_{\text{old}}^{-1} \cdot [\phi]_{\text{old}} \cdot [\sigma]_{\text{old}}.$$

Corollary 1.4. *In the special case $V = W$ and $\sigma = \tau$, we have*

$$[\phi]_{\text{new}} = S^{-1} \cdot [\phi]_{\text{old}} \cdot S.$$

In particular, taking $\phi = \sigma$, we find

$$[\sigma]_{\text{new}} := S^{-1}SS = S = [\sigma]_{\text{old}}.$$

Thus,

$$S = [\mathbf{e}']_{\mathbf{e}}, \qquad S^{-1} = [\mathbf{e}]_{\mathbf{e}'}.$$

Definition 1.5. $A, B \in M_n(F)$ are *similar* if there exist $S, S^{-1} \in M_n(F)$ such that $B = S^{-1}AS$. If A and B are similar, we write $A \sim B$.

Exercise 1.6. This is an equivalence relation.

Observation 1.7. A, B describe the same linear transformation with respect to some bases if and only if $A \sim B$.

Exercise 1.8. If $A \sim B$, then $\operatorname{rk}(A) = \operatorname{rk}(B)$, $\operatorname{tr}(A) = \operatorname{tr}(B)$, and $\det(A) = \det(B)$.

This shows that we can define the rank, trace, and determinant of a *linear transformation* as the corresponding quantity of the matrix that describes our linear transformation in some basis. These concepts are well-defined, i.e., they depend only on the linear transformation and not on the choice if the basis. A major theme in linear algebra is consequently to understand matrices up to equivalence.

Exercise 1.9 (!)**.** For a linear map $\phi : V \to W$ we define $\operatorname{rk}(\phi) := \dim \operatorname{im}(\phi)$. Let e be a basis of V and f is a basis of W. Set $A = [\phi]_{\mathrm{e},\mathrm{f}}$. Prove:

$$\operatorname{rk}(A) = \operatorname{rk}(\phi).$$

In particular, $\operatorname{rk}(A)$ does not depend on the choice of the bases e, f.

Recall that the rank-nullity theorem asserts

$$\underbrace{\dim(\ker \phi)}_{\text{nullity}} + \underbrace{\dim(\operatorname{im} \phi)}_{\text{rank}} = \dim(V).$$

2. Euler's φ function

Definition 2.1. *Euler's φ function* is the function

$$\varphi(n) := \left|\{i \mid 1 \le i \le n, \gcd(i, n) = 1\}\right|.$$

Problem 102. Prove that $\varphi(n)$ is equal to the number of primitive n-th roots of unity.

Problem 103. Show that for the matrix

$$D_n = (\gcd(i, j))_{1 \le i,j \le n},$$

we have

$$\det(D_n) = \prod_{i=1}^{n} \varphi(i).$$

Hint: Find an upper-triangular $(0, 1)$-matrix Z such that

$$D_n = Z^T \begin{pmatrix} \varphi(1) & & 0 \\ & \ddots & \\ 0 & & \varphi(n) \end{pmatrix} Z.$$

3. Eigenvalues, Eigenvectors, the Characteristic Polynomial

Definition 3.1. If $\phi : V \to V$ is a linear transformation, we say $v \in V$ is an *eigenvector* to the *eigenvalue* $\lambda \in F$ if $v \neq 0$ and $\phi(v) = \lambda v$.

Definition 3.2. If $A \in M_n(F)$, then $v \in F^n$is an *eigenvector* to *eigenvalue* $\lambda \in F$ if $v \neq 0$ and $Av = \lambda v$.

Definition 3.3. v is an *eigenvector* of ϕ if $v \neq 0$ and there exists $\lambda \in F$ such that $\phi(v) = \lambda v$. Similarly for matrices.

Definition 3.4. $\lambda \in F$ is an *eigenvalue* of ϕ if there exists $v \neq 0$ such that $\phi(v) = \lambda v$. Similarly for matrices.

Observation 3.5. $v \in V$ is an eigenvector of $\phi : V \to V$ to eigenvalue λ if and only if $[v]_{\mathbf{b}}$ is an eigenvector of $[\phi]_{\mathbf{b}}$.

Exercise 3.6. For any $\lambda \in F$ the set

$$U_\lambda = \{x \in F^n \mid Ax = \lambda x\} = \ker(\lambda I - A)$$

is a linear subspace. The nonzero elements of U_λ are precisely the eigenvectors of A to eigenvalue λ. So λ is an eigenvalue of A precisely if $\dim(U_\lambda) \geq 1$.

Observation 3.7. By the rank-nullity theorem, we have

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

Theorem 3.8. *The following are equivalent:*

(1) λ *is an eigenvalue of* $A \in M_n(F)$
(2) $\dim U_\lambda \geq 1$
(3) $\operatorname{rk}(\lambda I - A) < n$
(4) $\lambda I - A$ *is singular*
(5) $\det(\lambda I - A) = 0$

Let us state the last equivalence as a separate theorem to emphasize it; for that is how we usually detect eigenvalues.

Theorem 3.9. λ *is an eigenvalue of* A *if and only if* $\det(\lambda I - A) = 0$.

Definition 3.10. $f_A(t) = \det(tI - A)$ is called the *characteristic polynomial* of A.

Observation 3.11. λ is an eigenvalue of A if and only if $f_A(\lambda) = 0$.

Observation 3.12. If $A \in M_n(F)$, then f_A is a polynomial in t of degree n. So we may write

$$f_A(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_0.$$

From the definition of f_A, one may easily compute $a_n = 1$, $a_{n-1} = -\operatorname{tr}(A)$, and $a_0 = (-1)^n \det(A)$.

Theorem 3.13. *Over* $\mathbb{C}$*, we may factor*

$$f_A(t) = (t - \lambda_1) \cdots (t - \lambda_n), \tag{1}$$

where $\lambda_1, ..., \lambda_n$ *are the eigenvalues of* A *(not necessarily distinct here). It follows that*

$$\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i,$$

and

$$\det(A) = \prod_{i=1}^{n} \lambda_i.$$

Definition 3.14. The *algebraic mulitplicity* of an eigenvalue λ is the number of times the factor $(t - \lambda)$ occurs in the factorization (1).

Definition 3.15. The *geometric multiplicity* of an eigenvalue λ is $\dim U_\lambda = n - \operatorname{rk}(\lambda I - A)$. This is the number of linearly independent eigenvectors to eigenvalue λ.

Problem 104. For all $A \in M_n(\mathbb{C})$ and for all $\lambda \in \mathbb{C}$, we have

$$\text{algebraic multiplicity of } \lambda \geq \text{geometric multiplicity of } \lambda.$$

Problem 105. If $A \sim B$, then $f_A(t) = f_B(t)$.

Problem 106. The converse to Problem 105 is false. *Hint:* Find A, B such that $\text{rk}(A) \neq \text{rk}(B)$ but $f_A(t) = f_B(t)$.

Problem 107. (a) Consider the matrices

$$A = \begin{pmatrix} 2 & 7 \\ 0 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Note that their characteristic polynomials are equal. (Why?) Are these two matrices similar?

(b) Same question for the matrices

$$A = \begin{pmatrix} 2 & 7 \\ 0 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Definition 3.16. A matrix A is *diagonalizable* if A is similar to a diagonal matrix, that is, there exist S, S^{-1} such that $S^{-1}AS$ is diagonal.

Definition 3.17. An *eigenbasis* for ϕ is a basis consisting of eigenvectors of ϕ.

Exercise 3.18. The matrix $[\phi]_{\mathbf{b}}$ is diagonal if and only if $\mathbf{b}$ is an eigenbasis for the linear transformation ϕ.

Theorem 3.19. *(a) $[\phi]_{\mathbf{b}}$ is diagonalizable if and only if there exists an eigenbasis for ϕ.*
(b) $A \in M_n(F)$ is diagonalizable if and only if there exists an eigenbasis for A.

Problem 108. Prove that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable.

Problem 109. Over $\mathbb{C}$, a matrix A is diagonalizable if and only if for all eigenvalues λ of A, we have

$$\text{alg.mult.}(\lambda) = \text{geom.mult.}(\lambda).$$