

Linear Algebra, 13th day, Thursday 7/15/04

REU 2004. Info:

<http://people.cs.uchicago.edu/~laci/reu04>.

Instructor: László Babai

Scribe: Nick Gurski

1 Explicit form of the inverse matrix

Let $A = (\alpha_{ij})$ be an $n \times n$ matrix over a field F . Then the adjoint of A , written $\text{adj}(A)$, is the matrix $(a_{ij}) = ((-1)^{i+j} \det A_{\hat{j}\hat{i}})$ where $A_{\hat{j}\hat{i}}$ is the matrix A with row j and column i deleted (note how i and j switch roles). This allows us to prove the formula $A \cdot \text{adj}(A) = \det A \cdot I$. If we call the lefthand side $B = (\beta_{ij})$, consider β_{11} . This is equal to

$$\sum_j \alpha_{1j} (-1)^{1+j} \det A_{\hat{1}\hat{j}},$$

which is equal to the determinant of A computed by expanding along the first row; the same calculation works for β_{ii} . To show that the off-diagonal entries are zero, we compute β_{12} and note that the same technique will work for any off-diagonal. Now β_{12} is

$$\sum_j \alpha_{1j} (-1)^{2+j} \det A_{\hat{2}\hat{j}},$$

which is the determinant of the matrix A' which is the same as A except row 2 is replaced by a copy of row 1. This is zero, since the determinant of any matrix with linearly dependent rows is zero, proving the formula. Thus we have shown that if the determinant of A is nonzero, then

$$A^{-1} = \frac{1}{\det A} \text{adj } A.$$

This formula answers the question of when a matrix with integer entries has an inverse with integer entries. Since the determinant of A^{-1} is $1/\det A$, we know that if A^{-1} has integer entries then the determinant of A is ± 1 ; by the formula above we find that the converse is true as well. The adjoint of an integral matrix is integral, and if $\det A = \pm 1$, then the righthand side of the formula for the inverse is a matrix with integer entries.

2 Gram-Schmidt orthogonalization

Now we will consider a process called Gram-Schmidt orthogonalization. This is a machine that takes as input a sequence of vectors v_1, v_2, \dots and gives as output a sequence of orthogonal vectors b_1, b_2, \dots . The defining properties of the Gram-Schmidt process are that

1. the span of v_1, \dots, v_k is the same as the span of b_1, \dots, b_k (which we shall call U_k),
2. for all $i \neq j$, $\langle b_i, b_j \rangle = 0$, and
3. $v_k - b_k \in U_{k-1}$.

Note that $U_0 = \{0\}$; thus $v_1 - b_1 = 0$ or $v_1 = b_1$. We then get that $v_2 - b_2 = \mu_{2,1}b_1$, or rewriting we get

$$v_2 = b_2 + \mu_{2,1}b_1.$$

The general formula then becomes

$$v_k = b_k + \mu_{k,k-1}b_{k-1} + \mu_{k,k-2}b_{k-2} + \dots + \mu_{k,1}b_1.$$

Thus we must determine the coefficients $\mu_{k,j}$. To calculate $\mu_{3,2}$, we consider its equation and take the inner product of both sides with b_2 . This yields

$$\langle b_2, v_3 \rangle = 0 + \mu_{3,2}\|b_2\|^2 + 0,$$

using orthogonality and knowledge of b_1, b_2 . This gives an inductive definition of the $\mu_{k,j}$ as

$$\mu_{k,j} = \frac{\langle b_j, v_k \rangle}{\|b_j\|^2}.$$

This verifies the uniqueness of the coefficients. Assuming we have b_1, \dots, b_{k-1} , define $b_k = v_k - \sum \mu_{k,j}b_j$. We must verify the three properties above. The first follows by an inductive argument using this definition of b_k ; the second is a consequence of the definition of the coefficients. The third is immediate from the definition.

Exercise 13.1. Show that

$$\det G(b_1, \dots, b_k) = \det G(v_1, \dots, v_k),$$

where G is the Gram matrix. If the field of definition is \mathbb{R} , show that

$$\text{vol}(b_1, \dots, b_k) = \text{vol}(v_1, \dots, v_k),$$

where $\text{vol}(v_1, \dots, v_k)$ stands for the volume of the parallelepiped spanned by v_1, \dots, v_k . This parallelepiped is the set $\{\sum \alpha_i v_i \mid 0 \leq \alpha_i \leq 1\}$.

Exercise 13.2. Show that $b_k = 0$ if and only if $v_k \in U_{k-1}$.

We also have that if the vectors v_i form a basis, then so do the new vectors b_i , and thus an orthogonal basis. By scaling, we have thus shown the existence of an orthonormal basis.

Exercise 13.3. If the vectors v_1, \dots, v_k are orthogonal, then $b_i = v_i$ for $i = 1, \dots, k$.

It is easy to see that if $V = \text{Span}\{v_1, \dots, v_n\}$, then the vectors b_1, \dots, b_n consist of an orthogonal basis for V plus some number of additional zero vectors.

Example 13.4. Real polynomials.

The set of all polynomials with real coefficients, $\mathbb{R}[x]$, is a vector space over \mathbb{R} . We can give it an inner product by picking a density function $\rho(x)$ (with the properties that $\rho \geq 0$ and $0 < \int_{-\infty}^{\infty} x^{2n} \rho(x) dx < \infty$ for all n); this defines the inner product

$$\langle f(x), g(x) \rangle = \int_{-\infty}^{\infty} f(x)g(x)\rho(x)dx.$$

The standard basis of $\mathbb{R}[x]$ is $1, x, x^2, \dots$. Using Gram-Schmidt, we get a new orthogonal basis (depending on ρ), consisting of the polynomials $f_0 = 1, f_1, f_2, \dots$, where $\deg f_k = k$. These give special kinds of polynomials depending on which ρ we choose. If $\rho(x) = \sqrt{1-x^2}$ for $-1 \leq x \leq 1$ and $\rho(x) = 0$ otherwise, these are called the Chebyshev polynomials of the first kind; if instead we take $\rho(x) = 1/\sqrt{1-x^2}$ on the interval $(-1, 1)$, these are Chebyshev polynomials of the second kind. If $\rho(x) = \exp(-x^2)$, we get the Hermite polynomials.

Exercise 13.5. Regardless of ρ , each f_k has k real distinct roots, and the roots of f_{k-1} interlace those of f_k .

3 Algebraic numbers, minimal polynomials

We return now to more algebraic considerations. We say that a field F is algebraically closed if every nonconstant polynomial in $F[x]$ has a root. An example of such is the complex numbers, \mathbb{C} , and this result has the name of the Fundamental Theorem of Algebra. The real numbers are not algebraically closed, as $x^2 + 1$ has no root.

Exercise 13.6. There exists a countable, algebraically closed field.

Exercise 13.7. Every field F is contained in an algebraically closed field; moreover, there exists a unique (up to the appropriate isomorphism) smallest such algebraically closed field. This latter field is called the algebraic closure of F .

Exercise 13.8. Let A be the algebraic closure of the rational numbers. Then A is countable. (This is just one way to solve Exercise 5.)

Definition 13.9. We say that a complex number α is **algebraic** if there is a nonzero polynomial $f(x) \in \mathbb{Z}[x]$ such that $f(\alpha) = 0$.

Exercise 13.10. Show that the the set of all algebraic numbers forms a field.

Exercise⁺ 13.11. The field of all algebraic numbers in Exercise 13.10 is algebraically closed.

A corollary of these last two exercises is that the field A from exercise 7 is actually the same as the field of all algebraic numbers.

Let α be an algebraic number. Then the minimal polynomial $m_\alpha(x)$ of α is the polynomial with integer coefficients of least degree such that α is a root of m_α .

Exercise 13.12. The minimal polynomial is unique up to constant multiples.

Exercise 13.13. For any polynomial $f(x)$ with rational coefficients, $f(\alpha) = 0$ if and only if m_α divides f .

Example 13.14. Some minimal polynomials. The minimal polynomial of $\sqrt{2}$ is $x^2 - 2$. Let ω be the primitive cube root of unity given by $\frac{-1}{2} + \frac{\sqrt{3}}{2}i$. Then $m_\omega = x^2 + x + 1$, not $x^3 - 1$ since this last polynomial is not of minimal degree.

Exercise 13.15. The polynomial m_α is irreducible over the rational numbers.

Exercise 13.16. If f is a polynomial with integer coefficients such that $f(\alpha) = 0$ and f is irreducible, then $f = m_\alpha$.

4 The minimal polynomial of a matrix

Now we shall try to understand the minimal polynomial for a matrix A . By the Cayley-Hamilton theorem, we know that A satisfies its own characteristic polynomial. If we let f_A denote the characteristic polynomial of A , this means that $f_A(A) = 0$. Without using this theorem, we can easily prove that every matrix A is the root of some polynomial. If $f(x) = a_0 + a_1x + \cdots + a_nx^n$, then recall that

$$f(A) = a_0 \cdot I + a_1A + a_2A^2 + \cdots + a_nA^n.$$

Thus to show that A is the root of some f , we only need to show that there is some k such that the matrices I, A, A^2, \dots, A^k are linearly dependent. If A is an $n \times n$ matrix, then we know that $k = n^2$ is such a value, since the vector space $M_n(F)$ is only n^2 -dimensional and I, A, \dots, A^{n^2} is a collection of $n^2 + 1$ elements.

Now that we know (in two different ways) that every matrix is the root of a polynomial, we can ask about a minimal such. This is a quite different question than finding minimal polynomials for algebraic numbers, indicated in the following exercise.

Exercise 13.17. Over the real numbers, there are infinitely many $n \times n$ -matrices ($n \geq 2$) A with $A^2 = I$; the same holds for $A^2 = -I$.

We define m_A to be the minimal polynomial (over F) for which A is a root, just as we did for algebraic numbers.

Exercise 13.18. $f(A) = 0$ if and only if m_A divides f .

Exercise 13.19. Let f be any polynomial, and D a diagonal matrix with entries λ_i . Then $f(D)$ is a diagonal matrix with entries $f(\lambda_i)$.

Using this, we can compute that the minimal polynomial of

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

is $(x - 1)(x - 2)$.

Exercise 13.20. If D is diagonal with entries λ_i , then $m_D = \prod (x - \lambda_i)$ where the product is taken over the distinct λ_i only, i.e., m_D has no repeated roots.

Exercise 13.21. λ is an eigenvalue of A if and only if $m_A(\lambda) = 0$.

Exercise 13.22. m_A divides f_A , and f_A divides $(m_A)^n$.

Exercise 13.23. If a matrix M is triangular with main diagonal entries λ_i , then $f(M)$ is triangular with main diagonal entries $f(\lambda_i)$ for any polynomial f .

Exercise 13.24. Let f be a polynomial, $S \in GL(n, F)$, and $A \in M_n(F)$. Then $f(S^{-1}AS) = S^{-1}f(A)S$.

Remember that we say that A and B are similar, $A \sim B$ if $B = S^{-1}AS$ for some matrix S . The previous exercise then says that if A and B are similar, so are $f(A)$ and $f(B)$ for any polynomial f .

Exercise 13.25. If $A \sim B$, then $m_A = m_B$.

A corollary is that if A is diagonalizable (that is, similar to a diagonal matrix), then m_A has no repeated roots. Now consider

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

The characteristic polynomial of this matrix is $(x - 1)^2$; since the minimal polynomial must divide the characteristic polynomial, the minimal polynomial must be one of 1 , $x - 1$, and $(x - 1)^2$. It is simple to check that it is not either of the first two, so the minimal polynomial must be $(x - 1)^2$; this polynomial has a repeated root, and thus is not diagonalizable. This leads us to the following theorem, to be proved later.

Theorem 13.26. *A is diagonalizable if and only if m_A has no repeated roots.*