Linear Algebra, 6th day, Tuesday 7/6/04
REU 2004. Info: http://people.cs.uchicago.edu/∼laci/reu04.

Instructor: László Babai
Scribe: Jeff Clouse

1 Inverse matrix, rank

Today we will be discussing inverse matrices and representation of linear maps with respect to different bases. To begin, let A be an $n \times n$ matrix.

Definition 6.1. Write $B = A^{-1}$ if $AB = I_n$.

Exercise 6.2. If $AB = I_n$, then $BA = I_n$. We call such a B an **inverse** for A.

Now suppose that $A = (\underline{a_1}, \ldots, \underline{a_n})$ is a $k \times n$ matrix over the field $\mathbb{F}$, that is $A \in \mathbb{F}^{k\times n}$. If $\underline{b} \in \mathbb{F}^k$, then $\exists \underline{x} \in \mathbb{F}^n$ such that $A\underline{x} = \underline{b} \iff \underline{b} \in \mathrm{Span}(\underline{a_1}, \ldots, \underline{a_n}) =$ column space of A. If R is an $n \times k$ matrix, then $AR = A[\underline{r_1}, \ldots, \underline{r_k}] = [A\underline{r_1}, \ldots, A\underline{r_k}]$. This leads us to the following defintion.

Definition 6.3. R is a **right inverse** of A if $AR = I_k$.

We are naturally led to wonder when a right inverse should exist. Letting $\{\underline{e_1}, \ldots, \underline{e_k}\}$ be the standard basis for $\mathbb{F}^k$, we see that R exists $\iff \forall i\ \exists \underline{r_i}$ such that $A\underline{r_i} = \underline{e_i}$. This condition is equivalent to $\underline{e_1}, \ldots, \underline{e_k} \in$ column space of $A \leq \mathbb{F}^k$. This means that the column space of A is equal to $\mathbb{F}^k$; in other words, A has rank k. So, a $k \times n$ matrix A has a right inverse $\iff \mathrm{rk}\, A = k$. We restate our findings as a theorem.

Theorem 6.4. *The following are equivalent for a $k \times n$ matrix A:*

1. *A has a right inverse.*
2. $\mathrm{rk}\, A = k$.
3. *A has full row rank.*
4. *A has linearly independent rows.*

Remember that if $A = (a_{ij})$, then the transpose of A is the matrix given by $A^T = (a_{ji})$.

Exercise 6.5. We have the formula $(AB)^T = B^T A^T$.

Using the exercise, we see that $AR = I \iff R^T A^T = I$. Thus, A has a right inverse iff A^T has a left inverse. We therefore have a new theorem.

Theorem 6.6. *The following are equivalent for a $k \times n$ matrix A:*

1. *A has a left inverse.*
2. $\operatorname{rk} A = n$.
3. *A has full column rank.*
4. *A has linearly independent columns.*

The set of $n \times n$ matrices over a field $\mathbb{F}$ is very important and has its own notation, $M_n(\mathbb{F})$. In this case, the previous two theorems coincide.

Corollary 6.7. *The following are equivalent for $A \in M_n(\mathbb{F})$:*

1. $\operatorname{rk} A = n$.
2. *A has a right inverse.*
3. *A has a left inverse.*
4. *A has an inverse.*
5. $\det A \neq 0$ *(A is* **nonsingular***).*

Exercise 6.8. Let $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \mathbb{F}$ be all distinct. Now let $\gamma_{i,j} = 1/(\alpha_i - \beta_j)$. The matrix $H = (\gamma_{i,j})$ has full rank. Matrices such as H are called Cauchy-Hilbert matrices.

The set of matrices considered in the previous corollary pervades the study of linear algebra, so we give it a name.

Definition 6.9. The set of nonsingular $n \times n$ matrices over $\mathbb{F}$ is called the **General Linear Group**, and is denoted by $GL(n, \mathbb{F})$.

To justify the name, notice that $GL(n, \mathbb{F})$ is a group under matrix multiplication. Only two axioms require effort to check. First, see that $(A^{-1})^{-1} = A$, so $GL(n, \mathbb{F})$ is closed under taking inverses. Second, see that if $A, B \in GL(n, \mathbb{F})$, then $(AB)^{-1} = B^{-1}A^{-1}$. Therefore $GL(n, \mathbb{F})$ is closed under the group operation. Associativity and the existence of an identity element are clear, so we see that the general linear group is indeed a group.

Now we can use our understanding of matrix inversion to learn about changes of basis. Let $\varphi : V^n \to W^k$ be a linear map, and suppose we have two bases for each vector space: $\underline{e}, \underline{e}'; \underline{f}, \underline{f}'$. Now consider the basis change transformations

$$\sigma : V \to V, \quad \sigma(\underline{e_i}) = \underline{e_i}' \tag{1}$$
$$\tau : W \to W, \quad \tau(\underline{f_i}) = \underline{f_i}' \tag{2}$$

Define $S := [\sigma]_{\underline{e}} = [[\underline{e_1}']_{\underline{e}}, \ldots, [\underline{e_n}']_{\underline{e}}]$ and $T := [\tau]_{\underline{f}} = [[\underline{f_1}']_{\underline{f}}, \ldots, [\underline{f_k}']_{\underline{f}}]$. Similarly, let $S' := [\sigma]_{\underline{e}'}$ and $T' := [\tau]_{\underline{f}'}$. Notice that all four of these matrices are nonsingular because their columns are vector space bases. Now define the matrices $A = [\varphi]_{\underline{e},\underline{f}}$ and $A' = [\varphi]_{\underline{e}',\underline{f}'}$. Note that if x is a column vector in V, then $[\varphi x]_{\underline{f}} = [\varphi]_{\underline{e},\underline{f}}[x]_{\underline{e}}$.

Our first goal is to compare $\underline{u} = [x]_{\underline{e}}$ with $\underline{u}' = [x]_{\underline{e}'}$. Let's write $\underline{u} = u_1\underline{e_1} + \ldots + u_n\underline{e_n}$. Now consider the following simple and surprising calculation:

$$\underline{u}' = \sigma x = \sigma(\sum u_i \underline{e_i}) = \sum u_i \sigma(\underline{e_i}) = \sum u_i \underline{e_i}'.$$

This tells us that

$$\underline{u} = [x]_{\underline{e}} = [\sigma x]_{\underline{e}'} = [\sigma]_{\underline{e}'}[x]_{\underline{e}'} = S'[x]_{\underline{e}'} = S'\underline{u}'.$$

So, $\underline{u} = S'\underline{u}'$ and $\underline{u}' = (S')^{-1}\underline{u}$, accomplishing our first goal.

Now we can turn to our second goal, which is to compare A with A'. Define $\underline{v} = A\underline{u}$ and $\underline{v}' = A'\underline{u}'$. Now we can see that

$$T'\underline{v}' = \underline{v} = A\underline{u} = AS'\underline{u}'.$$

In other words,

$$(T')^{-1}AS'\underline{u}' = \underline{v}' = A'\underline{u}'.$$

Therefore, we have the formula

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

We can actually clean this formula up a bit by considering the case where $A = S$ and $A' = S'$. In this case, $\tau = \sigma$, so what above were T and T' are now S and S'. So the formula now reads: $S' = (S')^{-1}SS'$. Multiplying on the right by $(S')^{-1}$ then on the left by S', we find that $S' = S$. We could do the same thing with T to find that $T' = T$, so our nicer formula has the form:

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

Exercise 6.10. If A is nonsingular, then $\operatorname{rk}(AB) = \operatorname{rk} B$ and $\operatorname{rk}(CA) = \operatorname{rk}(C)$.

Exercise 6.11. $\operatorname{rk}(AB) \leq \max\{\operatorname{rk} A, \operatorname{rk} B\}$.

Exercise 6.12. $\operatorname{rk}(A + B) \leq \operatorname{rk} A + \operatorname{rk} B$.

2 Similarity of matrices, characteristic polynomial

Let A be an $n \times n$ matrix representing a linear map $V \to V$. Such a linear map is called a **linear transformation**. A change of basis matrix $S \in GL(n, \mathbb{F})$ represents a linear transformation. If A and A' represent the same linear transformation with respect to the two bases between which S changes, then we have $A' = S^{-1}AS$. This important concept leads us to a definition.

Definition 6.13. If $A, B \in M_n(\mathbb{F})$, then they are **similar** (or **conjugate**) if $\exists S \in GL(n, \mathbb{F})$ such that $B = S^{-1}AS$. This is denoted by $A \sim B$.

Theorem 6.14. *Let V be a vector space and φ a linear transformation. Then for any two bases $(\underline{e}, \underline{e}')$, $[\varphi]_{\underline{e}} \sim [\varphi]_{\underline{e}'}$.*

Exercise 6.15. Similarity of matrices is an equivalence relation.

Recall the determinant function $\det : M_n(\mathbb{F}) \to \mathbb{F}$.

Exercise 6.16. $\det(AB) = \det A \det B$

We have a neat formula for the determinant of an inverse matrix. Consider

$$AA^{-1} = I \Rightarrow \det(AA^{-1}) = \det I = 1.$$

Then, $\det(AA^{-1}) = \det A \det A^{-1} \Rightarrow \det A^{-1} = 1/\det A$.

Exercise 6.17. If $A \sim B$, then $\det A = \det B$.

Now recall that for an $n \times n$ matrix A, the trace of A is given by the formula $\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}$.

Exercise 6.18. For $A \in \mathbb{F}^{k \times n}$ and $B \in \mathbb{F}^{n \times k}$, we have $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Now let $A \in M_n(\mathbb{F})$ and x be a variable in $\mathbb{F}$.

Definition 6.19. The **characteristic matrix** of A is the matrix $xI - A$. The **characteristic polynomial** of A is the polynomial $f_A(x) := \det(xI - A)$.

Example 6.20. Let $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$, so $\det A = -2$ and $\operatorname{tr} A = 5$. Then the characteristic matrix of A is $xI - A = \left(\begin{smallmatrix} x-1 & -2 \\ -3 & x-4 \end{smallmatrix}\right)$. Then the characteristic polynomial of A is $f_A(x) = \left|\begin{smallmatrix} x-1 & -2 \\ -3 & x-4 \end{smallmatrix}\right| = (x-1)(x-4) - 6 = x^2 - 5x - 2 = x^2 - \operatorname{tr} A + \det A$.

Exercise 6.21. The characteristic polynomial of A is actually given by

$$f_A(x) = x^n - \operatorname{tr} A x^{n-1} + \cdots + (-1)^n \det A.$$

Exercise 6.22. If $A \sim B$, then $f_A(x) = f_B(x)$.

Since matrices which represent the same linear map with respect to different bases are similar, we can make the following definition.

Definition 6.23. Let $\varphi : V \to V$ be linear. The **characteristic polynomial** of φ is given by $f_\varphi(x) := f_A(x)$, where $A = [\varphi]$ in some basis.

Finally, let $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$.

Exercise 6.24. Calculate $f_A(A) = A^2 - (a + d)A + (ad - bc)I$ to find a curious result.