7 Seventh Class: Wed. 7/9/14

7.1 Roots vs. coefficients of a polynomial, elementary symmetric polynomials, roots vs. minors of the characteristic polynomial

Rotation matrix: With respect to any orthonormal\(^1\) basis \(e = (e_1, e_2)\), where \(e_2\) is counterclockwise to \(e_1\), the rotation matrix is the following:

\[
A := [R_\theta]_e = \begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}
\]

On the other hand, if \(f_1\) and \(f_2\) are unit vectors with an angle \(\theta \neq k\pi\) to \(f_1\) between them, the rotation matrix with respect to the basis \(f = (f_1, f_2)\) is as follows:

\[
B := [R_\theta]_f = \begin{bmatrix}
0 & -1 \\
1 & 2\cos \theta
\end{bmatrix}
\]

To derive this, note that \(R_\theta(f_1) = 0 \cdot f_1 + 1 \cdot f_2 = f_2\) and \(R_\theta(f_2) = -f_1 + (2\cos \theta)f_2\). Recall that the matrix of a transformation \(\varphi : V \rightarrow V\) with respect to the basis \(e = (e_1, \ldots, e_n)\) is

\[
[\varphi]_e = \begin{bmatrix}
[\varphi(e_1)]_e \\
[\varphi(e_2)]_e \\
\vdots \\
[\varphi(e_n)]_e
\end{bmatrix}
\]

where we associate with a vector \(x \in V\) an \(n \times 1\) matrix

\[
[x]_e = \begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_n
\end{bmatrix} \in F^n
\]

where the \(\alpha_i\) satisfy

\[
\alpha_1 e_1 + \cdots + \alpha_n e_n = x.
\]

Now we can compare:

<table>
<thead>
<tr>
<th>(A)</th>
<th>(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{Tr}(A) = \sum a_{ii} = 2\cos \theta)</td>
<td>(\text{Tr}(B) = \sum a_{ii} = 2\cos \theta)</td>
</tr>
<tr>
<td>(\text{det}(A) = (\cos \theta)^2 + (\sin \theta)^2 = 1)</td>
<td>(\text{det}(B) = 1)</td>
</tr>
</tbody>
</table>

It is a fact that the matrices of a linear transformation with respect to two different bases will have the same determinant and the same trace. Determinants and traces are just special cases of a more general invariant: the characteristic polynomial (which has the determinant and the trace as two of its coefficients, up to sign). That is, the characteristic polynomial is an invariant here.

\(^1\)“ortho-” means “perpendicular” and “-normal” means “of unit length”
**Division Theorem:** For polynomials, we have the following:

\[(\forall f, g \in F[x], g \neq 0)(\exists q, r \in F[x])(f = g \cdot q + r, \deg r < \deg g).\]

Let’s apply the division theorem to a polynomial of degree one: Let \(g(x) = x - \alpha\). Then we get

\[f(x) = (x - \alpha) \cdot q(x) + r(x).\]

Since \(\deg r < 1\), we can only have \(\deg r = 0\) (for \(r\) a non-zero constant) or \(\deg r = -\infty\) (for \(r = 0\)). In either case, \(r\) is constant. Thus we only have

\[f(x) = (x - \alpha) \cdot q(x) + r.\]

To solve for \(r\), plug in \(\alpha\) for \(x\) and obtain

\[f(\alpha) = r.\]

We have then obtained the following theorem:

**Theorem:**

\[(\forall f, \alpha)(\exists q)(f(x) = (x - \alpha) \cdot q(x) + f(\alpha)).\]

This is equivalent to

\[(\forall f, \alpha)(x - \alpha \mid f(x) - f(\alpha)).\]

**Corollary:**

\[f(\alpha) = 0 \iff x - \alpha \mid f(x).\]

**Fundamental Theorem of Algebra:**

\[(\forall f \in \mathbb{C}[x]) (\text{if } \deg f \geq 1 \text{ then } (\exists \alpha \in \mathbb{C})(f(\alpha) = 0)).\]

In other words, every polynomial over \(\mathbb{C}\) of degree at least one has a root in \(\mathbb{C}\).

We can repeat to obtain:

\[f_1(x) = (x - \alpha_1)f_2(x) = (x - \alpha_1)(x - \alpha_2)f_3(x) = \cdots = (x - \alpha_1) \cdots (x - \alpha_n) \cdot f_n(x) = a_n \prod_{i=1}^{n}(x - \alpha_i)\]

since \(f_n(x)\), as a polynomial of degree zero, is a non-zero constant.

For some \(a_n = 1\), suppose

\[f(x) = a_0 + a_1x + \cdots + a_n x^n = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n).\]

What is the relation of the coefficients \(a_i\) and the roots \(\alpha_i\)? (Recall that the roots can have multiplicity greater than one). If we multiply out the second line, we obtain

\[a_{n-1} = -(\alpha_1 + \cdots + \alpha_n) = -\sum_{i=1}^{n} \alpha_i.\]
Similarly, we obtain
\[ a_0 = (-1)^n \alpha_1 \cdots \alpha_n = (-1)^n \prod_{i=1}^n \alpha_i. \]

Continuing:
\[ a_{n-2} = \alpha_1 \alpha_2 + \alpha_1 \alpha_3 + \cdots + \alpha_1 \alpha_n + \alpha_2 \alpha_3 + \cdots + \alpha_2 \alpha_n + \cdots + \alpha_{n-1} \alpha_n \]
\[ = \sum_{i<j} \alpha_i \alpha_j. \]

Note that the sum above has \( \binom{n}{2} \) terms. Similarly:
\[ a_{n-3} = - \sum_{i<j<k} \alpha_i \alpha_j \alpha_k \]
\[ \vdots \]
\[ a_{n-t} = (-1)^t \sum_{i_1<i_2<\cdots<i_t} \alpha_{i_1} \cdots \alpha_{i_t} \]

The last equation, with \( \binom{n}{t} \) terms in the sum, uses \( \sigma_t(\alpha_1, \ldots, \alpha_n) \), called the degree-\( t \) elementary symmetric polynomial of \( \alpha_1, \ldots, \alpha_n \). Examples of this:
\[ \sigma_0(x_1, \ldots, x_n) = 1 \]
\[ \sigma_1(x_1, \ldots, x_n) = \sum_{i=1}^t x_i \]
\[ \sigma_2(x_1, \ldots, x_n) = \sum_{i<j} x_i x_j \]
\[ \sigma_t(x_1, \ldots, x_n) = \sum_{i_1<\cdots<i_t} x_{i_1} x_{i_2} \cdots x_{i_t} \]
\[ \sigma_n(x_1, \ldots, x_n) = x_1 \cdots x_n. \]

These elementary symmetric polynomials are called symmetric because we can permute the \( x_1, \ldots, x_n \) and obtain the same polynomials. There are, of course, other symmetric polynomials. For example,
\[ x_1^2 + \cdots + x_n^2 \]
is a non-elementary symmetric polynomial, where “non-elementary” just means “not on the list we have decided to call elementary.” That being said, can we express this symmetric polynomial given the elementary symmetric polynomials? In fact:
\[ (x_1 + \cdots + x_n)^2 = \sum x_i^2 + 2 \sum_{i<j} x_i x_j \]
\[ \sigma_1^2 = \sum x_i^2 + 2 \sigma_2 \]
\[ \sum x_i^2 = \sigma_1^2 - 2 \sigma_2. \]

**Exercise 7.1.** Express \( \sum x_i^3 \) by the \( \sigma_i \).

**Treasure:** To find a certain treasure chest, we need to discuss the roots of a degree-100 polynomial which begins \( f(x) = x^{100} + 5x^{99} + 13x^{98} + \cdots \). The rest of the fragment has been lost to time.

**Exercise 7.2.** Prove that not all the roots of this polynomial are real.
Exercise 7.3. Take \( f(x) \in \mathbb{Z}[x] \) where \( f(x) = a_0 + a_1 x + \cdots + a_n x^n \) where \( a_n \neq 0 \). Suppose \( f(x) = 0 \) for \( r, s \in \mathbb{Z} \) satisfying \( \gcd(r, s) = 1 \). Prove that:

\[
\begin{align*}
r &\mid a_0 \\
s &\mid a_n
\end{align*}
\]

This is a necessary but not sufficient condition on rational roots. As an example, if we were to look for the roots of the polynomial \( 5x^7 + \cdots + 8 \), we would need only try numbers satisfying

\[
\pm\{1, 2, 4, 8\} \setminus \{1, 5\}.
\]

We are reducing the question of finding rational roots to a finite search problem.

**Characteristic polynomials:** Given \( A \in M_n(F) \), we look at the characteristic polynomial

\[
f_A(t) = \det(tI - A) = a_0 + a_1 t + \cdots + a_n t^n
\]

This satisfies:

\[
\begin{align*}
a_n &= 1 \\
a_{n-1} &= -\text{Tr}(A) \\
a_{n-2} &= \text{det}(A)
\end{align*}
\]

To obtain an expression for \( a_{n-2} \), first try a \( 2 \times 2 \) matrix

\[
A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}
\]

to obtain the characteristic polynomial

\[
f_A(t) = \begin{vmatrix} t - \alpha_{11} & -\alpha_{12} \\ -\alpha_{21} & t - \alpha_{22} \end{vmatrix}
\]

\[
= (t - \alpha_{11})(t - \alpha_{22}) - \alpha_{12}\alpha_{21}
\]

\[
= t^2 - (\alpha_{11} + \alpha_{22}) + (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21})
\]

Hence

\[
a_{n-2} = \sum_{i<j} \frac{(\alpha_{ii}\alpha_{jj} - \alpha_{ij}\alpha_{ji})}{\det(A)}.
\]

This is the sum of the determinants of the symmetric \( 2 \times 2 \) minors. Generally, we will have

\[
a_{n-r} = (-1)^r \sum (\text{determinants of } r \times r \text{ symmetric minors}).
\]

This sum has \( \binom{n}{r} \) terms.

Since

\[
f_A(t) = \prod_{\lambda_i}(t - \lambda_i),
\]

We have

\[
\text{det}(A) = (-1)^{n-2} (-1)^{n-2}.
\]
we can again compare roots and coefficients:
\[
\sum \lambda_i = -a_{n-1} = \text{Tr}(A)
\]
and
\[
\sigma_2(\lambda_1, \ldots, \lambda_n) = \sum \text{determinants of } 2 \times 2 \text{ symmetric minors}
\]
\[
\vdots
\]
\[
\sigma_r(\lambda_1, \ldots, \lambda_n) = \sum r \times r.
\]
These are the two most important things to remember:
\[
\boxed{\lambda_1 \cdots \lambda_n = \det(A)}
\]
\[
\boxed{\sum \lambda_i = \text{Tr}(A)}
\]

### 7.2 Change of basis, similar matrices, diagonalizability

**Change of basis: example.** If \( e = (e_1, \ldots, e_n) \) is our old basis and \( e' = (2e_1, \ldots, 2e_n) \) our new basis, then from
\[
[x]_{\text{old}} = \begin{pmatrix}
\alpha_1 \\
\vdots \\
\alpha_n
\end{pmatrix}
\]
we obtain
\[
[x]_{\text{new}} = \begin{pmatrix}
\frac{\alpha_1}{2} \\
\vdots \\
\frac{\alpha_n}{2}
\end{pmatrix}
\]
since \( \sum_{i=1}^{n} \alpha_i e_i = x = \sum \frac{\alpha_i}{2} \cdot (2e_i) \). Since basis change is a transformation \( \sigma : e_i \mapsto e'_i \) where
\[
S = [\sigma]_e = 2I = \begin{pmatrix} 2 & 0 \\ 2 & \ddots \\ 0 & \ldots & 2 \end{pmatrix},
\]
then we note that
\[
S^{-1} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \frac{\alpha_1}{2} \\ \vdots \\ \frac{\alpha_n}{2} \end{pmatrix}.
\]

**Change of basis: general case.**

**Exercise 7.4.** (!) For any change of basis, prove that
\[
(\forall v \in V) ([v]_{\text{new}} = S^{-1}[v]_{\text{old}})
\]
where \( S \) denotes the basis change matrix
\[
S = \begin{bmatrix} [e'_1]_e & \cdots & [e'_n]_e \end{bmatrix}
\]
In other words, \( S = [\sigma]_{\text{old}} \) where \( \sigma \) is the basis change transformation: \( \sigma(e_i) = e'_i \).

**Change of basis for a linear map: example.** Now, consider two vector spaces \( V \) and \( W \) over the same field \( F \) and a linear map \( \varphi : V \to W \). If we take the bases \( e = (e_1, \ldots, e_k) \) and \( f = (f_1, \ldots, f_\ell) \) to be bases for \( V \) and \( W \), respectively, where \( k = \dim V \) and \( \ell = \dim W \), then we consider the \( \ell \times k \) matrix
\[
A = [\varphi]_e^f = [\varphi]_{\text{old}}.
\]
These are the old bases. We take the new bases \( e' = (e'_1, \ldots, e'_k) \) and \( f' = (f'_1, \ldots, f'_\ell) \) to make the new matrix
\[
A' = [\varphi]_{e'}^{f'} = [\varphi]_{\text{new}}.
\]
Consider the basis change transformations
\[
\sigma : e \mapsto e', \quad \tau : f \mapsto f'
\]
and associated basis change matrices
\[
S_{k \times k} = [e']_e, \quad T_{\ell \times \ell} = [f']_f.
\]
To change \( A \) to \( A' \), we need left-multiply the \( \ell \times k \) matrix \( A \) by an \( \ell \times \ell \) matrix, likely \( T \) or \( T^{-1} \), and right-multiply \( A \) by a \( k \times k \) matrix, likely \( S \) or \( S^{-1} \). To determine whether we are using the basis change matrices or their inverses, we take the simple case \( e'_i = 2e_i \). Then
\[
A' = 2A = (2I_\ell) \cdot A = A \cdot (2I_k) \cdot S.
\]

**Change of basis for a linear map: general case.**

**Theorem:**
\[
A' = T^{-1}AS.
\]

**Exercise 7.5.** Prove theorem. To use this, use Exercise 7.4 and the fact that
\[
(\forall x)(Ax = Bx) \implies A = B.
\]

**Linear transformations:** For a linear transformation \( \varphi : V \to V \), we have
\[
A' = S^{-1}AS.
\]
This is such a crucial relation between \( n \times n \) matrices that it has a name:

**Definition:** The matrices \( A, B \in M_n(F) \) are similar if \( \exists S \in M_n(F) \) and \( \exists S^{-1} \) such that
\[
B = S^{-1}AS.
\]
We write \( A \sim B \). This means that \( A \) and \( B \) describe the same linear transformation under different bases. Similarity is a basic equivalence relation among square matrices. What are the invariants of this equivalence relation?

**Question:** Does every invertible matrix correspond to a change of basis? Yes.
Exercise 7.6. \( A \sim B \implies \det(A) = \det(B) \).

Proof. We use the multiplicativity of the determinant \( (\det(CD) = \det(C) \cdot \det(D)) \) and the fact that \( 1 = \det(I) = \det(SS^{-1}) = \det(S)\det(S^{-1}) \) so \( \det(S^{-1}) = \det(S)^{-1} \). Then
\[
det(B) = det(S^{-1}AS) \\
= det(S^{-1})\det(A)\det(S) \\
= det(S^{-1})\det(S)\det(A) \\
= \det(A).
\]

\[\square\]

Exercise 7.7. \( A \sim B \implies \text{Tr}(A) = \text{Tr}(B) \).

Proof. This follows from the following exercise, which has already been assigned.

Exercise 7.8. \( C \in F^{k \times \ell}, D \in F^{\ell \times k} \implies \text{Tr}(CD) = \text{Tr}(DC) \).

Exercise 7.9. If \( A \sim B \), then \( f_A(t) = f_B(t) \). This implies Exercises 7.6 and 7.7.

If \( D = \text{diag}(\alpha_{11}, \ldots, \alpha_{nn}) \), then \( f_D(t) = \prod(t - \alpha_{ii}) \).

Exercise 7.10. Find \( 2 \times 2 \) matrices \( A, B \) such that \( f_A(t) = f_B(t) \) but \( A \not\sim B \).

Exercise 7.11. (a) Over \( \mathbb{C} \), every \( n \times n \) matrix is similar to a triangular matrix. (b) This is false over \( \mathbb{R} \) (hint: non-real roots).

We can compute the eigenvalues of triangular matrices just by looking at them; the characteristic polynomial is just \( \prod(t - \alpha_{ii}) \).

Exercise 7.12. Find an eigenvector corresponding to \( \lambda = 2 \) in the matrix
\[
\begin{pmatrix}
3 & 3 & 5 \\
0 & 2 & -7 \\
0 & 0 & 2
\end{pmatrix}.
\]

We might want to look for a candidate for 7.10.

\[
\begin{pmatrix}
3 & 7 \\
0 & 2
\end{pmatrix} \sim \begin{pmatrix}
3 & -1 \\
0 & 2
\end{pmatrix}
\]

Look at the linear transformation \( \varphi : V \to V \) and the matrix
\[
[\varphi]_e = \begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \ddots & 0 \\
0 & \cdots & \lambda_n
\end{pmatrix}.
\]

What can we say about \( e \)?

Def. We say that the basis \( e = (e_1, \ldots, e_n) \) of \( V \) is an eigenbasis of the linear transformation \( \varphi : V \to V \) if each \( e_i \) is an eigenvector of \( \varphi \). Note: An eigenbasis is associated with a linear transformation, but it is a basis not of the transformation but of the space.

Proposition: \([\varphi]_e \) is diagonal iff \( e \) is an eigenbasis of \( \varphi \).

Definition: \( A \in M_n(F) \) is diagonalizable if \( A \) is similar to a diagonal matrix.
**Exercise 7.13.** \( A \) is diagonalizable iff \( A \) has an eigenbasis.

Here we view \( A \) as a linear transformation of \( F^n \) under the rule \( x \mapsto Ax \) (\( x \in F^n \)).

**Exercise 7.14.**

\[
A = \begin{pmatrix} 3 & 7 \\ 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 3 & -1 \\ 0 & 2 \end{pmatrix} = B.
\]

These matrices are diagonalizable since they have two distinct eigenvalues which are therefore linearly independent, which means these vectors form a basis and thus an eigenbasis. It follows that,

\[
A \sim \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \sim B.
\]

So to solve Ex. 7.10, we have to avoid distinct eigenvalues.

**Exercise 7.15.** Decide:

\[
\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix}.
\]

**Exercise 7.16.** (Reward problem) Find the eigenvalues and the eigenvectors of the rotation matrix over \( \mathbb{C} \). \([R_\theta]\) is diagonalizable over \( \mathbb{C} \).