9/26 Draw some nice corollaries of P6 and P7.

Cor. 5.1 If $A \in \mathbb{R}^{n\times n}$ is nonsingular, then $\det A^{-1} = \frac{1}{\det A}$.

Pf.: By Cor. 4.7 ($\det I = 1$) and P7 ($\det AB = \det A \det B$),

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

Also, recall that for $A \in \mathbb{R}^{n\times n}$, the following are equivalent:

(a) $Ax = b$ has a unique solution $\forall\, b \in \mathbb{R}^n$.
(b) The only solution of $Ax = 0$ is $x = 0$.
(c) $Ax = b$ has at least one solution for each $b \in \mathbb{R}^n$.
(d) A has an inverse matrix A^{-1}.
(e) The column rank of A is n.

Can now add an additional equivalent condition: (f) $\det A \neq 0$.

Want to define $\det(T)$ for a linear $T: \mathcal{V} \to \mathcal{V}$, where $\mathcal{V}$ is finite-dimensional. First, a digression.

Changes of coordinates.

Recall that if $V = \{v_1, \ldots, v_n\}$ is a basis of $\mathcal{V}$ ($\dim \mathcal{V} = n$), then the coordinate vector of $v \in \mathcal{V}$ is the vector

$$a = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n \text{ such that } v = \sum_{j=1}^{n} \alpha_j v_j .$$

Suppose we have a new basis $W = \{w_1, \ldots, w_n\}$. Now have a coordinate vector for v given by

$$b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \text{ such that } v = \sum_{i=1}^{n} \beta_i w_i .$$

How are a and b related?

Suppose we write

$$v_j = \sum_{i=1}^{m} m_{ij} w_i \quad , \; 1 \le j \le n.$$

Set
$$M = \begin{pmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & & \vdots \\ m_{n1} & \cdots & m_{nn} \end{pmatrix}.$$

Then

$$v = \sum_j \alpha_j v_j = \sum_j \alpha_j \Big(\sum_i m_{ij} w_i\Big) = \sum_i \Big(\sum_j m_{ij}\alpha_j\Big) w_i$$

Since $v = \sum_i \beta_i w_i$, we have that

$$\beta_i = \sum_j m_{ij}\alpha_j \quad , \; 1 \le i \le n \; , \quad \text{or} \quad b = Ma.$$

Note: M is the matrix representation of the identity transformation on $\mathcal{V}$ with respect to the bases V and W.

The relation $b = Ma$ shows how coordinates with respect to V are changed to coordinates with respect to W. We can go the other way.

Suppose $w_j = \sum_i n_{ij} v_i$ for $1 \le j \le n$.

Set $N = \begin{pmatrix} n_{11} & \cdots & n_{1n} \\ \vdots & & \vdots \\ n_{n1} & \cdots & n_{nn} \end{pmatrix}$

Then $v = \sum_j \beta_j w_j = \sum_j \beta_j \Big(\sum_i n_{ij} v_i\Big) = \sum_i \Big(\sum_j n_{ij}\beta_j\Big) v_i$

Since $v = \sum_i \alpha_i v_i$, we have that

$$\alpha_i = \sum_j n_{ij}\beta_j \; , \; 1 \le i \le n, \; \text{or} \; a = Nb.$$

How are M and N related?

Changing bases from $\mathcal{V}$ to $\mathcal{W}$ and back to $\mathcal{V}$ gives

$$a = NMa \quad \text{for all } a \in \mathbb{R}^n.$$

Changing from $\mathcal{W}$ to $\mathcal{V}$ and back to $\mathcal{W}$ gives

$$b = MNb \quad \text{for all } v \in \mathbb{R}^n.$$

So we must have $MN = NM = I$, i.e., $N = M^{-1}$.

<u>Example</u>: on $\mathbb{R}^2$, suppose $\mathcal{V} = \{e_1, e_2\}$ is the natural basis ($e_1 = \begin{pmatrix}1\\0\end{pmatrix}$, $e_2 = \begin{pmatrix}0\\1\end{pmatrix}$) and $\mathcal{W} = \{w_1, w_2\}$, where

$$w_1 = \begin{pmatrix}1\\1\end{pmatrix} \text{ and } w_2 = \begin{pmatrix}-1\\1\end{pmatrix}.$$

See: $e_1 = \begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{2}w_1 - \frac{1}{2}w_2$, $e_2 = \begin{pmatrix}0\\1\end{pmatrix} = \frac{1}{2}w_1 + \frac{1}{2}w_2$.

Then $M = \begin{pmatrix}\frac{1}{2} & \frac{1}{2}\\ -\frac{1}{2} & \frac{1}{2}\end{pmatrix}$ and, from $\begin{pmatrix}a & b\\ c & d\end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix}d & -b\\ -c & a\end{pmatrix}$,

$$M^{-1} = \begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}.$$

So a general $a = \begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} \in \mathbb{R}^n$ can be written as

$$a = \beta_1 w_1 + \beta_2 w_2, \text{ where } b = \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} = Ma.$$

<u>Similarity Transformations</u>.

Suppose we have bases $\mathcal{V} = \{v_1, -, v_n\}$ and $\mathcal{W} = \{w_1, -, w_n\}$ of $\mathcal{V}$ as above, and suppose $T: \mathcal{V} \to \mathcal{V}$ is a linear transformation. Let A be the matrix representation of T with respect to $\mathcal{V}$, i.e.,

$$T(v_j) = \sum_{i=1}^{n} a_{ij} v_i, \quad 1 \le i \le n \Rightarrow A = \begin{pmatrix} a_{11} & - & a_{1n}\\ | & & | \\ a_{n1} & - & a_{nn}\end{pmatrix}$$

Note: Here, both the domain and range spaces are $\mathcal{V}$, and both the domain and range bases are $\mathcal{V}$.

Similarly, let B be the matrix representation of T with respect to $\mathcal{W}$, i.e.,

$$T(w_j) = \sum_{i=1}^{n} b_{ij} w_i, \; 1 \le j \le n \Rightarrow B = \begin{pmatrix} b_{11} & - & b_{1n} \\ | & & | \\ b_{n1} & - & b_{nn} \end{pmatrix}.$$

How are A and B related?

Let M be the matrix that transforms coordinates with respect to $\mathcal{V}$ into coordinates with respect to $\mathcal{W}$, i.e.,

$$v_j = \sum_i m_{ij} w_i, \; 1 \le j \le n \Rightarrow M = \begin{pmatrix} m_{11} & - & m_{1n} \\ | & & | \\ m_{n1} & - & m_{nn} \end{pmatrix}$$

Then

$$B = M A M^{-1} \qquad (5.1)$$

This is easy to see. It simply says that the action of T on a coordinate vector with respect to $\mathcal{W}$ is the same as first converting it to a coordinate vector with respect to $\mathcal{V}$, then applying the action of T on the converted coordinate vector, and finally converting the result back to a coordinate vector with respect to $\mathcal{W}$.

Def. 5.2. If $A, B \in \mathbb{R}^{n \times n}$ satisfy a relationship of the form (5.1), then A and B are similar.

Note: If $B = MAM^{-1}$, then $A = M^{-1}BM$, so neither A nor B plays a preferred role in the definition.

Def. 5.3. If $M \in \mathbb{R}^{n\times n}$ is nonsingular, then the transformation $A \to MAM^{-1}$ for $A \in \mathbb{R}^{n\times n}$ is a similarity transformation.

Note: A similarity transformation is linear on $\mathbb{R}^{n\times n}$.

Prop 5.4. The determinant is invariant under linear transformations, i.e., if $B = MAM^{-1}$, then $\det B = \det A$.

Pf.: Using Cor. 5.1 and P7, we have

$$\det B = \det MAM^{-1} = \det M \det A \det M^{-1} = \det A.$$

Now to define $\det T$ for a general linear transformation $T: \mathcal{V} \to \mathcal{V}$. Let $V = \{v_1, -, v_n\}$ be any basis of $\mathcal{V}$, and let $A \in \mathbb{R}^{n\times n}$ be the matrix representation of T with respect to V. Then define $\det T = \det A$.

Note: Since matrix representations of T with respect to different bases are similar and, consequently, have the same determinant, this definition of $\det T$ is independent of the choice of basis.

Recall Th. 3.9 (with $\mathcal{V} = \mathcal{W}$): The following are equivalent:

(a) $T(v) = w$ has a unique sol^n for every $w \in \mathcal{V}$.

(b) T is 1-1.

(c) T is onto.

(d) T has an inverse.

Can now add (e) $\det T \neq 0$.

Eigenvalues and eigenvectors

Suppose $T : \mathcal{V} \to \mathcal{V}$ is a linear transformation on $\mathcal{V}$.

Def. 5.5 A scalar λ is an eigenvalue of T if there is a nonzero $x \in \mathcal{V}$ such that $T(x) = \lambda x$. In this case, x is an eigenvector of T with eigenvalue λ.

Note: It is essential that x be nonzero in order for definition to be nontrivial. In general, the eigenvalues of T constitute a distinguished subset of the scalars called the spectrum of T, denoted $\sigma(T)$.

Prop. 5.6: If $\lambda \in \sigma(T)$, then the set of all eigenvectors with eigenvalue λ (together with the zero vector) is a subspace of $\mathcal{V}$, namely $\mathcal{N}(T - \lambda I)$, where I denotes the identity transformation on $\mathcal{V}$.

Pf.: $T(x) = \lambda x \Leftrightarrow (T - \lambda I)(x) = 0 \Leftrightarrow x \in \mathcal{N}(T - \lambda I)$.

It follows immediately that $\lambda \in \sigma(T) \Leftrightarrow \dim \mathcal{N}(T - \lambda I) > 0$.

Terminology: If $\lambda \in \sigma(T)$, then $\dim \mathcal{N}(T - \lambda I)$ is the geometric multiplicity of λ.

Now assume that $\mathcal{V}$ is finite-dimensional, say $\dim \mathcal{V} = n$.

Prop. 5.7. $\lambda \in \sigma(T) \Leftrightarrow \det(T - \lambda I) = 0$.

From here, the properties of eigenvalues and eigenvectors to be discussed will derive from the corresponding properties of the matrix representations of the transformations under consideration. They are often most easily discussed directly in terms of matrices, so focus on eigenvalues and eigenvectors of matrices.

Eigenvalues and eigenvectors of matrices.

Remember that (almost) all results we derive will pertain not only to matrices in $\mathbb{R}^{n\times n}$ but also to any linear transformations that they represent.

Suppose $A \in \mathbb{R}^{n\times n}$. Denote the eigenvalues of A by $\sigma(A)$.

Suppose $B = MAM^{-1}$ for nonsingular $M \in \mathbb{R}^{n\times n}$. Have

$$Ax = \lambda x \Leftrightarrow MAM^{-1}Mx = \lambda Mx = BMx = \lambda Mx,$$

and have shown

Prop. 5.8 If $B = MAM^{-1}$ for nonsingular $M \in \mathbb{R}^{n\times n}$, then $\sigma(B) = \sigma(A)$.

Note that, additionally, if (x, λ) is an eigenpair for A, then (Mx, λ) is an eigenpair for B.

Know $\lambda \in \sigma(A) \Longleftrightarrow \det(A - \lambda I) = 0$.

Recall: $\det A = \sum_{\pi} \sigma(\pi)\, a_{1\pi(1)} \cdots a_{n\pi(n)}$

Then $\det(A - \lambda I)$ is obtained from this by substituting $(a_{ii} - \lambda)$ for a_{ii} wherever a term a_{ii} appears in a summand, i.e., $\pi(i) = i$ for some i. Then, clearly,

$$\det(A - \lambda I) = p(\lambda),$$

a polynomial in λ (the characteristic polynomial).

See: $\deg p(\lambda) \le n$, since no term in the sum can involve more than n factors $(a_{ii} - \lambda)$.

Note that there is exactly one term in the sum that has n factors $(a_{ii}-\lambda)$. This corresponds to the identity permutation, which has sign $+1$. Then

$$p(\lambda) = \det(A-\lambda I) = \prod_{i=1}^{n}(a_{ii}-\lambda) + (\text{poly. of deg.} < n)$$

$$= (-1)^n \lambda^n + (\text{poly. of degree} < n).$$

So the characteristic poly. is a poly. of degree n.
Eigenvalues are roots of the <u>characteristic equation</u> $p(\lambda)=0$.

<u>Algebra Fact #1</u>: Every poly. of degree > 0 has a root <u>in</u> $\mathbb{C}$.

Follows that every $A \in \mathbb{R}^{n\times n}$ has an eigenvalue <u>in</u> $\mathbb{C}$.
Some implications for $A \in \mathbb{R}^{n\times n}$:

(1) We may be forced to consider complex eigenvalues, even if A is real.
<u>Ex.</u>: $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ (rotation) has only complex eigenvalues.

(2) However, if A is real, then the complex eigenvalues occur in <u>complex conjugate pairs</u>, i.e., if $\lambda = \mu + i\nu$ is an eigenvalue ($\mu, \nu \in \mathbb{R}$), then so is $\bar{\lambda} = \mu - i\nu$ (the complex conjugate).

Why? Have $p(\lambda) = a_0 + a_1\lambda + \dots + a_n\lambda^n$ for real $a_0, \dots, a_n$.

Then $\overline{p(\lambda)} = \overline{a_0} + \overline{a_1\lambda} + \dots + \overline{a_n\lambda^n}$

$= a_0 + a_1\bar{\lambda} + \dots + a_n(\bar{\lambda})^n = p(\bar{\lambda}).$

So $p(\lambda) = 0 \Leftrightarrow p(\bar{\lambda}) = 0$.

Ex. (cont.) Characteristic equation is

$$0 = p(\lambda) = \det(A - \lambda I) = \det\begin{pmatrix} c-\lambda & -s \\ s & c-\lambda \end{pmatrix} =$$

$$= (c-\lambda)^2 + s^2 = c^2 - 2c\lambda + \lambda^2 + s^2 = \lambda^2 - 2c\lambda + 1.$$

Roots (eigenvalues) are $\frac{2c \pm \sqrt{4c^2-4}}{2} = c \pm \sqrt{(-1)(1-c^2)}$
$= \cos\theta \pm i\sin\theta.$

(3) If $\lambda \in \sigma(A)$ is complex, then every eigenvector with eigenvalue λ must also be complex.

Why? Suppose $Ax = \lambda x$ for real $x \neq 0$. Then
$\overline{Ax} = \overline{\lambda x} \Rightarrow \bar{A}\bar{x} = \bar{\lambda}\bar{x} \Rightarrow Ax = \bar{\lambda}x.$
Writing $\lambda = \mu + i\nu$, we can add these to get
$2Ax = (\lambda + \bar{\lambda})x \Rightarrow 2Ax = 2\mu x \Rightarrow Ax = \mu x.$
Then $\mu x = Ax = \lambda x \Rightarrow \lambda = \mu$ (i.e., $\nu = 0$).

Algebra fact #2: A polynomial $p(x) = a_0 + a_1 x + \dots + a_n x^n$ (with $a_n \neq 0$) can be written as

$$p(x) = a_n \prod_{i=1}^{k} (x - x_i)^{m_i}$$

where $x_1, \dots, x_k$ are the distinct roots of p.

m_i is called the multiplicity of x_i.
x_i is a simple root if $m_i = 1$
Since $\deg p = n$, we must have $m_1 + \dots + m_k = n$.

So we can write the characteristic polynomial as

$$p(\lambda) = (-1)^n \prod_{i=1}^{k} (\lambda - \lambda_i)^{m_i}$$

where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues.

Note: A consequence is that $A \in \mathbb{R}^{n \times n}$ can have no more than n distinct eigenvalues.

Terminology: m_i is the algebraic multiplicity of λ_i.

Note: The constant term in $p(\lambda)$ is just

$$p(0) = (-1)^n \prod_{i=1}^{n} (-\lambda_i)^{m_i} = (-1)^n (-1)^{\Sigma m_i} \prod_{i=1}^{n} \lambda_i^{m_i}$$

$$= \prod_{i=1}^{n} \lambda_i^{m_i}.$$

Example: Suppose A is diagonal, e.g., $A = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$

Then

$$p(\lambda) = \det(A - \lambda I) = \det \begin{pmatrix} a_{11}-\lambda & & 0 \\ & \ddots & \\ 0 & & a_{nn}-\lambda \end{pmatrix}$$

$$= \prod_{i=1}^{n} (a_{ii} - \lambda).$$

So $p(\lambda) = 0 \iff \lambda = a_{ii}$ for some i, i.e., the eigenvalues are just the diagonal entries. Algebraic multiplicity of $\lambda \in \sigma(A)$ is just the number of times λ appears among the a_{ii}. The geometric multiplicity is, in general, $\leq$ the algebraic multiplicity (proof later). Consider $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Only eigenvalue is 1, algebraic multiplicity is 2, but geometric multiplicity is 1.