MA 511, Session 32

Similarity Transformations

Let A be a $n \times n$ matrix. If S is a nonsingular $n \times n$ matrix, then $A \longrightarrow S^{-1}AS$ is called a similarity transformation.

We saw that this related, e.g. to changes of variables in systems of differential equations: Suppose $B = S^{-1}AS$, then with $y = S^{-1}x$,

$$\frac{dA}{dt} = Ax \text{ is equivalent to } \frac{dB}{dt} = By.$$

Definition: Is $B = S^{-1}AS$, we say that A and B are similar.

Remark: Similarity of matrices is an equivalence relation, i.e. it is reflexive, symmetric, and transitive.

Theorem: Similar matrices have the same eigenvalues.

Proof: The result will follow after we show that they actually have the same characteristic polynomial.

$$\begin{aligned}
p(\lambda) &= \det(B - \lambda\mathbb{I}) = \det(S^{-1}AS - \lambda\mathbb{I}) \\
&= \det(S^{-1}AS - \lambda S^{-1}\mathbb{I}S) \\
&= \det\big(S^{-1}(A - \lambda\mathbb{I})S\big) \\
&= (\det S^{-1})\big(\det(A - \lambda\mathbb{I})\big)(\det S) \\
&= \det(A - \lambda\mathbb{I}).
\end{aligned}$$

Similarity and Change of Basis

The standard basis in $\mathbb{R}^3$ is $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, and when we see a vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ we take it as represented in this basis:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

However, we can take another basis, for example $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, and then the <u>same</u> point would be represented by the (different) triple $\begin{pmatrix} y \\ x \\ z \end{pmatrix}$ (note that even the ordering of the basis vectors is important, not just what vectors they are!).

Let now $T : V \longrightarrow V$ be a linear transformation from the vector space V into itself, and let $\beta = \{v_1, \ldots, v_n\}$ be a basis for V. Then, T is entirely determined by its action on β and this is summarized in the matrix of T with respect to the basis β, $(T)_{\beta\beta} = A = (a_{ij})$. Recall that this means

$$\begin{cases} T(v_1) = a_{11}v_1 + \cdots + a_{n1}v_n = w_1 \\ \quad\quad\vdots \\ T(v_n) = a_{1n}v_1 + \cdots + a_{nn}v_n = w_n, \end{cases}$$

and A transforms *coordinates* of vectors in the basis β into *coordinates* of the transformed vector in the same basis β. For example, if $x = c_1 v_1 + \cdots + c_n v_n$, then

$$T(x) = c_1 w_1 + \cdots + c_n w_n,$$

has coordinates Ac in the basis β.

If we change the basis β to another basis $\tilde{\beta} = \{\tilde{v}_1, \ldots, \tilde{v}_n\}$, the representation of the linear transformation changes, though the transformation itself does not. If

$$\begin{cases} v_1 = m_{11}\tilde{v}_1 + \cdots + m_{n1}\tilde{v}_n \\ \quad\vdots \\ v_n = m_{1n}\tilde{v}_1 + \cdots + m_{nn}\tilde{v}_n, \end{cases}$$

then the matrix $M = (m_{ij})$ represents the identity transformation ($\mathrm{id}\colon V \longrightarrow V, \quad \mathrm{id}\,(x) = x$) in the bases β and $\tilde{\beta}$ in that order. This means that $M = (\mathrm{id}\,)_{\beta\tilde{\beta}}$ and also $M^{-1} = (\mathrm{id}\,)_{\tilde{\beta}\beta}$. Let $D = (T)_{\tilde{\beta}\tilde{\beta}}$ be the representation of T with respect to the basis $\tilde{\beta}$.

The vector $\tilde{c}$ of coordinates of x in the basis $\tilde{\beta}$ is now given by $\tilde{c} = Mc$ (i.e. $M^{-1}\tilde{c} = c$). Indeed, let

$B = (\, v_1 \quad \ldots \quad v_n\,)$ and $\tilde{B} = (\, \tilde{v}_1 \quad \ldots \quad \tilde{v}_n\,)$. Then, $B = \tilde{B}M$ (that is $BM^{-1} = \tilde{B}$) and

$$T(x) = BAc = BAM^{-1}\tilde{c} = \tilde{B}MAM^{-1}\tilde{c},$$

which means that

$$(T)_{\tilde{\beta}\tilde{\beta}} = D = MAM^{-1} = (\mathrm{id}\,)_{\beta\tilde{\beta}}(T)_{\beta\beta}(\mathrm{id}\,)_{\tilde{\beta}\beta}.$$

We see again a similarity transformation from A to D that now represents the change of basis from β to $\tilde{\beta}$.

Theorem: If U_1, U_2 are unitary $n\times n$ matrices, then so is U_1U_2.

Proof: $(U_1U_2)^H(U_1U_2) = U_2^H U_1^H U_1 U_2 = U_2^H U_2 = \mathbb{I}$.

Theorem: For any $n\times n$ matrix A, there is a unitary $n\times n$ matrix U such that $U^H AU$ is upper triangular.

Proof: Let λ_1 be an eigenvalue of A and x_1 a corresponding eigenvector of unit length, $\|x_1\| = 1$. Let now U_1 be a $n \times n$ unitary matrix having x_1 as first column. It follows that $AU_1 = U_1A_1$, where A_1 has its first column equal to $\lambda_1 e_1$, that is $A_1 = U_1^{-1} A U_1$ has the first column as desired.

Next, consider the $(n-1) \times (n-1)$ matrix $\tilde{A}_1$ obtained from A_1 by removing the first row and first column, and let λ_2 be an eigenvalue for it with a corresponding associated eigenvector $x_2 \in \mathbb{R}^{n-1}$. Let $\tilde{U}_2$ be any $(n-1) \times (n-1)$ unitary matrix having x_2 for first column and define the $n \times n$ unitary matrix U_2 with e_1 as first column and $(e_1)^T$ as first row, and $\tilde{U}_2$ for the remaining rows and columns. It follows that $A_2 = U_2^{-1}(U_1^{-1} A U_1) U_2$ has for first column $\lambda_1 e_1$ and for second column $c e_1 + \lambda_2 e_2$, as desired.

Repeating this process $n-1$ times, we have unitary matrices $U_1, \ldots, U_{n-1}$ such that

$$U_{n-1}^{-1} \ldots U_1^{-1} A U_1 \ldots U_{n-1} = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & \lambda_2 & \ldots & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

We now define the $n \times n$ unitary matrix

$$U = U_1 \dots U_{n-1},$$

and it follows that $U^{-1}AU$ is upper triangular.

We can now prove the spectral theorem we stated in the last session for Hermitian matrices

Corollary: If A is Hermitian, it is unitarily similar to a diagonal matrix.

Proof: Take a unitary U such that $U^H AU = T$ is upper triangular. Now, $(U^H AU)^H = T^H$ is lower triangular, but it is also $(U^H AU)^H = U^H AU = T$. Thus, T is actually diagonal.

Corollary: If A is unitary, it is unitarily similar to a diagonal matrix.

Proof: Take a unitary U such that $U^H AU = T$ is upper triangular. Since the product of unitary matrices is unitary, T is unitary. This means $T^H = T^{-1}$. Thus,

$$\begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ 0 & t_{22} & \dots & t_{2n} \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & 0 & \dots & t_{nn} \end{pmatrix} \begin{pmatrix} \bar{t}_{11} & 0 & \dots & 0 \\ \bar{t}_{12} & \bar{t}_{22} & \dots & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ \bar{t}_{1n} & \bar{t}_{2n} & \dots & \bar{t}_{nn} \end{pmatrix} = \mathbb{I}_n,$$

which, using the last row of T, gives the relations $|t_{nn}|^2 = 1$, and $t_{jn} = 0$ for $1 \leq j < n$. Then, using the penultimate row of T, we see that $t_{j,n-1} = 0$ for $1 \leq j < n-1$, and $|t_{n-1,n-1}|^2 = 1$. Repeating this process, we see that T is diagonal (and also $|t_{jj}|^2 = 1$ for $1 \leq j \leq n$).

Example: Find the eigenvalues and eigenspaces of the operator T defined on $V = \mathcal{C}(-\infty, \infty)$ by $T(f) = \int_0^x f(t)\, dt$.

<u>Solution</u>: We want scalars λ such that

$$T(f) = \int_0^x f(t)\, dt = \lambda\, f(x), \qquad -\infty < x < \infty$$

for some nonzero function $f \in V$.

It follows from the fundamental theorem of calculus that

$$f(x) = \lambda f'(x).$$

If $\lambda = 0$, then $f \equiv 0$ says there is no eigenvector (eigenvectors are nonzero by definition). Thus, 0 is not an eigenvalue.

If $\lambda \neq 0$, then $\dfrac{f'}{f} = \dfrac{1}{\lambda}$ and, by elementary ODEs we know $f(x) = ce^{\frac{1}{\lambda}x}$ for some constant c. But at $x = 0$ we have $Tf(0) = 0 = \lambda c$ which implies $c = 0$ and thus $f \equiv 0$ cannot be an eigenvector. Therefore, this operator T has no eigenvalues.

Example: Show that all real numbers are eigenvalues of the operator T defined on $V = \mathcal{C}^2(-\infty, \infty)$ by $T(f) = f''$, and find all the eigenspaces.
Solution: We want scalars λ such that

$$T(f) = f'' = \lambda\, f(x), \qquad -\infty < x < \infty$$

for some nonzero function $f \in V$.

There are three different cases, $\lambda < 0$, $\lambda = 0$, and $\lambda > 0$. In any case, the characteristic polynomial of the ODE is $r^2 - \lambda$.

(i) In the first case, there is a pair of complex conjugate roots $r = \pm\sqrt{-\lambda}\, i$ and the general solution is $f(x) = c_1 \cos\sqrt{-\lambda}x + c_2 \sin\sqrt{-\lambda}x$.

(ii) In the second case there is a double root $r = 0$ and the general solution is $f(x) = c_1 + c_2 x$.

(iii) In the third case its roots are real, $r = \pm\sqrt{\lambda}$ and the general solution is $f(x) = c_1 e^{\sqrt{\lambda}x} + c_2 e^{-\sqrt{\lambda}x}$.

As for the eigenspaces, they are all 2-dimensional (planes), with the following bases:

(i) $\lambda < 0, \quad S_\lambda = \text{span}\,\{\cos\sqrt{-\lambda}x, \sin\sqrt{-\lambda}x\}$.

(ii) $\lambda = 0, \quad S_\lambda = \text{span}\,\{1, x\}$.

(iii) $\lambda > 0, \quad S_\lambda = \text{span}\,\{e^{\sqrt{\lambda}x}, e^{-\sqrt{\lambda}x}\}$.