7.2 Orthogonal Diagonalization

Def – A, B $n \times n$ matrices

A and B are orthogonal similar if there is an orthogonal matrix P such that

$$P^T A P = B$$

Def:- If A is orthogonal similar to a diagonal matrix D

$$P^T A P = D,$$

we say that A is orthogonal diagonalizable and that P orthogonally diagonalizes A.

Thm 7.2.1 A $n \times n$ matrix. The following statements are equivalent:

a) A is orthogonally diagonalizable.

b) A has an orthonormal set of n eigenvectors.

c) A is symmetric.

Before proving this Jump to Exerc. #7) in page 11.

Proof.- a) $\Rightarrow$ b)

a) $\Rightarrow$ There exists P orthogonal ($P^{-1} = P^T$) Such that

$$P^T A P = D \quad \text{or} \quad P^{-1} A P = D$$

If $P = [\vec{p}_1 \dots \vec{p}_n]$ using thm 5.2.1 we conclude that the set $S = \{\vec{p}_1, .., \vec{p}_n\}$ is a set of n linearly indep eigenvectors. Moreover, using thm 7.1.1, applied to P we arrive to the conclusion that S is also orthonormal. Therefore,

$S = \{\vec{p}_1 \dots, \vec{p}_n\}$ is an orthonormal set of n eigenvectors of A. ✓

b) $\Rightarrow$ a) If b) $\overset{\text{Thm 6.3.1}}{\Rightarrow}$ this set ($S = \{\vec{p}_1, .., \vec{p}_n\}$ of eigenvectors) is lin. indep

and by thm 5.2.1 A is diagonalizable

and $P^{-1} A P = D$, where $P = [\vec{p}_1 \cdots \vec{p}_n]$

Since the columns of P are orthonormal, using thm 7.1.1 we conclude that P is orthogonal: $P^{-1} = P^T$

Therefore, $D = P^{-1} A P = P^T A P$

and A is orthogonal diagonalizable ✓

Corollary.- If A is orthogonal diagonalizable ($P^T A P = D$) then $P = [\vec{p}_1\, \vec{p}_2 \cdots \vec{p}_n]$ where the set $S = \{\vec{p}_1, .., \vec{p}_n\}$ is a set of n orthonormal eigenvectors.

The proof is contained above.

a) $\Rightarrow$ c)

If (a) $\Rightarrow$ $P^TAP = D$, where P is orthogonal.

$$\Rightarrow \quad A = PDP^T \Rightarrow A^T = (PDP^T)^T = (P^T)^T D^T P^T = $$
$$= PDP^T = A. \checkmark$$

c) $\Rightarrow$ a) more complicated.

<u>Thm 7.2.2.</u>

a) A Symm. $\Rightarrow$ Eigenvalues of A are real.

b) A Symm. $\Rightarrow$ Eigenvectors from different eigenspaces are orthogonal.

<u>Proof</u> (b) Recall that

$$A\vec{x} \cdot \vec{y} = \vec{y}^T(A\vec{x}) = (\vec{y}^TA)\vec{x} = (A^T\vec{y})^T\vec{x} =$$
$$= \vec{x} \cdot A^T\vec{y} \qquad (*)$$

Therefore, if $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors of A corresponding to the eigenvalues λ_1 and λ_2, respectively,

$$A\vec{v}_1 \cdot \vec{v}_2 \overset{\text{prev. result}}{=} \vec{v}_1 \cdot A^T\vec{v}_2 \overset{\text{Hyp.}}{=} \vec{v}_1 \cdot A\vec{v}_2$$

$$\Rightarrow \lambda_1\vec{v}_1 \cdot \vec{v}_2 = \vec{v}_1 \cdot \lambda_2\vec{v}_2 \overset{\lambda_2 \text{ is real}}{=} \lambda_2(\vec{v}_1 \cdot \vec{v}_2)$$

$$\Rightarrow (\lambda_1 - \lambda_2)(\vec{v}_1 \cdot \vec{v}_2) = 0, \quad \text{Since } \lambda_1 \neq \lambda_2$$

$$\Downarrow \; \vec{v}_1 \cdot \vec{v}_2 = 0. \Rightarrow \vec{v}_1 \perp \vec{v}_2 \checkmark$$

Hint:
Show
$A\vec{v}_1 \cdot \vec{v}_2 = A\vec{v}_2 \cdot \vec{v}_1$

Natural try: (All real)

$A\vec{v}_1 \cdot \vec{v}_2 = \lambda_1\vec{v}_1 \cdot \vec{v}_2$

$A\vec{v}_2 \cdot \vec{v}_1 = \lambda_2\vec{v}_2 \cdot \vec{v}_1$

$A\vec{v}_1 \cdot \vec{v}_2 - A\vec{v}_2 \cdot \vec{v}_1 = (\lambda_1 - \lambda_2)\vec{v}_1 \cdot \vec{v}_2 \overset{?}{=} 0$

Now $A\vec{v}_2 \cdot \vec{v}_1 \overset{\text{real}}{=} \vec{v}_1 \cdot A\vec{v}_2 =$
$= \vec{v}_1 \cdot A^T\vec{v}_2 \overset{(*)}{=} A\vec{v}_1 \cdot \vec{v}_2$

7.2 Ex. 7) <u>Find P that orthogonally diagonalizes A, and determine $P^{-1}AP = P^TAP$.</u>

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Clearly, <u>A is symmetric</u>

then, thm 7.2.1 establishes that <u>A is orthog. diag.</u>

We need to construct P whose columns are formed by eigenvectors of A that form an orthonormal set

$$|\lambda I - A| = \begin{vmatrix} \lambda-2 & +1 & +1 \\ +1 & \lambda-2 & +1 \\ 1 & 1 & \lambda-2 \end{vmatrix} = (\lambda-2)^3 + 2 - (\lambda-2) - (\lambda-2) - (\lambda-2)$$

$$= (\lambda-2)^3 - 3(\lambda-2) + 2 = 0$$

$$= (\lambda-2)(\lambda^2-4\lambda+4) - 3\lambda + 6 + 2 = \lambda^3 - 6\lambda^2 + 12\lambda - \cancel{8} - 3\lambda + \cancel{8} = 0$$

$$= \lambda(\lambda^2 - 6\lambda + 9) = \lambda(\lambda-3)^2 \Rightarrow \begin{cases} \lambda_1 = 0 \\ \lambda_2 = 3 \end{cases} \text{Eigenvalues}$$

For $\boxed{\lambda_2 = 3}$ eigenvectors satisfy

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$x_1 = -x_2 - x_3$

$$\tilde{x} = \begin{bmatrix} -x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} -x_3 \\ 0 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Then, any eigenvector of A for $\lambda_2=3$ is a linear combination of $\vec{u}_2=\begin{bmatrix}-1\\1\\0\end{bmatrix}$ and $\vec{u}_3=\begin{bmatrix}-1\\0\\1\end{bmatrix}$

Since they are lin. indep, they form a basis for the eigenspace corresponding to $\lambda_2=3$.

On the other hand, for $\boxed{\lambda_1=0}$ the eigenvectors satisfy

$$(\overset{=0}{\lambda_1}I-A)\vec{x}=\begin{bmatrix}-2&1&1\\1&-2&1\\1&1&-2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

or

$$\begin{bmatrix}-2&1&1&0\\1&-2&1&0\\1&1&-2&0\end{bmatrix}\sim\begin{bmatrix}1&1&-2&0\\1&-2&1&0\\-2&1&1&0\end{bmatrix}\sim\begin{bmatrix}1&1&-2&0\\0&3&-3&0\\0&3&-3&0\end{bmatrix}\sim$$

$$\sim\begin{bmatrix}1&1&-2&0\\0&3&-3&0\\0&0&0&0\end{bmatrix}\Rightarrow\begin{aligned}&3x_2-3x_3=0\\&\boxed{x_2=x_3}\\&x_1=2x_3-x_2=x_3\end{aligned}$$

then

$$\vec{x}=\begin{bmatrix}x_3\\x_3\\x_3\end{bmatrix}=x_3\begin{bmatrix}1\\1\\1\end{bmatrix}.$$

and $\vec{u}_1=\begin{bmatrix}1\\1\\1\end{bmatrix}$ is an eigenvector corresponding to $\lambda_1=0$.

Now, the matrix A is Symm. Therefore (Thm 7.2.2) the eigenvector $\vec{u}_1 \perp \vec{u}_2$ and $\vec{u}_1 \perp \vec{u}_3$ (This can be easily verified directly) and the set $S = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is a set of Linearly independent eigenvectors. Then, using thm 5.2.1

A is diagonalizable or

$$\boxed{P^{-1}AP = D}$$

where

$$P = \begin{array}{c} \begin{matrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{matrix} \\ \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \end{array} \quad \text{and} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Since A is Symmetric, we can do better and find $\hat{P}$ orthogonal that orthogonally diagonalizes A: $\boxed{\hat{P}^T A \hat{P} = D}$

According to thm 7.2.1 what we need is 3 orthonormal eigenvectors. They can be easily obtained from

$S = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$.

First, we find $\hat{u}_2$ and $\hat{u}_3$ from $\vec{u}_2$ and $\vec{u}_3$ using Gram-Schmidt process

In fact,

$$\hat{u}_2 = \frac{\vec{u}_2}{\|\vec{u}_2\|} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$\langle \hat{u}_3, \hat{u}_2 \rangle = [-1\ 0\ 1] \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} = \\ = 1/\sqrt{2}$$

$$\vec{u}_3{}' = \vec{u}_3 - \langle \vec{u}_3, \hat{u}_2 \rangle \hat{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - 1/\sqrt{2} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} =$$

$$= \begin{bmatrix} -1 + 1/2 \\ -1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \end{bmatrix}, \quad \|\vec{u}_3{}'\| = \sqrt{\tfrac{1}{4} + \tfrac{1}{4} + 1} = \\ = \tfrac{1}{2}\sqrt{6} = \sqrt{6}/2$$

then an eigenvector $\hat{u}_3$ such that $\|\hat{u}_3\| = 1$ is given by $\hat{u}_3 = \frac{\vec{u}_3{}'}{\|\vec{u}_3{}'\|}$

or $$\hat{u}_3 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

Since $\vec{u}_1 \perp \vec{u}_2$ and $\vec{u}_1 \perp \vec{u}_3$ then $\vec{u}_1$ is also orthog. to $\hat{u}_2$ and $\hat{u}_3$. Then

$$\hat{S} = \{ \hat{u}_1, \hat{u}_2, \hat{u}_3 \} = \left\{ \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix} \right\}$$ is an orthonormal set of eigenvectors

---

<u>Remark:</u> It can be easily proven that $\vec{u}_3{}'$ is also an eigenvector assoc. to $\lambda_2 = 3$

In fact, $$A\vec{u}_3{}' = A\vec{u}_3 - \frac{1}{\sqrt{2}} A\hat{u}_2 = \lambda_2 \vec{u}_3 - \frac{1}{\sqrt{2}} \lambda_2 \hat{u}_2 = \lambda_2 \left[\vec{u}_3 - \frac{1}{\sqrt{2}} \hat{u}_2 \right] = \lambda_2 \vec{u}_3{}'.$$

Thus, using thm 7.2.1 A is orthogonally diagonalizable

There is $\hat{P}$ such that $\boxed{\hat{P}^T A \hat{P} = D}$

where

$$\hat{P} = \begin{array}{c} \begin{array}{ccc} \hat{u}_1 & \hat{u}_2 & \hat{u}_3 \end{array} \\ \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix} \end{array} \quad \text{and} \quad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Spectral Decomposition of a Matrix A.

First, consider the product of

$$\begin{array}{c} \begin{array}{cc} \vec{c}_1 & \vec{c}_2 \end{array} \\ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \end{array} \begin{array}{cc} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} & \begin{array}{c} \vec{r}_1 \\ \vec{r}_2 \end{array} \end{array} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{21}b_{11} & a_{21}b_{12} \end{bmatrix} + \begin{bmatrix} a_{12}b_{21} & a_{12}b_{22} \\ a_{22}b_{21} & a_{22}b_{22} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \end{bmatrix} + \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} \begin{bmatrix} b_{21} & b_{22} \end{bmatrix} =$$

$$= \vec{c}_1 \vec{r}_1 + \vec{c}_2 \vec{r}_2 .$$

This result is true for general A, B $n \times n$ matrices.

Now, if for a given matrix $A_{n\times n}$ Symmetric

$\lambda_1, .., \lambda_n$ are eigenvalues with Corresp orthonormal eigenvectors $\vec{u}_1, .., \vec{u}_n$, then

$$D = P^T A P$$

$$\Rightarrow \quad A = PDP^T = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{u}_1^T \rightarrow \\ \vec{u}_2^T \rightarrow \\ \vdots \\ \vec{u}_n^T \rightarrow \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \vec{u}_1 & \lambda_2 \vec{u}_2 & \cdots & \lambda_n \vec{u}_n \end{bmatrix} \begin{bmatrix} \vec{u}_1^T \rightarrow \\ \vec{u}_2^T \rightarrow \\ \vdots \\ \vec{u}_n^T \rightarrow \end{bmatrix}$$

From previous Comput.

$$= \lambda_1 \vec{u}_1 \vec{u}_1^T + \lambda_2 \vec{u}_2 \vec{u}_2^T + \cdots + \lambda_n \vec{u}_n \vec{u}_n^T$$

Summarizing,

$$\boxed{A = \lambda_1 \vec{u}_1 \vec{u}_1^T + \lambda_2 \vec{u}_2 \vec{u}_2^T + \cdots + \lambda_n \vec{u}_n \vec{u}_n^T} \quad (16.1)$$

This expression is called Spectral Decomposition of A

In our previous exercise,

$$\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} = 0 \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} + 3 \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} +$$

$$+ 3 \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix} \begin{bmatrix} -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \end{bmatrix}$$

$$= 3 \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} + 3 \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & -\frac{2}{6} \\ \frac{1}{6} & \frac{1}{6} & -\frac{2}{6} \\ -\frac{2}{6} & -\frac{2}{6} & \frac{4}{6} \end{bmatrix}$$

Geometric Interpretation of a Spectral Decomposition.

It can be proved that by multiplying

$$\hat{u}\,\hat{u}^T \vec{x} = \text{proj}_{\text{Span}\{\hat{u}\}} \vec{x}$$

Therefore, the spectral decomposition (16.1) applied to a vector $\vec{x}$ can be obtained by projecting $\vec{x}$ orthogonally on the lines determined by the eigenvectors of A, then multiplying these projections by the eigenvalues, and finally adding the scaled projections.

SHOW Figure 7.2.1 for example 2 (book)

$$A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$$

Eigenvalues

$$\lambda_1 = -3 \rightarrow \vec{u}_1 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

$$\lambda_2 = 2 \rightarrow \vec{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

orthonormal set of eigenvectors

For example

$$A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \left(\lambda_1 \vec{u}_1 \vec{u}_1^T + \lambda_2 \vec{u}_2 \vec{u}_2^T\right) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix} = -3\begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{4}{5} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

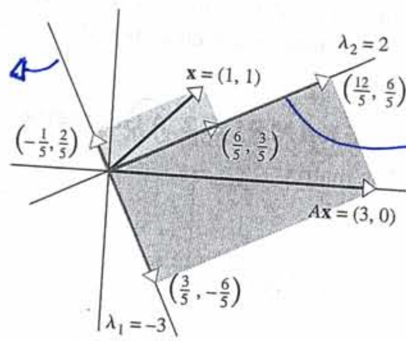
$$= (-3)\begin{bmatrix} -1/5 \\ 2/5 \end{bmatrix} + (2)\begin{bmatrix} 6/5 \\ 3/5 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 6/5 \end{bmatrix} + \begin{bmatrix} 12/5 \\ 6/5 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

Notice that

$$\text{proj}_{(1/\sqrt{5},\,-2/\sqrt{5})} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \left[(1,1)\cdot\left(1/\sqrt{5}, -2/\sqrt{5}\right)\right]\left(1/\sqrt{5}, -2/\sqrt{5}\right) = \frac{-1}{\sqrt{5}}\left(1/\sqrt{5}, -2/\sqrt{5}\right) = \left(-\frac{1}{5}, \frac{2}{5}\right)$$

Formulas (9) and (10) provide two different ways of viewing the image of the vector (1, 1) under multiplication by A: Formula (9) tells us directly that the image of this vector is (3, 0), whereas Formula (10) tells us that this image can also be obtained by projecting (1, 1) onto the eigenspaces corresponding to $\lambda_1 = -3$ and $\lambda_2 = 2$ to obtain the vectors $\left(-\frac{1}{5}, \frac{2}{5}\right)$ and $\left(\frac{6}{5}, \frac{3}{5}\right)$, then scaling by the eigenvalues to obtain $\left(\frac{3}{5}, -\frac{6}{5}\right)$ and $\left(\frac{12}{5}, \frac{6}{5}\right)$, and then adding these vectors (see Figure 7.2.1). ◀

direction of unit vector $(1/\sqrt{5}, -2/\sqrt{5})$



direction of unit vector $\left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$

▶ Figure 7.2.1

The Nondiagonalizable Case

If A is an $n \times n$ matrix that is not orthogonally diagonalizable, it may still be possible to achieve considerable simplification in the form of P^TAP by choosing the orthogonal matrix P appropriately. We will consider two theorems (without proof) that illustrate this. The first, due to the German mathematician Issai Schur, states that every square matrix A is orthogonally similar to an upper triangular matrix that has the eigenvalues of A on the main diagonal.

THEOREM 7.2.3 Schur's Theorem

If A is an $n \times n$ matrix with real entries and real eigenvalues, then there is an orthogonal matrix P such that P^TAP is an upper triangular matrix of the form

$$P^TAP = \begin{bmatrix} \lambda_1 & \times & \times & \cdots & \times \\ 0 & \lambda_2 & \times & \cdots & \times \\ 0 & 0 & \lambda_3 & \cdots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{11}$$

in which $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of the matrix A repeated according to multiplicity.

Issai Schur (1875–1941)

Historical Note The life of the German mathematician Issai Schur is a sad reminder of the effect that Nazi policies had on Jewish intellectuals during the 1930s. Schur was a brilliant mathematician and a popular lecturer who attracted many students and researchers to the University of Berlin, where he worked and taught. His lectures sometimes attracted so many students that opera glasses were needed to see him from the back row. Schur's life became increasingly difficult under Nazi rule, and in April of 1933 he was forced to "retire" from the university under a law that prohibited non-Aryans from holding "civil service" positions. There was an outcry from many of his students and colleagues who respected and liked him, but it did not stave off his complete dismissal in 1935. Schur, who thought of himself as a loyal German never understood the persecution and humiliation he received at Nazi hands. He left Germany for Palestine in 1939, a broken man. Lacking in financial resources, he had to sell his beloved mathematics books and lived in poverty until his death in 1941.

[Image: Courtesy Electronic Publishing Services, Inc., New York City]

It is common to denote the upper triangular matrix in (11) by S (for Schur), in which case that equation can be rewritten as

$$A = PSP^T \tag{12}$$

which is called a ***Schur decomposition*** of A.

The next theorem, due to the German mathematician and engineer Karl Hessenberg (1904–1959), states that every square matrix with real entries is orthogonally similar to a matrix in which each entry below the first ***subdiagonal*** is zero (Figure 7.2.2). Such a matrix is said to be in ***upper Hessenberg form.***

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}$$

First subdiagonal

▲ **Figure 7.2.2**

THEOREM 7.2.4 Hessenberg's Theorem

If A is an $n \times n$ matrix, then there is an orthogonal matrix P such that P^TAP is a matrix of the form

$$P^TAP = \begin{bmatrix} \times & \times & \cdots & \times & \times & \times \\ \times & \times & \cdots & \times & \times & \times \\ 0 & \times & \cdots & \times & \times & \times \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \times & \times & \times \\ 0 & 0 & \cdots & 0 & \times & \times \end{bmatrix} \tag{13}$$

Note that unlike those in (11), the diagonal entries in (13) are usually *not* the eigenvalues of A.

It is common to denote the upper Hessenberg matrix in (13) by H (for Hessenberg), in which case that equation can be rewritten as

$$A = PHP^T \tag{14}$$

which is called an ***upper Hessenberg decomposition*** of A.

Remark In many numerical algorithms the initial matrix is first converted to upper Hessenberg form to reduce the amount of computation in subsequent parts of the algorithm. Many computer packages have built-in commands for finding Schur and Hessenberg decompositions.

Concept Review

- Orthogonally similar matrices
- Orthogonally diagonalizable matrix
- Spectral decomposition (or eigenvalue decomposition)
- Schur decomposition
- Subdiagonal
- Upper Hessenburg form
- Upper Hessenburg decomposition

Skills

- Be able to recognize an orthogonally diagonalizable matrix.
- Know that eigenvalues of symmetric matrices are real numbers.
- Know that for a symmetric matrix eigenvectors from different eigenspaces are orthogonal.
- Be able to orthogonally diagonalize a symmetric matrix.
- Be able to find the spectral decomposition of a symmetric matrix.
- Know the statement of Schur's Theorem.
- Know the statement of Hessenburg's Theorem.