Diagonalization

A $n\times n$ matrix, P nonsingular $n\times n$. Then $P^{-1}AP$ is called a <u>similarity transformation</u> of A, and the resulting matrix $\hat{A}=P^{-1}AP$ is said to be <u>similar</u> to A.

<u>Theorem</u>. If $\hat{A}$ is similar to A, then $\hat{A}$ and A have the same eigenvalues.

The proof is easy:

$$\det(A-\lambda I)=0 \Rightarrow \det(P^{-1}(A-\lambda I)P)=0$$

$$\Rightarrow \det(P^{-1}AP-\lambda P^{-1}IP)=0$$

$$\Rightarrow \det(\hat{A}-\lambda I)=0.$$

Theorem. If $x_1, \ldots, x_k$ are eigenvectors corresponding to different eigenvalues $\lambda_1, \ldots, \lambda_k$ then $x_1, \ldots, x_k$ are linearly independent.

The case where all λ_j's are different, i.e. if $n \times n$ matrix A has n different eigenvalues $\lambda_1, \ldots, \lambda_n$ then there is a set of eigenvectors $x_1, \ldots, x_n$ which are linearly independent.

The same is true for unitary and Hermitian matrices even if there are repeated roots.

This means that for unitary and Hermitian matrices, for every λ_j

geometric multiplicity = algebraic mult.

Even more is true for unitary and Hermitian matrices.

If A is Hermitian, then eigenvectors corresponding to different eigenvectors are orthogonal. Using Gram-Schmidt on the bases for eigenspaces corresponding to multiple eigenvalues we get a simple algorithm for finding an orthonormal set of eigenvectors $x_1, \ldots, x_n$ for an $n \times n$

1. Compute eigenvalues for A $\lambda_1, \ldots, \lambda_k$
2. Find eigenvectors corresponding to those λ_j having algebraic multiplicity 1 and take $u_j = \dfrac{x_j}{\|x_j\|}$.
3. For an eigenvalue λ of algebraic multiplicity $r > 1$, take $x_1, \ldots, x_r$ as a basis for the

corresponding eigenspace (remember geometric mult = algebraic mult for Hermitian A). Perform Gram-Schmidt on $x_1, \dots, x_r$ to obtain orthonormal vectors $v_1, \dots, v_r$ which also span the eigenspace.

4. Putting together the vectors u_j obtained in 2 with the v_j from 3 we get n orthonormal vectors.

The orthonormal set of eigenvectors is called a <u>unitary system</u>.

<u>Ex.</u> $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, $p(\lambda) = (1-\lambda)(\lambda^2-1)$

$\lambda = 1$ mult. 2 $\qquad\qquad$ $\lambda = -1$ mult 1.

For $\lambda = 1$, $A - \lambda I = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}$ so in echelon form we have

$$\begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$ with basis for soln. space $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

For $\lambda = -1$, $A - \lambda I = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ in echelon form is

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$ so a basis is $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

The unitary system of eigenvectors is then

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

Now we are going to consider the notion of <u>diagonalizing</u> a matrix

· A matrix A is <u>diagonalizable</u> if

There is a nonsingular P such that

$$P^{-1}AP = D \quad \text{diagonal.}$$

Then, since similarity transformations do not change eigenvalues, and diagonal matrices have their eigenvalues on the diagonal,

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 \cdots & 0 \\ & & \ddots & 0 \\ & 0 & & \lambda_n \end{pmatrix} \qquad \lambda_1, \dots, \lambda_n \text{ eigenvalues}$$

where the λ_j's on diagonal are repeated as many times as their algebraic multiplicity.

Not all matrices are diagonalizable,

e.g. $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is not (as we shall see in a minute)

To understand which matrices are diagonalizable we must understand the relation between P and the eigen-vectors of A. In fact

The <u>columns</u> <u>of</u> <u>P</u> <u>are</u> <u>eigenvectors</u> of A.

Think about it: Let the columns of A be denoted by $v_1, \ldots, v_n$. Then,

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

$$\Rightarrow \quad AP = PD$$

or

$$A\begin{pmatrix} v_1 & v_2 & \cdots & v_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix} = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

Look at the first column of left and right hand side:

$$Av_1 = \lambda_1 v_2 \ ;$$

second column is

$$Av_2 = \lambda_2 v_2 \ ; \qquad \text{etc.}$$

With this observation we see that an $n \times n$ matrix A is diagonalizable $\Longleftrightarrow$ A has n linearly independent eigenvectors.

If A has n linearly independent eigenvectors then A is called <u>nondefective</u>. Otherwise it is <u>defective</u>. From lesson 7 we have that

$$B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

are <u>defective</u> and thus, not diagonalizable.

If the algebraic multiplicities of all the eigenvalues are 1, the the matrix is nondefective, and hence diagonalizable. However, it can happen that a matrix might be that a matrix is nondefective even though some eigenvalues have higher multiplicity. For instance, Hermitian and unitary matrices are always nondefective and therefore diagonalizable even though they may have repeated eigenvalues.

Ex. $A = \begin{pmatrix} 0 & 2 \\ -1 & 2 \end{pmatrix}$. From lesson 7,

$\lambda_1 = 1+i$ has eigenvector $\begin{pmatrix} 1-i \\ 1 \end{pmatrix}$

$\lambda_2 = 1-i$ " " $\begin{pmatrix} 1+i \\ 1 \end{pmatrix}$.

Take $P = \begin{pmatrix} 1-i & 1+i \\ 1 & 1 \end{pmatrix}$. Find P^{-1}

$$\left(\begin{array}{cc|cc} 1-i & 1+i & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & \frac{1+i}{1-i} & \frac{1}{1-i} & 0 \\ 1 & 1 & 0 & 1 \end{array}\right)$$

$$= \left(\begin{array}{cc|cc} 1 & i & \frac{1+i}{2} & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & i & \frac{1+i}{2} & 0 \\ 0 & 1-i & -\frac{(1+i)}{2} & 1 \end{array}\right)$$

$$\to \left(\begin{array}{cc|cc} 1 & i & \frac{1+i}{2} & 0 \\ 0 & 1 & -\frac{i}{2} & \frac{1+i}{2} \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 0 & i/2 & \frac{1-i}{2} \\ 0 & 1 & -i/2 & \frac{1+i}{2} \end{array}\right)$$

$$\begin{pmatrix} i/2 & (1-i)/2 \\ -i/2 & (1+i)/2 \end{pmatrix} A \begin{pmatrix} 1-i & 1+i \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}$$

Let $Q = x^T A x$, A symmetric.

Suppose $P^{-1}AP = D$. Let $x = Py$

<u>and</u> P orthogonal

$(\text{so } P^{-1} = P^T)$

With the variable y, Q becomes

$$Q = y^T P^T A P y = y^T P^{-1} A P y = y^T D y.$$

If we write $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ and

$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ we have

$Q = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$. This shows that for Q to be positive definite, it is sufficient that all λ_j's be positive.

To show it is necessary, let j be any integer from 1 to n, and calculate Q for the vector $\begin{array}{cl} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} & \leftarrow j^{\text{th}} \text{ row.} \end{array}$

For this vector we get

$$0 < Q = 0 + \cdot 0 + \lambda_j + 0 \cdots + 0$$

so $\lambda_j > 0$. Therefore, the condition that all λ_j's are positive is also necessary.

Appendix: <u>Gram</u> <u>Schmidt</u>

$v_1, \dots, v_k$ linearly independent

$$u_1 = \frac{v_1}{\|v_1\|}$$

$$v_2' = v_2 - (v_2, u_1)\, u_1 , \quad u_2 = \frac{v_2'}{\|v_2'\|}$$

$$v_3' = v_3 - (v_3, u_1)\, u_1 - (v_3, u_2)\, u_2 , \quad u_3 = \frac{v_3'}{\|v_3'\|}$$

etc.