Orthogonal diagonalizability

Symmetric matrices can be diagonalized by an orthogonal matrix

Orthogonal similarity

- If P is orthogonal, then $P^{-1}=P^T$.

Definition 8.3.1 If A and C are square matrices with the same size, then we say that ***C is orthogonally similar to A*** if there exists an orthogonal matrix P such that $C = P^TAP$.

Theorem 8.3.2 *Two matrices are orthogonally similar if and only if there exist orthonormal bases with respect to which the matrices represent the same linear operator.*

The Orthogonal Diagonalization Problem Given a square matrix A, does there exist an orthogonal matrix P for which P^TAP is a diagonal matrix, and if so, how does one find such a P? If such a matrix P exists, then A is said to be ***orthogonally diagonalizable***, and P is said to ***orthogonally diagonalize*** A.

- A must be symmetric so that A is orthogonally diagonalizable.
- $D=P^TAP$. $A=PDP^T$.
- $A^T=PD^TP^T=A$ since $D^T=D$.

Theorem 8.3.3 *An $n \times n$ matrix A is orthogonally diagonalizable if and only if there exists an orthonormal set of n eigenvectors of A.*

- Proof: ->) Let P=[p_1,p_2,...,p_n] be the orthogonal matrix. Then {p_1,p_2,..,p_n} is an orthonormal basis and are eigenvectors of A.
- (<-): Given an orthonormal set of eigenvectors, we can form P. P diagonalizes A.

Theorem 8.3.4

(*a*) *A matrix is orthogonally diagonalizable if and only if it is symmetric.*

(*b*) *If A is a symmetric matrix, then eigenvectors from different eigenspaces are orthogonal.*

- Proof: (a) -> done
- (a) <-. (not prove.)
- (b) $l_1v_1.v_2 = (l_1v_1)^Tv_2=(Av_1)^Tv_2=v_1^TA^Tv_2=v_1^TAv_2$
- $=v_1^Tl_2v_2=l_2v_1.v_2$. Since l_1 is different from l_2, this means $v_1.v_2=0$.

A method of diagonalization

Orthogonally Diagonalizing an $n \times n$ Symmetric Matrix

Step 1. Find a basis for each eigenspace of A.

Step 2. Apply the Gram–Schmidt process to each of these bases to produce orthonormal bases for the eigenspaces.

Step 3. Form the matrix $P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n]$ whose columns are the vectors constructed in Step 2. The matrix P will orthogonally diagonalize A, and the eigenvalues on the diagonal of $D = P^T AP$ will be in the same order as their corresponding eigenvectors in P.

- Example 1. From any symmetric matrix, we find eigenvalues and eigenvectors and this will work.

Spectral decompositions

- A symmetric. Orthogonally diagonalizable.
- P=[u_1,u_2,..,u_n] u_i eigenvector corr to l_i.
- $D=P^TAP$.
- Then $A=PDP^T=$

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 u_1 & \lambda_2 u_2 & \cdots & \lambda_n u_n \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix}$$

- $A=l_1u_1u_1^T+l_2u_2u_2^T+\dots+l_nu_nu_n^T$.
- This is called the spectral decomposition of A.
- $u_iu_i^T$ is a projection to Span$\{u_i\}$.
- There are eigenspaces...
- If A is not symmetric, the eigenspaces are not necessarily orthogonal.
- Example 2:

Power of a diagonalizable matrices

- Suppose that A is diagonalizable.
- Then $A=PDP^{-1}$.
- $A^2=PDP^{-1}PDP^{-1}=PD^2P^{-1}$
- $A^k = PDP^{-1}PDP^{-1}\ldots PDP^{-1}=PD^kP^{-1}$.
- If A is symmetric, and $A=l\_1u\_1u\_1^T+\ldots+l\_nu\_nu\_n^T$.
  - $A^k=l\_1^k u\_1u\_1^T+\ldots+l\_n^k u\_nu\_n^T$. (This follows by midle cancellations also.)
- Example 3:

Cayley-Hamilton theorem

Theorem 8.3.5 (*Cayley–Hamilton Theorem*) *Every square matrix satisfies its characteristic equation; that is, if A is an $n \times n$ matrix whose characteristic equation is*

$$\lambda^n + c_1\lambda^{n-1} + \cdots + c_n = 0$$

then

$$A^n + c_1 A^{n-1} + \cdots + c_n I = 0 \tag{14}$$

We $A(-A^{n-1}/c\_n-\ldots-c\_n-1I/c\_n)=I$. We can obtain the inverse of A.

- Example 4.

Exponent of a matrix

Theorem 8.3.6 *Suppose that A is an $n \times n$ diagonalizable matrix that is diagonalized by P and that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A corresponding to the successive column vectors of P. If f is a real-valued function whose Maclaurin series converges on some interval containing the eigenvalues of A, then*

$$f(A) = P \begin{bmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{bmatrix} P^{-1} \tag{21}$$

- Example 5:

- A symmetric with a spectral decomposition:
- A=l_1u_1u_1^T+l_2u_2u_2^T+…+l_nu_nu_n^T.
- f(A)
  =f(l_1)u_1u_1^T+f(l_2)u_2u_2^T+…+f(l_n)u_nu_n^T.

Diagonalization and linear systems

- $Ax=b$. A diagonalizable
- $P^{-1}AP=D$
- $x=Py$ substiture
- $APy=b$, $P^{-1}APy=P^{-1}b$.
- $Dy=P^{-1}b$. This is easy to solve.
- Such algorithms are time saving once we know the diagonalizability and P.

The nondiagonalizable case

- $A = PSP^T$, S upper triangular matrix.

Theorem 8.3.7 **(*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*

$$P^TAP = \begin{bmatrix} \lambda_1 & \times & \times & \cdots & \times \\ 0 & \lambda_2 & \times & \cdots & \times \\ 0 & 0 & \lambda_3 & \cdots & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{24}$$

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

- $A = PHP^T$.

Theorem 8.3.8 (***Hessenberg's Theorem***) *Every square matrix with real entries is orthogonally similar to a matrix in upper Hessenberg form; that is, 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*

$$P^TAP = \begin{bmatrix} \times & \times & \cdots & \times & \times & \times \\ \times & \times & \cdots & \times & \times & \times \\ 0 & \times & \ddots & \times & \times & \times \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \times & \times & \times \\ 0 & 0 & \cdots & 0 & \times & \times \end{bmatrix} \tag{26}$$

- There are algorithmically different.
- Hessenberg->Schur->Eigenvalues...