

Section 8.3 Diagonalization of Symmetric Matrices and QR factorization

First we should define a symmetric matrix.

An $n \times n$ matrix A is symmetric if $A = [a_{ij}]$ with $a_{ij} = a_{ji}$. In other words the like off diagonal elements are the same. This also means that when A is symmetric, $A = A^T$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}$$

Example $\begin{bmatrix} 1 & :) & :(\\ :) & 2 & :\backslash \\ :(& :\backslash & 3 \end{bmatrix}$ What is the transpose?

Something worth knowing. Let $\vec{x}, \vec{y} \in \mathbb{R}^n$ then $\vec{x} \cdot \vec{y} = (\vec{x})^T (\vec{y})$

Ex: Calculate $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ using the formula above.

We will use this notation to help us prove the following theorem.

Let A be a symmetric matrix, then eigenvectors corresponding to distinct eigen values are orthogonal.

This theorem is one of the reasons symmetric matrices are so great!

Section 8.3 Diagonalization of Symmetric Matrices and QR factorization

Let A be a symmetric matrix, then eigenvectors corresponding to distinct eigen values are orthogonal.

Proof: What we want to do is show that for eigenvectors $\vec{u}_1, \vec{u}_2$ such that $A\vec{u}_1 = \lambda_1\vec{u}_1$ and $A\vec{u}_2 = \lambda_2\vec{u}_2$ with $\lambda_1 \neq \lambda_2$ that their dot product $\vec{u}_1 \cdot \vec{u}_2 = 0$ note that $\vec{u}_1^T \vec{u}_2 = \vec{u}_2^T \vec{u}_1 = \vec{u}_1 \cdot \vec{u}_2$, we will obtain our dot products in this notation.

To do this we will be super clever and examine the quantity $(A\vec{u}_1)^T \vec{u}_2$ using the associative property to calculate it in two different ways keeping in mind that $(AB)^T = B^T A^T$.

So if we end up with n distinct eigen values then all of the eigen vectors of A will be orthongal. This says something really cool about the matrix P in the diagonalization of $A = PDP^{-1}$ provided we normalize the eigen vectors first.

Section 8.3 Diagonalization of Symmetric Matrices and QR factorization

Definition: A matrix P is called orthonormal when the column vectors are orthogonal and of length 1

If P is an orthogonal matrix then $P^{-1} = P^T$

Proof:

Let $P = [\vec{p}_1 \ \vec{p}_2 \ \dots \ \vec{p}_n]$ with all columns orthonormal, meaning $\vec{p}_i \cdot \vec{p}_j = 0$ and $||\vec{p}_i|| = 1$.

What do we get when we calculate PP^T and P^TP

$$P^TP = \begin{bmatrix} \vec{p}_1^T \\ \vec{p}_2^T \\ \vdots \\ \vec{p}_n^T \end{bmatrix} \begin{bmatrix} \vec{p}_1 & \vec{p}_2 & \dots & \vec{p}_n \end{bmatrix} = \begin{bmatrix} \vec{p}_1^T \vec{p}_1 & \vec{p}_1^T \vec{p}_2 & \dots & \vec{p}_1^T \vec{p}_n \\ \vec{p}_2^T \vec{p}_1 & \vec{p}_2^T \vec{p}_2 & \dots & \vec{p}_2^T \vec{p}_n \\ \vdots & \vdots & \ddots & \vdots \\ \vec{p}_n^T \vec{p}_1 & \vec{p}_n^T \vec{p}_2 & \dots & \vec{p}_n^T \vec{p}_n \end{bmatrix}$$

For PP^T we note that P^TP above is symmetric! So this means $(P^TP)^T = P^TP$ but from properties of the transpose we also have that $(P^TP)^T = P^T(P^T)^T = P^TP$ so we would get the exact same answer as above! Thus $P^T = P^{-1}$ when it is made of orthonormal columns.

Section 8.3 Diagonalization of Symmetric Matrices and QR factorization

A square matrix A is said to be orthogonally diagonalizable if there exists an orthogonal matrix P and diagonal matrix D such that $A = PDP^{-1}$

If a matrix A is orthogonally diagonalizable if and only if it is symmetric

Proof: we will only prove 1 way here, the other way is very involved!

Let A be orthogonally diagonalizable then there is an orthonormal matrix P and diagonal matrix D such that $A = PDP^{-1}$.

To show A is symmetric, we can just show that $A^T = A$

The above theorem is often called the "spectral theorem", but there are lots of other spectral theorems...

Section 8.3 Diagonalization of Symmetric Matrices and QR factorization

Example: Find an orthogonal diagonalization for $A = \begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix}$

Section 8.3 Diagonalization of Symmetric Matrices and QR factorization

If A is a real matrix, then AA^T has positive eigen values.

Proof:

Let $\vec{u}$ be an eigen vector of AA^T , that is $AA^T\vec{u} = \lambda\vec{u}$

Consider $\|A\vec{u}\|^2 = (A\vec{u}) \cdot (A\vec{u}) = (A\vec{u})^T (A\vec{u}) =$

Thus since $0 \leq \|A\vec{u}\|^2 = \lambda \|\vec{u}\|^2$ and $\|\vec{u}\|^2 \geq 0$ the only way this equation can hold is if $\lambda \geq 0$

QR factorization. Suppose A is an $n \times m$ real matrix with linearly independent columns. Then there exists an $n \times m$ matrix $Q = [\vec{q}_1 \ \vec{q}_2 \ \cdots \ \vec{q}_m]$ with orthonormal columns and $n \times m$ upper triangular matrix R such that

$$A = QR$$

Before we prove this just a word on why it is useful.... Suppose your matrix A is very large and you need to solve $A\vec{x} = \vec{b}$. One way would be to use $\vec{x} = A^{-1}\vec{b}$ however two horrible things can happen, 1. it is very computationally expensive to find A^{-1} and 2. The $\det(A)$ may be small meaning large blow ups in error due to numerical precision issues. If we instead look at the system $A\vec{x} = Q(R\vec{x}) = \vec{b}$ Then we can solve it as

$R\vec{x} = Q^T\vec{b}$ since $Q^{-1} = Q^T$ taking a transposition gives no loss in numerical accuracy and a machine does it very quickly. Then since R is upper triangular one can just back solve avoiding a lot of round off error! In applied settings we are always worried about these issues and clever algorithms like QR factorization end up being created so we can use linear algebra to it's full potential, plus its cool.

Section 8.3 Diagonalization of Symmetric Matrices and QR factorization

QR factorization. Suppose A is an $n \times m$ real matrix with linearly independent columns. Then there exists an $n \times m$ matrix $Q = [\vec{q}_1 \ \vec{q}_2 \ \cdots \ \vec{q}_m]$ with orthonormal columns and $n \times m$ upper triangular matrix R with positive entries on the diagonal such that

$$A = QR$$

Proof:

Since the columns of A are assumed to be linearly independent we know that we could use Gram-Schmidt to find an orthonormal basis for the same span. So let $\{\vec{q}_1 \dots \vec{q}_m\}$ be the orthonormal vectors obtained from the columns of A via the G-S algorithm.

Now define $Q = [\vec{q}_1 \ \vec{q}_2 \ \cdots \ \vec{q}_m]$

Now we employ this theorem from sec 8.1

Let $V = \{\vec{v}_1 \dots \vec{v}_k\}$ be an orthogonal basis for a subspace S , then for any vector $\vec{s} \in S$ we can write $\vec{s}$ as

$$\vec{s} = c_1 \vec{v}_1 + c_2 \vec{v}_2 + \cdots + c_k \vec{v}_k \text{ where } c_i = \frac{\vec{v}_i \cdot \vec{s}}{\|\vec{v}_i\|^2}$$

And write each vector of $A = [\vec{a}_1 \ \vec{a}_2 \ \cdots \ \vec{a}_m]$ in the following form

$$\vec{a}_i = (\vec{q}_1 \cdot \vec{a}_i) \vec{q}_1 + (\vec{q}_2 \cdot \vec{a}_i) \vec{q}_2 + \cdots + (\vec{q}_m \cdot \vec{a}_i) \vec{q}_m$$

So we now have a representation of each column vector of A in terms of the $\vec{q}'$ s

Now define $R = [r_{ij}]$ where $r_{ij} = \vec{q}_i \cdot \vec{a}_j$ for i, j above or on the diagonal and $r_{ij} = 0$ below the diagonal. Then.

$$QR = [\vec{q}_1 \ \vec{q}_2 \ \cdots \ \vec{q}_m] \begin{bmatrix} \vec{q}_1 \cdot \vec{a}_1 & \vec{q}_1 \cdot \vec{a}_2 & \cdots & \vec{q}_1 \cdot \vec{a}_m \\ 0 & \vec{q}_2 \cdot \vec{a}_2 & \cdots & \vec{q}_2 \cdot \vec{a}_m \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & \vec{q}_m \cdot \vec{a}_m \end{bmatrix} = [\vec{a}_1 \ \vec{a}_2 \ \cdots \ \vec{a}_m]$$

Since the j th column of $QR = \vec{a}_j = (\vec{q}_1 \cdot \vec{a}_j) \vec{q}_1 + (\vec{q}_2 \cdot \vec{a}_j) \vec{q}_2 + \cdots + (\vec{q}_m \cdot \vec{a}_j) \vec{q}_m$

We can ensure that each diagonal entry is positive by changing $\vec{q}_j$ to $-\vec{q}_j$ in the original orthogonal set if we find a r_{jj} that is negative.

Section 8.3 Diagonalization of Symmetric Matrices and QR factorization

It seems like it would be a lot of work to find R , so many dot products! However once we have Q we can find R easily since $Q^T = Q^{-1}$

Let $A = QR$ then $Q^T A = Q^T QR = IR = R$ so $Q^T A = R$

Find the QR factorization for the matrix A . (Enter `sqrt(n)` for $\sqrt{n}$.)

$$A = \begin{bmatrix} 0 & -5 \\ 8 & -5 \\ 10 & 4 \end{bmatrix}$$

8a

$$Q = \begin{bmatrix} \boxed{} & \boxed{} \\ \boxed{} & \boxed{} \\ \boxed{} & \boxed{} \end{bmatrix}$$

 `[0, -5/66*sqrt(66); 4/41*sqrt(41), -5/66*sqrt(66); 5/41*sqrt(41), 2/33*sqrt(66)]`

8b

$$R = \begin{bmatrix} \boxed{} & \boxed{} \\ \boxed{} & \boxed{} \end{bmatrix}$$

 `[2*sqrt(41), 0; 0, sqrt(66)]`