How Do We Really Find Eigenvalues and Why Should You Care?

Thomas R. Cameron

Department of Mathematics
Washington State University

March 12, 2015

Outline

Definition of Eigenvalues and Eigenvectors

Let $A \in \mathbb{C}^{n \times n}$. We say that $\lambda \in \mathbb{C}$ is an eigenvalue and $x \in \mathbb{C}^n \setminus \{0\}$ is a corresponding eigenvector, if

$$Ax = \lambda x.$$

Existence of Eigenvalues

- $\lambda \in \mathbb{C}$ is an eigenvalue of $A \in \mathbb{C}^{n\times n}$ if and only if $\lambda I - A$ is singular.
- if and only if $\det(\lambda I - A) = 0$.
- if and only if λ is a root of the characteristic polynomial $p(\lambda) = \det(\lambda I - A)$
- $p(\lambda)$ is a polynomial of degree n. By the fundamental theorem of algebra $p(\lambda)$ has n roots, counting multiplicities.

Existence of Eigenvalues

- $\lambda \in \mathbb{C}$ is an eigenvalue of $A \in \mathbb{C}^{n\times n}$ if and only if $\lambda I - A$ is singular.
- if and only if $\det(\lambda I - A) = 0$.
- if and only if λ is a root of the characteristic polynomial $p(\lambda) = \det(\lambda I - A)$
- $p(\lambda)$ is a polynomial of degree n. By the fundamental theorem of algebra $p(\lambda)$ has n roots, counting multiplicities.

Existence of Eigenvalues

- $\lambda \in \mathbb{C}$ is an eigenvalue of $A \in \mathbb{C}^{n\times n}$ if and only if $\lambda I - A$ is singular.
- if and only if $\det(\lambda I - A) = 0$.
- if and only if λ is a root of the characteristic polynomial $p(\lambda) = \det(\lambda I - A)$
- $p(\lambda)$ is a polynomial of degree n. By the fundamental theorem of algebra $p(\lambda)$ has n roots, counting multiplicities.

Existence of Eigenvalues

- $\lambda \in \mathbb{C}$ is an eigenvalue of $A \in \mathbb{C}^{n\times n}$ if and only if $\lambda I - A$ is singular.
- if and only if $\det(\lambda I - A) = 0$.
- if and only if λ is a root of the characteristic polynomial $p(\lambda) = \det(\lambda I - A)$
- $p(\lambda)$ is a polynomial of degree n. By the fundamental theorem of algebra $p(\lambda)$ has n roots, counting multiplicities.

Example

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

- The characteristic polynomial is $p(\lambda) = \lambda^2 - 2\lambda - 3$.
- The spectrum of A is $\sigma(A) = \{3, -1\}$

Example

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

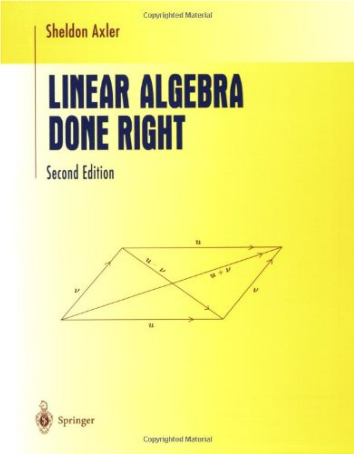
- The characteristic polynomial is $p(\lambda) = \lambda^2 - 2\lambda - 3$.
- The spectrum of A is $\sigma(A) = \{3, -1\}$

Example

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

- The characteristic polynomial is $p(\lambda) = \lambda^2 - 2\lambda - 3$.
- The spectrum of A is $\sigma(A) = \{3, -1\}$

Linear Algebra Done Right



No Determinants

- Linear algebra can be done better without determinants (Sheldon Axler, Down with Determinants!)
- I find it hard to conceive of a situation in which the numerical value of a determinant is needed (Henry Thacher, SIAM News, September 1988).

No Determinants

- Linear algebra can be done better without determinants (Sheldon Axler, Down with Determinants!)
- I find it hard to conceive of a situation in which the numerical value of a determinant is needed (Henry Thacher, SIAM News, September 1988).

Existence of Eigenvalues

Proof.

Let $x \in \mathbb{C}^n \setminus \{0\}$ and consider the Krylov sequence

$$x, Ax, A^2x, \ldots.$$

There exists a minimal integer $1 \leq k < n$ such that $\{x, Ax, \ldots, A^kx\}$ is a set of linearly dependent vectors. There exists constants $a_0, a_1, \ldots, a_k$ such that

$$a_0x + a_1Ax + \cdots + a_kA^kx = 0.$$

Define $p(z) = a_0 + a_1z + \cdots + a_kz^k = c(z - r_1)\cdots(z - r_m)$, then

$$c(A - r_1I)\cdots(A - r_mI)x = 0.$$

$\square$

Outline

Companion Matrix

Let $p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ be a monic scalar polynomial. We define the companion matrix of p to be the $n \times n$ matrix

$$A_p = \begin{bmatrix} -a_{n-1} & \cdots & -a_1 & -a_0 \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{bmatrix}.$$

Eigenvalues of Companion Matrix

Theorem

The eigenvalues of A_p are the roots of the polynomial $p(\lambda)$.

Proof.

The equation $A_p v = \lambda v$ is equivalent to

$$\begin{bmatrix} -a_{n-1}v_1 - \cdots - a_1 v_{n-1} - a_0 v_n \\ v_1 \\ \vdots \\ v_{n-1} \end{bmatrix} = \begin{bmatrix} \lambda v_1 \\ \lambda v_2 \\ \vdots \\ \lambda v_n \end{bmatrix}.$$

Therefore, $v_1 = \lambda^{n-1} v_n$, $v_2 = \lambda^{n-2} v_n, \ldots, v_{n-1} = \lambda v_n$. □

Summary

- The eigenvalues of a matrix A are the roots of the characteristic polynomial.
- The roots of a polynomial are the eigenvalues of a companion matrix.
- We cannot expect to solve the eigenvalue problem in a finite number of steps.

Summary

- The eigenvalues of a matrix A are the roots of the characteristic polynomial.
- The roots of a polynomial are the eigenvalues of a companion matrix.
- We cannot expect to solve the eigenvalue problem in a finite number of steps.

Summary

- The eigenvalues of a matrix A are the roots of the characteristic polynomial.
- The roots of a polynomial are the eigenvalues of a companion matrix.
- We cannot expect to solve the eigenvalue problem in a finite number of steps.

Outline

Definition of dominant eigenvalue

We say that the eigenvalue λ is dominant, if there exists a positive number $r < |\lambda|$ such that

$$\sigma(A) \setminus \{\lambda\} \subseteq \{z \in \mathbb{C} : |z| \leq r\}.$$

- If λ is dominant, then every eigenvector v associated with λ is called a dominant eigenvector.

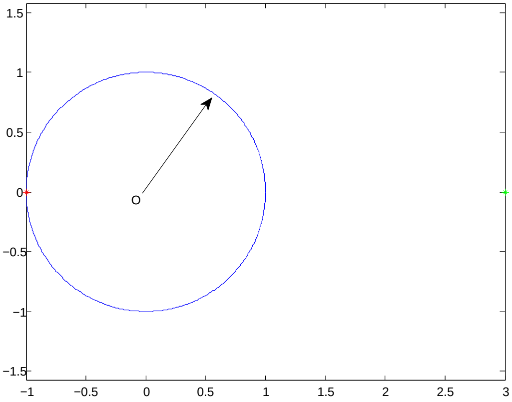
Definition of dominant eigenvalue

We say that the eigenvalue λ is dominant, if there exists a positive number $r < |\lambda|$ such that

$$\sigma(A) \setminus \{\lambda\} \subseteq \{z \in \mathbb{C} : |z| \leq r\}.$$

- If λ is dominant, then every eigenvector v associated with λ is called a dominant eigenvector.

Example



The Power Method

Given a nonzero vector $x \in \mathbb{C}^n$, the power method forms the Krylov sequence of vectors

$$x, Ax, A^2x, \ldots$$

- If x is not an unlucky choice and if A has a dominant eigenvector, then the Krylov sequence will converge to a dominant eigenvector of A.

The Power Method

Given a nonzero vector $x \in \mathbb{C}^n$, the power method forms the Krylov sequence of vectors

$$x, Ax, A^2x, \dots$$

- If x is not an unlucky choice and if A has a dominant eigenvector, then the Krylov sequence will converge to a dominant eigenvector of A.

Example

Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ and $x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

- Then we update via

$$x_{n+1} = \frac{Ax_n}{\|Ax_n\|}$$

where $\|\cdot\|$ is the vector 2 norm.

Example

Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ and $x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

- Then we update via

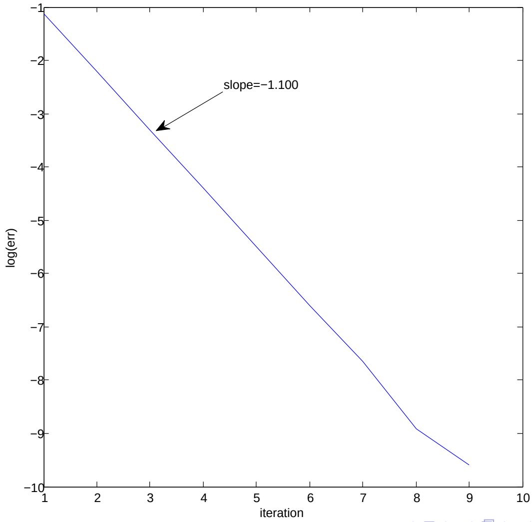
$$x_{n+1} = \frac{Ax_n}{\|Ax_n\|}$$

where $\|\cdot\|$ is the vector 2 norm.

Example

n	x_n	n	x_n
1	$\begin{bmatrix} 0.4472 \\ 0.8944 \end{bmatrix}$	6	$\begin{bmatrix} 0.7081 \\ 0.7061 \end{bmatrix}$
2	$\begin{bmatrix} 0.7809 \\ 0.6247 \end{bmatrix}$	7	$\begin{bmatrix} 0.7068 \\ 0.7074 \end{bmatrix}$
3	$\begin{bmatrix} 0.6805 \\ 0.7328 \end{bmatrix}$	8	$\begin{bmatrix} 0.7072 \\ 0.7070 \end{bmatrix}$
4	$\begin{bmatrix} 0.7158 \\ 0.6983 \end{bmatrix}$	9	$\begin{bmatrix} 0.7071 \\ 0.7071 \end{bmatrix}$
5	$\begin{bmatrix} 0.7042 \\ 0.7100 \end{bmatrix}$	10	$\begin{bmatrix} 0.7071 \\ 0.7071 \end{bmatrix}$

Example



Summary

- The power method generally converges to a dominant eigenvector of A.
- The rate of convergence is linear and has contraction factor $\frac{r}{|\lambda|}$.

Summary

- The power method generally converges to a dominant eigenvector of A.
- The rate of convergence is linear and has contraction factor $\frac{r}{|\lambda|}$.

Outline

Subspace Iteration

- When computing eigenvectors, our real object of interest are eigenspaces.
- If the eigenspace is one-dimensional, then a single eigenvector v is a basis for the space.
- We can view the power method as

$$S, AS, A^2S, \ldots$$

where $S = \text{span}\{x\}$, and the above sequence converges to $\text{span}\{v\}$.

Subspace Iteration

- When computing eigenvectors, our real object of interest are eigenspaces.
- If the eigenspace is one-dimensional, then a single eigenvector v is a basis for the space.
- We can view the power method as

$$S, AS, A^2S, \ldots$$

where $S = \text{span}\{x\}$, and the above sequence converges to $\text{span}\{v\}$.

Subspace Iteration

- When computing eigenvectors, our real object of interest are eigenspaces.
- If the eigenspace is one-dimensional, then a single eigenvector v is a basis for the space.
- We can view the power method as

$$S, AS, A^2S, \ldots$$

where $S = \operatorname{span}\{x\}$, and the above sequence converges to $\operatorname{span}\{v\}$.

Definition of Invariant Subspace

Let $A \in \mathbb{F}^{n \times n}$. A subspace S of $\mathbb{F}^n$ is invariant under A if $Ax \in S$ whenever $x \in S$. That is,

$$AS \subseteq S.$$

Eigenvectors form Invariant Subspaces

Let $v_1, \ldots, v_k$ be eigenvectors of A corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_k$.

- For $x \in \operatorname{span}\{v_1, \ldots, v_k\}$ we have $x = c_1 v_1 + \cdots + c_k v_k$.
- $Ax = c_1 A v_1 + \cdots + c_k A v_k = c_1 \lambda_1 v_1 + \cdots + c_k \lambda_k v_k \in \operatorname{span}\{v_1, \ldots, v_k\}$.
- If A is semisimple (A has n linearly independent eigenvectors), then every invariant subspace under A has this form.

Eigenvectors form Invariant Subspaces

Let $v_1, \ldots, v_k$ be eigenvectors of A corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_k$.

- For $x \in \text{span}\{v_1, \ldots, v_k\}$ we have $x = c_1 v_1 + \cdots + c_k v_k$.
- $Ax = c_1 A v_1 + \cdots + c_k A v_k = c_1 \lambda_1 v_1 + \cdots + c_k \lambda_k v_k \in \text{span}\{v_1, \ldots, v_k\}$.
- If A is semisimple (A has n linearly independent eigenvectors), then every invariant subspace under A has this form.

Eigenvectors form Invariant Subspaces

Let $v_1, \ldots, v_k$ be eigenvectors of A corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_k$.

- For $x \in \text{span}\{v_1, \ldots, v_k\}$ we have $x = c_1 v_1 + \cdots + c_k v_k$.
- $Ax = c_1 A v_1 + \cdots + c_k A v_k = c_1 \lambda_1 v_1 + \cdots + c_k \lambda_k v_k \in \text{span}\{v_1, \ldots, v_k\}$.
- If A is semisimple (A has n linearly independent eigenvectors), then every invariant subspace under A has this form.

Eigenvectors form Invariant Subspaces

Let $v_1, \ldots, v_k$ be eigenvectors of A corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_k$.

- For $x \in \operatorname{span}\{v_1, \ldots, v_k\}$ we have $x = c_1 v_1 + \cdots + c_k v_k$.
- $Ax = c_1 A v_1 + \cdots + c_k A v_k = c_1 \lambda_1 v_1 + \cdots + c_k \lambda_k v_k \in \operatorname{span}\{v_1, \ldots, v_k\}$.
- If A is semisimple (A has n linearly independent eigenvectors), then every invariant subspace under A has this form.

Example

Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ and $S = \operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

- Note that S is not spanned by eigenvectors of A.
- For $x \in S$, we have $x = \begin{bmatrix} 0 \\ a \\ b \end{bmatrix}$.
- $Ax = \begin{bmatrix} 0 \\ 2a+b \\ 2b \end{bmatrix} \in S$.

Example

Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ and $S = \operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

- Note that S is not spanned by eigenvectors of A.
- For $x \in S$, we have $x = \begin{bmatrix} 0 \\ a \\ b \end{bmatrix}$.
- $Ax = \begin{bmatrix} 0 \\ 2a+b \\ 2b \end{bmatrix} \in S$.

Example

Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ and $S = \operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

- Note that S is not spanned by eigenvectors of A.
- For $x \in S$, we have $x = \begin{bmatrix} 0 \\ a \\ b \end{bmatrix}$.
- $Ax = \begin{bmatrix} 0 \\ 2a+b \\ 2b \end{bmatrix} \in S$.

Example

Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ and $S = \operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

- Note that S is not spanned by eigenvectors of A.
- For $x \in S$, we have $x = \begin{bmatrix} 0 \\ a \\ b \end{bmatrix}$.
- $Ax = \begin{bmatrix} 0 \\ 2a + b \\ 2b \end{bmatrix} \in S$.

Definition of Dominant Invariant Subspace

A $k-$dimensional invariant subspace S is dominant if the eigenvalues $\lambda_1, \ldots, \lambda_k$ of the restricted operator $A\,|_S$ dominate the spectrum of A.

- That is, there exists an r such that $r < |\lambda_j|$ for $j = 1, \ldots, k$ and the rest of the spectrum lies within the disc $\{z \in \mathbb{C} : |z| \leq r\}$.

Definition of Dominant Invariant Subspace

A $k-$dimensional invariant subspace S is dominant if the eigenvalues $\lambda_1, \ldots, \lambda_k$ of the restricted operator $A\,|_S$ dominate the spectrum of A.

- That is, there exists an r such that $r < |\lambda_j|$ for $j = 1, \ldots, k$ and the rest of the spectrum lies within the disc $\{z \in \mathbb{C} : |z| \leq r\}$.

Subspace Iteration

Let S be a $k-$dimensional space, and not an unlucky choice. Then the sequence

$$S, AS, A^2S, \ldots$$

will converge to a $k-$dimensional dominant invariant subspace under A, if A has one.

Deflating the Problem

Theorem

Let $S = span\{x_1, \ldots, x_k\}$ *be invariant under* $A \in \mathbb{F}^{n\times n}$ *and*

$$\{x_1, \ldots, x_k, x_{k+1}, \ldots, x_n\}$$

be a basis for $\mathbb{F}^n$*. Define* $X = [x_1, \ldots, x_n]$ *and* $B = X^{-1}AX$*, then*

$$B = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}.$$

Deflating the Problem

Proof.

The equation $B = X^{-1}AX$ is equivalent to $AX = XB$. The j^{th} column of this equation is equivalent to

$$Ax_j = \sum_{i=1}^{n} x_i b_{ij}.$$

If S is invariant under A, then $Ax_j \in \operatorname{span}\{x_1, \ldots, x_k\}$ for $j = 1, \ldots, k$. Therefore,

$$b_{ij} = 0$$

for $1 \leq j \leq k$ and $k+1 \leq i \leq n$. $\square$

Simultaneous Iteration

For $i = 0, 1, 2, \ldots$, let $S^{(i)=} A^i S$.

- If $\left\{q_1^{(i)}, \ldots, q_k^{(i)}\right\}$ is a basis for $S^{(i)}$, then $\left\{Aq_1^{(i)}, \ldots, Aq_k^{(i)}\right\}$ is a basis for $S^{(i+1)}$.
- In practice we specify that $\left\{q_1^{(i)}, \ldots, q_k^{(i)}\right\}$ be orthonormal.
- We can obtain an orthonormal basis $\left\{q_1^{(i+1)}, \ldots, q_k^{(i+1)}\right\}$ for $S^{(i+1)}$ by applying the Gram-Schmidt process to $\left\{Aq_1^{(i)}, \ldots, Aq_k^{(i)}\right\}$.

Simultaneous Iteration

For $i = 0, 1, 2, \ldots$, let $S^{(i)} = A^i S$.

- If $\left\{q_1^{(i)}, \ldots, q_k^{(i)}\right\}$ is a basis for $S^{(i)}$, then $\left\{Aq_1^{(i)}, \ldots, Aq_k^{(i)}\right\}$ is a basis for $S^{(i+1)}$.
- In practice we specify that $\left\{q_1^{(i)}, \ldots, q_k^{(i)}\right\}$ be orthonormal.
- We can obtain an orthonormal basis $\left\{q_1^{(i+1)}, \ldots, q_k^{(i+1)}\right\}$ for $S^{(i+1)}$ by applying the Gram-Schmidt process to $\left\{Aq_1^{(i)}, \ldots, Aq_k^{(i)}\right\}$.

Simultaneous Iteration

For $i = 0, 1, 2, \ldots$, let $S^{(i)=} A^i S$.

- If $\left\{q_1^{(i)}, \ldots, q_k^{(i)}\right\}$ is a basis for $S^{(i)}$, then $\left\{Aq_1^{(i)}, \ldots, Aq_k^{(i)}\right\}$ is a basis for $S^{(i+1)}$.
- In practice we specify that $\left\{q_1^{(i)}, \ldots, q_k^{(i)}\right\}$ be orthonormal.
- We can obtain an orthonormal basis $\left\{q_1^{(i+1)}, \ldots, q_k^{(i+1)}\right\}$ for $S^{(i+1)}$ by applying the Gram-Schmidt process to $\left\{Aq_1^{(i)}, \ldots, Aq_k^{(i)}\right\}$.

Simultaneous Iteration

For $i = 0, 1, 2, \ldots$, let $S^{(i)=}A^i S$.

- If $\left\{q_1^{(i)}, \ldots, q_k^{(i)}\right\}$ is a basis for $S^{(i)}$, then $\left\{Aq_1^{(i)}, \ldots, Aq_k^{(i)}\right\}$ is a basis for $S^{(i+1)}$.
- In practice we specify that $\left\{q_1^{(i)}, \ldots, q_k^{(i)}\right\}$ be orthonormal.
- We can obtain an orthonormal basis $\left\{q_1^{(i+1)}, \ldots, q_k^{(i+1)}\right\}$ for $S^{(i+1)}$ by applying the Gram-Schmidt process to $\left\{Aq_1^{(i)}, \ldots, Aq_k^{(i)}\right\}$.

Important Benefits

- The Gram-Schmidt process preserves all lower dimensional subspaces

$$\operatorname{span}\left\{q_1^{(i+1)}, \ldots, q_j^{(i+1)}\right\} = \operatorname{span}\left\{Aq_1^{(i)}, \ldots, Aq_j^{(i)}\right\}$$

  for $j = 1, \ldots, k$.

- We are simultaneously performing subspace iteration on subspaces of dimension $1, \ldots, k$.
- We are always looking for the opportunity to break off a smaller problem of size $j \times j$, for $j = 1, \ldots, k$.

Important Benefits

- The Gram-Schmidt process preserves all lower dimensional subspaces

$$\text{span}\left\{q_1^{(i+1)}, \ldots, q_j^{(i+1)}\right\} = \text{span}\left\{Aq_1^{(i)}, \ldots, Aq_j^{(i)}\right\}$$

  for $j = 1, \ldots, k$.
- We are simultaneously performing subspace iteration on subspaces of dimension $1, \ldots, k$.
- We are always looking for the opportunity to break off a smaller problem of size $j \times j$, for $j = 1, \ldots, k$.

Important Benefits

- The Gram-Schmidt process preserves all lower dimensional subspaces

$$\operatorname{span}\left\{q_1^{(i+1)}, \ldots, q_j^{(i+1)}\right\} = \operatorname{span}\left\{Aq_1^{(i)}, \ldots, Aq_j^{(i)}\right\}$$

  for $j = 1, \ldots, k$.
- We are simultaneously performing subspace iteration on subspaces of dimension $1, \ldots, k$.
- We are always looking for the opportunity to break off a smaller problem of size $j \times j$, for $j = 1, \ldots, k$.

Outline

Deflating the Problem

Theorem

Let $S = span\{q_1, \ldots, q_k\}$ be invariant under $A \in \mathbb{F}^{n\times n}$ and

$$\{q_1, \ldots, q_k, q_{k+1}, \ldots, q_n\}$$

be an orthonormal basis for $\mathbb{F}^n$. Define $Q = [q_1, \ldots, q_n]$ and $B = Q^*AQ$, then

$$B = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}.$$

Change of Coordinate System

Consider a step of simultaneous iteration with the standard basis vectors $e_1, e_2, \ldots, e_n$.

- We get $Ae_1, Ae_2, \ldots, Ae_n$, and then orthonormalize to get $q_1, q_2, \ldots, q_n$.
- We perform the similarity transformation $\hat{A} = Q^*AQ$, where $Q = [q_1, q_2, \ldots, q_n]$.
- The vectors $q_1, q_2, \ldots, q_n$ in the old coordinate system are $e_1, e_2, \ldots, e_n$ in the new coordinate system.

Change of Coordinate System

Consider a step of simultaneous iteration with the standard basis vectors $e_1, e_2, \ldots, e_n$.

- We get $Ae_1, Ae_2, \ldots, Ae_n$, and then orthonormalize to get $q_1, q_2, \ldots, q_n$.
- We perform the similarity transformation $\hat{A} = Q^*AQ$, where $Q = [q_1, q_2, \ldots, q_n]$.
- The vectors $q_1, q_2, \ldots, q_n$ in the old coordinate system are $e_1, e_2, \ldots, e_n$ in the new coordinate system.

Change of Coordinate System

Consider a step of simultaneous iteration with the standard basis vectors $e_1, e_2, \ldots, e_n$.

- We get $Ae_1, Ae_2, \ldots, Ae_n$, and then orthonormalize to get $q_1, q_2, \ldots, q_n$.
- We perform the similarity transformation $\hat{A} = Q^* AQ$, where $Q = [q_1, q_2, \ldots, q_n]$.
- The vectors $q_1, q_2, \ldots, q_n$ in the old coordinate system are $e_1, e_2, \ldots, e_n$ in the new coordinate system.

Change of Coordinate System

Consider a step of simultaneous iteration with the standard basis vectors $e_1, e_2, \dots, e_n$.

- We get $Ae_1, Ae_2, \dots, Ae_n$, and then orthonormalize to get $q_1, q_2, \dots, q_n$.
- We perform the similarity transformation $\hat{A} = Q^*AQ$, where $Q = [q_1, q_2, \dots, q_n]$.
- The vectors $q_1, q_2, \dots, q_n$ in the old coordinate system are $e_1, e_2, \dots, e_n$ in the new coordinate system.

QR Algorithm

Simultaneous iteration with a change of coordinate system

$$A_i = Q_{i+1}R_{i+1}, \quad A_{i+1} = Q_{i+1}^* A_i Q_{i+1},$$

for $i = 1, 2, \ldots$.

Example

i	A
0	$\begin{bmatrix} 4 & 2 \\ 3 & 3 \end{bmatrix}$
1	$\begin{bmatrix} 6.0400 & 0.2800 \\ -0.7200 & 0.9600 \end{bmatrix}$
2	$\begin{bmatrix} 6.0205 & -0.8832 \\ 0.1168 & 0.9795 \end{bmatrix}$
3	$\begin{bmatrix} 6.0038 & 0.9807 \\ -0.0193 & 0.9962 \end{bmatrix}$
4	$\begin{bmatrix} 6.0006 & -0.9968 \\ 0.0032 & 0.9994 \end{bmatrix}$

i	A
5	$\begin{bmatrix} 6.0001 & 0.9995 \\ -0.0005 & 0.9999 \end{bmatrix}$
6	$\begin{bmatrix} 6.0000 & -0.9999 \\ 0.0001 & 1.0000 \end{bmatrix}$
7	$\begin{bmatrix} 6.0000 & 1.0000 \\ -0.0000 & 1.0000 \end{bmatrix}$

Francis's Algorithm

- J.G. Francis. The QR transformation, part 1. Computer J., 4:265-272, 1961.
- J.G. Francis. The QR transformation, part 2. Computer J., 4:332-245, 1961.
- The Best of the 20th Century: Editors Name top 10 Algorithms. SIAM News, May 16, 2000.
- David S. Watkins. Francis's Algorithm. American Math Monthly, May 2011.

Francis's Algorithm

- J.G. Francis. The QR transformation, part 1. Computer J., 4:265-272, 1961.
- J.G. Francis. The QR transformation, part 2. Computer J., 4:332-245, 1961.
- The Best of the 20th Century: Editors Name top 10 Algorithms. SIAM News, May 16, 2000.
- David S. Watkins. Francis's Algorithm. American Math Monthly, May 2011.

Francis's Algorithm

- J.G. Francis. The QR transformation, part 1. Computer J., 4:265-272, 1961.
- J.G. Francis. The QR transformation, part 2. Computer J., 4:332-245, 1961.
- The Best of the 20th Century: Editors Name top 10 Algorithms. SIAM News, May 16, 2000.
- David S. Watkins. Francis's Algorithm. American Math Monthly, May 2011.

Francis's Algorithm

- J.G. Francis. The QR transformation, part 1. Computer J., 4:265-272, 1961.
- J.G. Francis. The QR transformation, part 2. Computer J., 4:332-245, 1961.
- The Best of the 20th Century: Editors Name top 10 Algorithms. SIAM News, May 16, 2000.
- David S. Watkins. Francis's Algorithm. American Math Monthly, May 2011.

Outline

The Problem of Sparsity

- A sparse matrix is one in which the vast majority of their entries are zero.
- We may be limited by storing constraints.
- Similarity transformations are a bad idea.

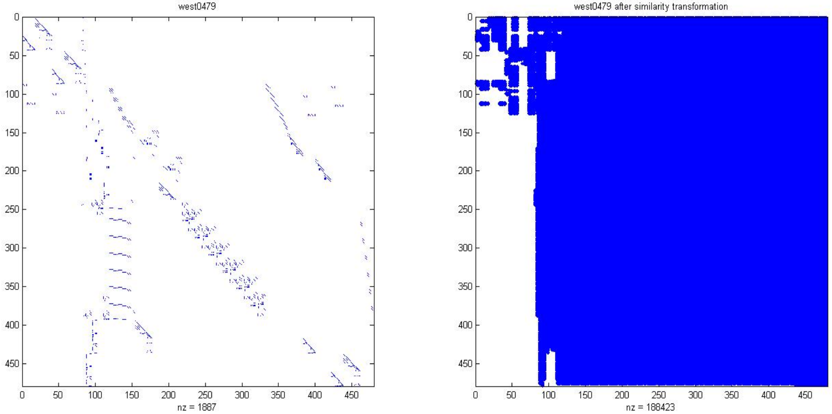
The Problem of Sparsity

- A sparse matrix is one in which the vast majority of their entries are zero.
- We may be limited by storing constraints.
- Similarity transformations are a bad idea.

The Problem of Sparsity

- A sparse matrix is one in which the vast majority of their entries are zero.
- We may be limited by storing constraints.
- Similarity transformations are a bad idea.

Example


west0479
nz = 1887
west0479 after similarity transformation
nz = 188423

The Idea behind Arnoldi's Process

Given $x \in \mathbb{C}^n \setminus \{0\}$, the power method forms the Krylov sequence of vectors

$$x, Ax, A^2x, \ldots.$$

- At the k^{th} iteration we only have $A^k x$.
- After k steps in the Arnoldi process we have

$$x, Ax, \ldots, A^k x.$$

- These vectors form a basis for the Krylov subspace

$$K_{k+1}(A, x) = \text{span}\left\{x, Ax, \ldots, A^k x\right\}.$$

The Idea behind Arnoldi's Process

Given $x \in \mathbb{C}^n \setminus \{0\}$, the power method forms the Krylov sequence of vectors

$$x, Ax, A^2x, \ldots.$$

- At the k^{th} iteration we only have $A^k x$.
- After k steps in the Arnoldi process we have

$$x, Ax, \ldots, A^k x.$$

- These vectors form a basis for the Krylov subspace

$$K_{k+1}(A, x) = \text{span}\left\{x, Ax, \ldots, A^k x\right\}.$$

The Idea behind Arnoldi's Process

Given $x \in \mathbb{C}^n \setminus \{0\}$, the power method forms the Krylov sequence of vectors

$$x, Ax, A^2x, \ldots.$$

- At the k^{th} iteration we only have A^kx.
- After k steps in the Arnoldi process we have

$$x, Ax, \ldots, A^kx.$$

- These vectors form a basis for the Krylov subspace

$$K_{k+1}(A,x) = \text{span}\left\{x, Ax, \ldots, A^kx\right\}.$$

The Idea behind Arnoldi's Process

Given $x \in \mathbb{C}^n \setminus \{0\}$, the power method forms the Krylov sequence of vectors

$$x, Ax, A^2x, \ldots.$$

- At the k^{th} iteration we only have A^kx.
- After k steps in the Arnoldi process we have

$$x, Ax, \ldots, A^kx.$$

- These vectors form a basis for the Krylov subspace

$$K_{k+1}(A,x) = \operatorname{span}\left\{x, Ax, \ldots, A^kx\right\}.$$

We want Orthonormal vectors

On the first step of the Arnoldi process we take $q_1 = \frac{x}{\|x\|}$.

- On the second step we take

$$\hat{q}_2 = Aq_1 - \langle Aq_1, q_1 \rangle q_1, \quad q_2 = \frac{\hat{q}_2}{\|\hat{q}_2\|}.$$

- On subsequent steps we take

$$\hat{q}_{k+1} = Aq_k - \sum_{j=1}^{k} q_j h_{jk}, \quad q_{k+1} = \frac{\hat{q}_{k+1}}{h_{k+1,k}},$$

where $h_{jk} = \langle Aq_k, q_j \rangle$ and $h_{k+1,k} = \|\hat{q}_{k+1}\|$.

We want Orthonormal vectors

On the first step of the Arnoldi process we take $q_1 = \frac{x}{\|x\|}$.

- On the second step we take

$$\hat{q}_2 = Aq_1 - \langle Aq_1, q_1 \rangle q_1, \quad q_2 = \frac{\hat{q}_2}{\|\hat{q}_2\|}.$$

- On subsequent steps we take

$$\hat{q}_{k+1} = Aq_k - \sum_{j=1}^{k} q_j h_{jk}, \quad q_{k+1} = \frac{\hat{q}_{k+1}}{h_{k+1,k}},$$

where $h_{jk} = \langle Aq_k, q_j \rangle$ and $h_{k+1,k} = \|\hat{q}_{k+1}\|$.

We want Orthonormal vectors

On the first step of the Arnoldi process we take $q_1 = \frac{x}{\|x\|}$.

- On the second step we take

$$\hat{q}_2 = Aq_1 - \langle Aq_1, q_1 \rangle q_1, \quad q_2 = \frac{\hat{q}_2}{\|\hat{q}_2\|}.$$

- On subsequent steps we take

$$\hat{q}_{k+1} = Aq_k - \sum_{j=1}^{k} q_j h_{jk}, \quad q_{k+1} = \frac{\hat{q}_{k+1}}{h_{k+1,k}},$$

where $h_{jk} = \langle Aq_k, q_j \rangle$ and $h_{k+1,k} = \|\hat{q}_{k+1}\|$.

Matrix Representation of Arnoldi Process

We can write the previous steps as

$$Aq_k = \sum_{j=1}^{k+1} q_j h_{jk}, \quad k = 1, 2, 3, \ldots$$

- We can write this as

$$A\left[q_1 \cdots q_m\right] = \left[q_1 \cdots q_m q_{m+1}\right] H_{m+1,m}$$

- Where

$$H_{m+1,m} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2m} \\ & h_{32} & & \vdots \\ & & \ddots & h_{mm} \\ & & & h_{m+1,m} \end{bmatrix}.$$

Matrix Representation of Arnoldi Process

We can write the previous steps as

$$Aq_k = \sum_{j=1}^{k+1} q_j h_{jk}, \quad k = 1, 2, 3, \ldots$$

- We can write this as

$$A[q_1 \cdots q_m] = [q_1 \cdots q_m q_{m+1}] \, H_{m+1,m}$$

- Where

$$H_{m+1,m} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2m} \\ & h_{32} & & \vdots \\ & & \ddots & h_{mm} \\ & & & h_{m+1,m} \end{bmatrix}.$$

Matrix Representation of Arnoldi Process

We can write the previous steps as

$$Aq_k = \sum_{j=1}^{k+1} q_j h_{jk}, \quad k = 1, 2, 3, \ldots$$

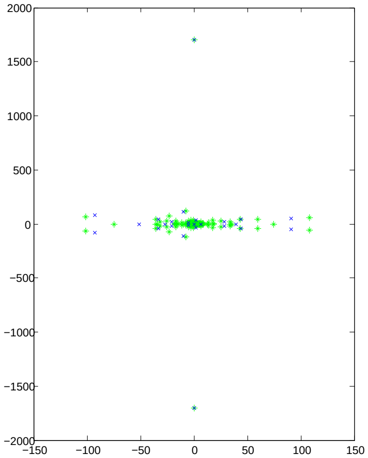
- We can write this as

$$A[q_1 \cdots q_m] = [q_1 \cdots q_m q_{m+1}]\, H_{m+1,m}$$

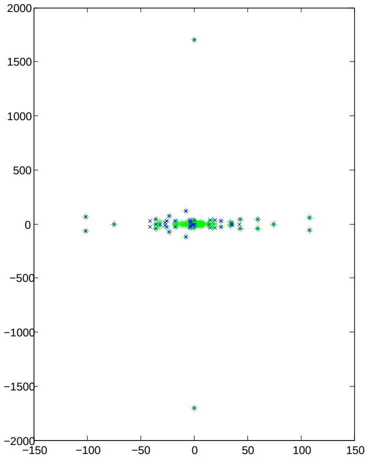
- Where

$$H_{m+1,m} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2m} \\ & h_{32} & & \vdots \\ & & \ddots & h_{mm} \\ & & & h_{m+1,m} \end{bmatrix}.$$

Degree 25 Eigenvalue Approximations



Degree 50 Eigenvalue Approximations



Outline

Arnoldi Simplified

Let $A \in \mathbb{R}^{n \times n}$ be symmetric and consider the matrix representation of A with respect to $(q_j)_{j=1}^{m}$.

$$H_{m+1,m} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2m} \\ & h_{32} & & \vdots \\ & & \ddots & h_{mm} \\ & & & h_{m+1,m} \end{bmatrix}.$$

Arnoldi Simplified

We can write H as follows

$$H_{m+1,m} = \begin{bmatrix} \alpha_1 & \beta_2 & & \\ \beta_2 & \alpha_2 & \ddots & \\ & \beta_3 & \ddots & \beta_m \\ & & \ddots & \alpha_m \\ & & & \beta_{m+1} \end{bmatrix},$$

where $\alpha_k = \langle Aq_k, q_k \rangle$ and $\beta_k = \langle Aq_k, q_{k-1} \rangle$.

Outline

A functional space example

Let $T : \mathscr{H} \to \mathscr{H}$ be the linear operator $Tf(x) = xf(x)$.

- $\mathscr{H} = L^2(-1,1)$ with inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx,$$

where f and g are real valued Lebesgue integrable functions over the interval $[-1, 1]$.

A functional space example

Let $T : \mathscr{H} \to \mathscr{H}$ be the linear operator $Tf(x) = xf(x)$.

- $\mathscr{H} = L^2(-1,1)$ with inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx,$$

where f and g are real valued Lebesgue integrable functions over the interval $[-1, 1]$.

Space of polynomials

Consider the Krylov sequence starting from $1 \in \mathscr{H}$

$$1, x, x^2, \ldots.$$

- The Symmetric Lanczos process will provide an orthonormal basis for this space!
- Starting from $q_1 \in \mathscr{H}$ such that $\|q_1\| = 1$, we have

$$\beta_{k+1} q_{k+1} = (x - \alpha_k) q_k - \beta_k q_{k-1}.$$

Space of polynomials

Consider the Krylov sequence starting from $1 \in \mathscr{H}$

$$1, x, x^2, \ldots.$$

- The Symmetric Lanczos process will provide an orthonormal basis for this space!
- Starting from $q_1 \in \mathscr{H}$ such that $\|q_1\| = 1$, we have

$$\beta_{k+1} q_{k+1} = (x - \alpha_k) q_k - \beta_k q_{k-1}.$$

Space of polynomials

Consider the Krylov sequence starting from $1 \in \mathscr{H}$

$$1, x, x^2, \ldots.$$

- The Symmetric Lanczos process will provide an orthonormal basis for this space!
- Starting from $q_1 \in \mathscr{H}$ such that $\|q_1\| = 1$, we have

$$\beta_{k+1} q_{k+1} = (x - \alpha_k) q_k - \beta_k q_{k-1}.$$

Legendre Polynomials

Let $p_1(x) = \frac{1}{\sqrt{2}}$.

- $\beta_2 p_2(x) = (x - \alpha_1)p_1(x)$,
- $\frac{1}{\sqrt{3}} p_2(x) = \frac{1}{\sqrt{2}} x \Rightarrow p_2(x) = \sqrt{\frac{3}{2}} x$.
- $\beta_3 p_3(x) = (x - \alpha_2)p_2(x) - \beta_2 p_1(x)$,
- $\frac{2}{\sqrt{15}} p_3(x) = \sqrt{\frac{3}{2}} (x^2 - 1) \Rightarrow p_3(x) = \sqrt{\frac{5}{2}} \left(\frac{3}{2}x^2 - \frac{1}{2}\right)$.

Legendre Polynomials

Let $p_1(x) = \frac{1}{\sqrt{2}}$.

- $\beta_2 p_2(x) = (x - \alpha_1)p_1(x)$,
- $\frac{1}{\sqrt{3}} p_2(x) = \frac{1}{\sqrt{2}} x \Rightarrow p_2(x) = \sqrt{\frac{3}{2}} x$.
- $\beta_3 p_3(x) = (x - \alpha_2)p_2(x) - \beta_2 p_1(x)$,
- $\frac{2}{\sqrt{15}} p_3(x) = \sqrt{\frac{3}{2}} (x^2 - 1) \Rightarrow p_3(x) = \sqrt{\frac{5}{2}} \left(\frac{3}{2}x^2 - \frac{1}{2}\right)$.

Legendre Polynomials

Let $p_1(x) = \frac{1}{\sqrt{2}}$.

- $\beta_2 p_2(x) = (x - \alpha_1)p_1(x)$,
- $\frac{1}{\sqrt{3}} p_2(x) = \frac{1}{\sqrt{2}} x \Rightarrow p_2(x) = \sqrt{\frac{3}{2}} x$.
- $\beta_3 p_3(x) = (x - \alpha_2)p_2(x) - \beta_2 p_1(x)$,
- $\frac{2}{\sqrt{15}} p_3(x) = \sqrt{\frac{3}{2}} \left(x^2 - 1\right) \Rightarrow p_3(x) = \sqrt{\frac{5}{2}} \left(\frac{3}{2}x^2 - \frac{1}{2}\right)$.

Legendre Polynomials

Let $p_1(x) = \frac{1}{\sqrt{2}}$.

- $\beta_2 p_2(x) = (x - \alpha_1)p_1(x)$,
- $\frac{1}{\sqrt{3}} p_2(x) = \frac{1}{\sqrt{2}} x \Rightarrow p_2(x) = \sqrt{\frac{3}{2}} x$.
- $\beta_3 p_3(x) = (x - \alpha_2)p_2(x) - \beta_2 p_1(x)$,
- $\frac{2}{\sqrt{15}} p_3(x) = \sqrt{\frac{3}{2}} \left(x^2 - 1\right) \Rightarrow p_3(x) = \sqrt{\frac{5}{2}} \left(\frac{3}{2} x^2 - \frac{1}{2}\right)$.

Legendre Polynomials

Let $p_1(x) = \frac{1}{\sqrt{2}}$.

- $\beta_2 p_2(x) = (x - \alpha_1)p_1(x)$,
- $\frac{1}{\sqrt{3}} p_2(x) = \frac{1}{\sqrt{2}} x \Rightarrow p_2(x) = \sqrt{\frac{3}{2}} x$.
- $\beta_3 p_3(x) = (x - \alpha_2)p_2(x) - \beta_2 p_1(x)$,
- $\frac{2}{\sqrt{15}} p_3(x) = \sqrt{\frac{3}{2}} \left(x^2 - 1\right) \Rightarrow p_3(x) = \sqrt{\frac{5}{2}} \left(\frac{3}{2}x^2 - \frac{1}{2}\right)$.

Legendre Polynomials

We are building the following infinite matrix

$$H = \begin{bmatrix} 0 & \frac{1}{\sqrt{3}} & & & \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{15}} & & \\ & \frac{2}{\sqrt{15}} & 0 & \frac{3}{\sqrt{35}} & \\ & & \frac{3}{\sqrt{35}} & 0 & \ddots \\ & & & \ddots & \ddots \end{bmatrix}.$$

Numerical Integration

Suppose we want to compute

$$I = \int_{-1}^{1} xsin(x)dx.$$

- We can approximate this integral by Gauss Quadrature

$$Q_m = 2\sum_{i=1}^{m} f(x_i)w_i$$

where $f(x) = xsin(x)$, x_i are interpolation points, and w_i are weights.

Numerical Integration

Suppose we want to compute

$$I = \int_{-1}^{1} xsin(x)dx.$$

- We can approximate this integral by Gauss Quadrature

$$Q_m = 2\sum_{i=1}^{m} f(x_i)w_i$$

where $f(x) = xsin(x)$, x_i are interpolation points, and w_i are weights.

Numerical Integration

The 4^{th} degree approximation to I is given by the eigenvalues and eigenvectors of

$$H_4 = \begin{bmatrix} 0 & \frac{1}{\sqrt{3}} & & \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{15}} & \\ & \frac{2}{\sqrt{15}} & 0 & \frac{3}{\sqrt{35}} \\ & & \frac{3}{\sqrt{35}} & 0 \end{bmatrix} = UDU^T$$

- The correct interpolation points and weights are $x_i = d_{ii}$ and $w_i = u_{1i}^2$.
- $Q_4 = 2\left(f(d_{11})u_{11}^2 + f(d_{22})u_{12}^2 + f(d_{33})u_{13}^2 + f(d_{44})u_{14}^2\right) = 0.60234$.

Numerical Integration

The 4^{th} degree approximation to I is given by the eigenvalues and eigenvectors of

$$H_4 = \begin{bmatrix} 0 & \frac{1}{\sqrt{3}} & & \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{15}} & \\ & \frac{2}{\sqrt{15}} & 0 & \frac{3}{\sqrt{35}} \\ & & \frac{3}{\sqrt{35}} & 0 \end{bmatrix} = UDU^T$$

- The correct interpolation points and weights are $x_i = d_{ii}$ and $w_i = u_{1i}^2$.
- $Q_4 = 2\left(f(d_{11})u_{11}^2 + f(d_{22})u_{12}^2 + f(d_{33})u_{13}^2 + f(d_{44})u_{14}^2\right) = 0.60234$.

Numerical Integration

The 4^{th} degree approximation to I is given by the eigenvalues and eigenvectors of

$$H_4 = \begin{bmatrix} 0 & \frac{1}{\sqrt{3}} & & \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{15}} & \\ & \frac{2}{\sqrt{15}} & 0 & \frac{3}{\sqrt{35}} \\ & & \frac{3}{\sqrt{35}} & 0 \end{bmatrix} = UDU^T$$

- The correct interpolation points and weights are $x_i = d_{ii}$ and $w_i = u_{1i}^2$.
- $Q_4 = 2\left(f(d_{11})u_{11}^2 + f(d_{22})u_{12}^2 + f(d_{33})u_{13}^2 + f(d_{44})u_{14}^2\right) = 0.60234$.