Some Linear Algeba Insights

The representation of a vector with respect to a basis

Let V be a real or complex finite dimensional vector space and $B = \{v_1, \ldots, v_n\}$ be a basis of V.

For $v \in V$, there exist unique scalars $a_1, \ldots, a_n$ such that $v = \sum_{i=1}^{n} a_i v_i$. The vector

$$[v]_B := \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{F}^n$$

is defined to be the representation of v with respect to the basis B.

Example: The representation of $p(x) = 1 + 2x - x^2$ with respect to the standard basis $B = \{1, x, x^2\}$ of the vector space $P_2(\mathbb{R})$ of real polynomials of degree at most 2 is

$$[p]_B = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

The representation of a vector with respect to a basis

Let $B = \{e_1, \ldots, e_n\}$ be the standard basis in $\mathbb{F}^n$.
For $x = [x_1 \ \cdots \ x_n]^T \in \mathbb{F}^n$, $[x]_B = x$ as $x = \sum_{i=1}^n x_i e_i$.

The representation of a vector with respect to a basis

Let $B = \{e_1, \ldots, e_n\}$ be the standard basis in $\mathbb{F}^n$.
For $x = [x_1 \ \cdots \ x_n]^T \in \mathbb{F}^n$, $[x]_B = x$ as $x = \sum_{i=1}^n x_i e_i$.

For another basis $B' = \{v_1, \ldots, v_n\}$ of $\mathbb{F}^n$, $[x]_{B'} = [y_1 \ \cdots \ y_n]^T$ if

$$x = \sum_{i=1}^n v_i y_i.$$

Let $V = [v_1 \ v_2 \ \cdots \ v_n]$. Then $x = Vy \Rightarrow y = V^{-1}x$. Thus

$$[x]_{B'} = V^{-1}x.$$

Matrix of a linear map with respect to a basis

Consider the linear map $T : \mathbb{R}^3 \mapsto \mathbb{R}^3$ defined by

$$T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 - x_3 \\ x_1 + x_2 \\ 2x_1 + 2x_3 + 3x_3 \end{bmatrix}.$$

Evidently, $Tx = Ax$ for all $x \in \mathbb{R}^3$ where

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ and } A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 0 \\ 2 & 2 & 3 \end{bmatrix}.$$

Observe that for the standard basis $B = \{e_1, e_2, e_3\}$ of $\mathbb{R}^3$,

$$[Te_1]_B = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, [Te_2]_B = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, [Te_3]_B = \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}.$$

A is defined to the matrix of T with respect to the basis B.

Matrix of a linear map with respect to a basis

Let V be a real or complex finite dimensional vector space and $T : V \mapsto V$ be a linear transformation.

Given a basis $B = \{v_1, \dots, v_n\}$, of V, the matrix of T with respect to B, is defined to be the $n \times n$ matrix whose i^{th} column is $[Tv_i]_B$ for each $i = 1, \dots, n$.

Consider the basis $B' = \{v_1, v_2, v_3\}$ of $\mathbb{R}^3$ where

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

For the linear map $T : \mathbb{R}^3 \mapsto \mathbb{R}^3$ in the previous slide,

$$Tv_1 = T \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{5}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

so that $[Tv_1]_{B'} = \begin{bmatrix} \frac{1}{2} \\ \frac{5}{2} \\ 4 \end{bmatrix}$. Likewise,

$$[Tv_2]_{B'} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{bmatrix},\ [Tv_3]_{B'} = \begin{bmatrix} -\frac{3}{2} \\ \frac{5}{2} \\ 5 \end{bmatrix}.$$

Therefore,

$$[T]_{B'} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{3}{2} \\ \frac{5}{2} & -\frac{1}{2} & \frac{5}{2} \\ 4 & 0 & 5 \end{bmatrix}.$$

Observe that

$$[T]_{B'} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 0 \\ 2 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = V^{-1}[T]_B V$$

where $V = [v_1\ v_2\ v_3]$.

Taking similarity transformations is equivalent to changing bases

Theorem Let $T : \mathbb{F}^n \mapsto \mathbb{F}^n$ be linear. Let $B = \{e_1, \ldots, e_n\}$ be the standard basis and $B' = \{v_1, \ldots, v_n\}$ be any other basis of $\mathbb{F}^n$. Then for $V = [v_1 \ \cdots \ v_n]$,

$$[T]_{B'} = V^{-1}[T]_B V.$$

Taking similarity transformations is equivalent to changing bases

Theorem Let $T : \mathbb{F}^n \mapsto \mathbb{F}^n$ be linear. Let $B = \{e_1, \ldots, e_n\}$ be the standard basis and $B' = \{v_1, \ldots, v_n\}$ be any other basis of $\mathbb{F}^n$. Then for $V = [v_1 \ \cdots \ v_n]$,

$$[T]_{B'} = V^{-1}[T]_B V.$$

In general, if $B_1 = \{w_1, \ldots, w_n\}$ and $B_2 = \{u_1, \ldots, u_n\}$ are any two bases of $\mathbb{F}^n$, then

$$[T]_{B_2} = S^{-1}[T]_{B_1} S$$

where $S = \left[[u_1]_{B_1} \ \cdots \ [u_n]_{B_1}\right].$

Taking similarity transformations is equivalent to changing bases

Theorem Let $T : \mathbb{F}^n \mapsto \mathbb{F}^n$ be linear. Let $B = \{e_1, \ldots, e_n\}$ be the standard basis and $B' = \{v_1, \ldots, v_n\}$ be any other basis of $\mathbb{F}^n$. Then for $V = [v_1 \ \cdots \ v_n]$,

$$[T]_{B'} = V^{-1}[T]_B V.$$

In general, if $B_1 = \{w_1, \ldots, w_n\}$ and $B_2 = \{u_1, \ldots, u_n\}$ are any two bases of $\mathbb{F}^n$, then

$$[T]_{B_2} = S^{-1}[T]_{B_1} S$$

where $S = \left[[u_1]_{B_1} \ \cdots \ [u_n]_{B_1}\right]$.

Corollary Let $T : \mathbb{F}^n \mapsto \mathbb{F}^n$ be linear and $B_1 = \{w_1, \ldots, w_n\}$ and $B_2 = \{u_1, \ldots, u_n\}$ be any two orthonormal bases of $\mathbb{F}^n$. Then $Q = [[u_1]_{B_1} \ \cdots \ [u_n]_{B_1}]$ is a unitary matrix such that $[T]_{B_2} = Q^*[T]_{B_1} Q$.

In particular if B_1 is the standard orthonormal basis $\{e_1, \ldots, e_n\}$, then $Q = [u_1 \ \cdots \ u_n]$.

Schur's Theorem: A fresh perspective

Schur's Theorem: Given any matrix $A \in \mathbb{C}^{n \times n}$, there exists a unitary matrix Q and an upper triangular matrix T such that $Q^* A Q = T$.

Schur's Theorem: A fresh perspective

Schur's Theorem: Given any matrix $A \in \mathbb{C}^{n \times n}$, there exists a unitary matrix Q and an upper triangular matrix T such that $Q^* A Q = T$.

Equivalently, given $A \in \mathbb{C}^{n \times n}$, there exists an orthonormal basis of $\mathbb{C}^n$ with respect to which the matrix of the linear operator represented by A is an upper triangular matrix.

Schur's Theorem: A fresh perspective

Schur's Theorem: Given any matrix $A \in \mathbb{C}^{n \times n}$, there exists a unitary matrix Q and an upper triangular matrix T such that $Q^* A Q = T$.

Equivalently, given $A \in \mathbb{C}^{n \times n}$, there exists an orthonormal basis of $\mathbb{C}^n$ with respect to which the matrix of the linear operator represented by A is an upper triangular matrix.

Spectral Theorem for Hermitian Matrices $A \in \mathbb{C}^{n \times n}$ is Hermitian if and only if there exists a unitary matrix Q and a real diagonal matrix D such that $Q^* A Q = D$.

Schur's Theorem: A fresh perspective

Schur's Theorem: Given any matrix $A \in \mathbb{C}^{n\times n}$, there exists a unitary matrix Q and an upper triangular matrix T such that $Q^*AQ = T$.

Equivalently, given $A \in \mathbb{C}^{n\times n}$, there exists an orthonormal basis of $\mathbb{C}^n$ with respect to which the matrix of the linear operator represented by A is an upper triangular matrix.

Spectral Theorem for Hermitian Matrices $A \in \mathbb{C}^{n\times n}$ is Hermitian if and only if there exists a unitary matrix Q and a real diagonal matrix D such that $Q^*AQ = D$.

Spectral Theorem for Normal Matrices $A \in \mathbb{C}^{n\times n}$ is normal if and only if there exists a unitary matrix Q and a diagonal matrix D such that $Q^*AQ = D$.

Matrix of a linear map when domain and range space may be different

Let $T : \mathbb{F}^n \mapsto \mathbb{F}^m$ be a linear transformation.

Suppose $B_1 = \{q_1, \ldots, q_n\}$ and $B_2 = \{q'_1, \ldots, q'_m\}$ are bases of $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively. The matrix of T with respect to B_1 and B_2 is defined as

$$[T]_{B_1,B_2} = \left[\begin{array}{ccc} [Tq_1]_{B_2} & \cdots & [Tq_n]_{B_2} \end{array} \right] \in \mathbb{F}^{m\times n}.$$

Matrix of a linear map when domain and range space may be different

Let $T : \mathbb{F}^n \mapsto \mathbb{F}^m$ be a linear transformation.

Suppose $B_1 = \{q_1, \ldots, q_n\}$ and $B_2 = \{q'_1, \ldots, q'_m\}$ are bases of $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively. The matrix of T with respect to B_1 and B_2 is defined as

$$[T]_{B_1,B_2} = \begin{bmatrix} [Tq_1]_{B_2} & \cdots & [Tq_n]_{B_2} \end{bmatrix} \in \mathbb{F}^{m \times n}.$$

Consider $T : \mathbb{R}^2 \mapsto \mathbb{R}^3$ defined by

$$T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ x_2 \\ x_1 - x_2 \end{bmatrix}.$$

Let $B_1 = \{e_1, e_2\}$ be the standard basis of $\mathbb{R}^2$ and $B_2 = \{v_1, v_2, v_3\}$ be a basis of $\mathbb{R}^3$ where

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \; v_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \; v_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Matrix of a linear map when domain and range spaces may be different

Then,

$$Te_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 0.v_1 + v_2 + v_3$$

$$Te_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 2v_1 + 0.v_2 - v_3$$

so that $[T]_{B_1,B_2} = \begin{bmatrix} 0 & 2 \\ 1 & 0 \\ 1 & -1 \end{bmatrix}$.

Change of basis when range and domain may be different

Theorem Let $T : \mathbb{F}^n \mapsto \mathbb{F}^m$ be linear. Let B_1, B_1' be two bases of $\mathbb{F}^n$ and B_2, B_2' be two bases of $\mathbb{F}^m$ where $B_1' = \{v_1, \ldots, v_n\}$, $B_2' = \{u_1, \ldots, u_m\}$ and B_1 and B_2 are the standard bases. Then for $V = [v_1 \ \cdots \ v_n]$ and $U = [u_1 \ \cdots \ u_m]$,

$$[T]_{B_1', B_2'} = U^{-1} [T]_{B_1, B_2} V.$$

Change of basis when range and domain may be different

Theorem Let $T : \mathbb{F}^n \mapsto \mathbb{F}^m$ be linear. Let B_1, B_1' be two bases of $\mathbb{F}^n$ and B_2, B_2' be two bases of $\mathbb{F}^m$ where $B_1' = \{v_1, \dots, v_n\}$, $B_2' = \{u_1, \dots, u_m\}$ and B_1 and B_2 are the standard bases. Then for $V = [v_1 \ \cdots \ v_n]$ and $U = [u_1 \ \cdots \ u_m]$,

$$[T]_{B_1',B_2'} = U^{-1}[T]_{B_1,B_2} V.$$

In general, if $B_1 = \{w_1, \dots, w_n\}$ and $B_2 = \{q_1, \dots, q_m\}$ are any two bases of $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively, then

$$[T]_{B_1',B_2'} = M_2[T]_{B_1,B_2} M_1$$

where $M_1 = [[v_1]_{B_1} \ \cdots \ [v_n]_{B_1}]$ and $M_2 = \left[[q_1]_{B_2'} \ \cdots \ [q_m]_{B_2'}\right]$.

Change of basis when range and domain may be different

Theorem Let $T : \mathbb{F}^n \mapsto \mathbb{F}^m$ be linear. Let B_1, B_1' be two bases of $\mathbb{F}^n$ and B_2, B_2' be two bases of $\mathbb{F}^m$ where $B_1' = \{v_1, \dots, v_n\}$, $B_2' = \{u_1, \dots, u_m\}$ and B_1 and B_2 are the standard bases. Then for $V = [v_1 \ \cdots \ v_n]$ and $U = [u_1 \ \cdots \ u_m]$,

$$[T]_{B_1', B_2'} = U^{-1}[T]_{B_1, B_2} V.$$

In general, if $B_1 = \{w_1, \dots, w_n\}$ and $B_2 = \{q_1, \dots, q_m\}$ are any two bases of $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively, then

$$[T]_{B_1', B_2'} = M_2[T]_{B_1, B_2} M_1$$

where $M_1 = [[v_1]_{B_1} \ \cdots \ [v_n]_{B_1}]$ and $M_2 = \left[[q_1]_{B_2'} \ \cdots \ [q_m]_{B_2'} \right]$.

Corollary Let $T : \mathbb{F}^n \mapsto \mathbb{F}^m$ be linear where $m \neq n$. Let B_1, B_1' be two orthonormal bases of $\mathbb{F}^n$ and B_2, B_2' be two orthonormal bases of $\mathbb{F}^m$ where $B_1' = \{v_1, \dots, v_n\}$, $B_2' = \{u_1, \dots, u_m\}$ and B_1 and B_2 are the standard bases. Then for $V = [v_1 \ \cdots \ v_n]$ and $U = [u_1 \ \cdots \ u_m]$,

$$[T]_{B_1', B_2'} = U^*[T]_{B_1, B_2} V.$$

The SVD: An alternative view

Singular Value Decomposition: The Singular Value Decomposition (SVD) of a matrix $A \in \mathbb{F}^{n \times m}$ is a decomposition of the form

$$A = U \Sigma V^*$$

where $U \in \mathbb{F}^{n \times n}$ and $V \in \mathbb{F}^{m \times m}$ are unitary matrices and $\Sigma := \mathrm{diag}(\sigma_1, \sigma_2, \ldots \sigma_p) \in \mathbb{R}^{n \times m}$ is a diagonal matrix with

$$\sigma_1 \geq \sigma_2 \geq \cdots \sigma_p \geq 0$$

for $p = \min\{n, m\}$.

The SVD: An alternative view

Singular Value Decomposition: The Singular Value Decomposition (SVD) of a matrix $A \in \mathbb{F}^{n\times m}$ is a decomposition of the form

$$A = U\Sigma V^*$$

where $U \in \mathbb{F}^{n\times n}$ and $V \in \mathbb{F}^{m\times m}$ are unitary matrices and $\Sigma := \mathrm{diag}(\sigma_1, \sigma_2, \ldots \sigma_p) \in \mathbb{R}^{n\times m}$ is a diagonal matrix with

$$\sigma_1 \geq \sigma_2 \geq \cdots \sigma_p \geq 0$$

for $p = \min\{n, m\}$.

Equivalently, given any $A \in \mathbb{F}^{n\times m}$, there exist suitable choices of orthonormal bases in $\mathbb{F}^n$ and $\mathbb{F}^m$ with respect to which the matrix of the linear transformation defined by A is a diagonal matrix $\Sigma := \mathrm{diag}(\sigma_1, \sigma_2, \ldots \sigma_p) \in \mathbb{R}^{n\times m}$ with

$$\sigma_1 \geq \sigma_2 \geq \cdots \sigma_p \geq 0$$

for $p = \min\{n, m\}$.

Invariant Subspaces

Consider $A : \mathbb{F}^n \mapsto \mathbb{F}^n$. A subspace $V \subset \mathbb{F}^n$ is said to be invariant with respect to A, if $Av \in V$ for all $v \in V$.

Invariant Subspaces

Consider $A : \mathbb{F}^n \mapsto \mathbb{F}^n$. A subspace $V \subset \mathbb{F}^n$ is said to be invariant with respect to A, if $Av \in V$ for all $v \in V$.

Examples:

1. For $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & -2 & 0 \\ 1 & 1 & -1 \end{bmatrix}$, $A\left(\mathrm{span}\{e_1, e_3\}\right) \subseteq \mathrm{span}\{e_1, e_3\}$.
2. For $A = \begin{bmatrix} 4 & 3 & -5 \\ 0 & -3 & 3 \\ 0 & -2 & 3 \end{bmatrix}$, $A(\mathrm{span}\{e_1\}) \subseteq \mathrm{span}\{e_1\}$ and $A\left(\mathrm{span}\{e_1 + e_2 + e_3, e_1 - e_2\}\right) \subseteq \mathrm{span}\{e_1 + e_2 + e_3, e_1 - e_2\}$.

Invariant Subspaces

Consider $A : \mathbb{F}^n \mapsto \mathbb{F}^n$. A subspace $V \subset \mathbb{F}^n$ is said to be invariant with respect to A, if $Av \in V$ for all $v \in V$.

Examples:

1. For $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & -2 & 0 \\ 1 & 1 & -1 \end{bmatrix}$, $A\left(\mathrm{span}\{e_1, e_3\}\right) \subseteq \mathrm{span}\{e_1, e_3\}$.
2. For $A = \begin{bmatrix} 4 & 3 & -5 \\ 0 & -3 & 3 \\ 0 & -2 & 3 \end{bmatrix}$, $A(\mathrm{span}\{e_1\}) \subseteq \mathrm{span}\{e_1\}$ and $A\left(\mathrm{span}\{e_1 + e_2 + e_3, e_1 - e_2\}\right) \subseteq \mathrm{span}\{e_1 + e_2 + e_3, e_1 - e_2\}$.

Facts:

1. The trivial subspaces $\mathbb{F}^n$ and $\{0\}$ are always invariant with respect to every $A \in \mathbb{F}^{n \times n}$.
2. $V \subseteq \mathbb{F}^n$ is a one dimensional subspace of $\mathbb{F}^n$ invariant with respect to $A \in \mathbb{F}^{n \times n}$ if and only if $V = \mathrm{span}\{v\}$ for some eigenvector v of A.
3. Eigenvectors of $A \in \mathbb{F}^{n \times n}$ span invariant subspaces.

Invariant Subspaces

Theorem: Let $A \in \mathbb{F}^{n\times n}$. The first k columns of an invertible $S \in \mathbb{F}^{n\times n}$ span a subspace of $\mathbb{F}^n$ invariant with respect to A if and only if

$$S^{-1}AS = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline & A_{22} \end{array}\right]$$

where $A_{11} \in \mathbb{F}^{k\times k}$, $A_{12} \in \mathbb{F}^{k\times n-k}$ and $A_{22} \in \mathbb{F}^{n-k\times n-k}$.

Invariant Subspaces

Theorem: Let $A \in \mathbb{F}^{n\times n}$. The first k columns of an invertible $S \in \mathbb{F}^{n\times n}$ span a subspace of $\mathbb{F}^n$ invariant with respect to A if and only if

$$S^{-1}AS = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline & A_{22} \end{array}\right]$$

where $A_{11} \in \mathbb{F}^{k\times k}, A_{12} \in \mathbb{F}^{k\times n-k}$ and $A_{22} \in \mathbb{F}^{n-k\times n-k}$.

Corollary: Let $A \in \mathbb{F}^{n\times n}$ and $S = [s_1 \ \cdots \ s_n] \in \mathbb{F}^{n\times n}$ be a invertible matrix. Then the first k columns of S span subspaces of $\mathbb{F}^n$ that are invariant with respect to A $\underline{\text{for each } k = 1, \ldots, n-1,}$ if and only if $S^{-1}AS$ is an upper triangular matrix.

Invariant Subspaces

Theorem: Let $A \in \mathbb{F}^{n\times n}$. The first k columns of an invertible $S \in \mathbb{F}^{n\times n}$ span a subspace of $\mathbb{F}^n$ invariant with respect to A if and only if

$$S^{-1}AS = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline & A_{22} \end{array}\right]$$

where $A_{11} \in \mathbb{F}^{k\times k}$, $A_{12} \in \mathbb{F}^{k\times n-k}$ and $A_{22} \in \mathbb{F}^{n-k\times n-k}$.

Corollary: Let $A \in \mathbb{F}^{n\times n}$ and $S = [s_1 \ \cdots \ s_n] \in \mathbb{F}^{n\times n}$ be a invertible matrix. Then the first k columns of S span subspaces of $\mathbb{F}^n$ that are invariant with respect to A for each $k = 1, \ldots, n-1$, if and only if $S^{-1}AS$ is an upper triangular matrix.

Schur's Theorem: Given $A \in \mathbb{C}^{n\times n}$, there exists an orthonormal basis $\{q_1, \ldots, q_n\}$ of $\mathbb{C}^n$ such that

$$A(\mathrm{span}\{q_1, \ldots, q_k\}) \subseteq \mathrm{span}\{q_1, \ldots, q_k\}$$

for each $k = 1, \ldots, n-1$.

Subspace Iteration

Given $A \in \mathbb{F}^{n \times n}$, recall that the Power Method is essentially about producing a series of vectors

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

which under suitable scaling converges to a dominant eigenvector of A under suitable conditions.

Subspace Iteration

Given $A \in \mathbb{F}^{n \times n}$, recall that the Power Method is essentially about producing a series of vectors

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

which under suitable scaling converges to a dominant eigenvector of A under suitable conditions.

The scalings will replace x by other vectors in $\mathcal{S} := \mathrm{span}\{x\}$ and depending on their choices the iterations will converge to some vector in the one dimensional invariant eigenspace $\mathcal{T} := \mathrm{span}\,\{v\}$ where v is a dominant eigenvector of A.

Subspace Iteration

Given $A \in \mathbb{F}^{n\times n}$, recall that the Power Method is essentially about producing a series of vectors

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

which under suitable scaling converges to a dominant eigenvector of A under suitable conditions.

The scalings will replace x by other vectors in $\mathcal{S} := \mathrm{span}\{x\}$ and depending on their choices the iterations will converge to some vector in the one dimensional invariant eigenspace $\mathcal{T} := \mathrm{span}\,\{v\}$ where v is a dominant eigenvector of A.

So in Power Method we are doing

$$\mathcal{S}, A(\mathcal{S}), A^2(\mathcal{S}), \dots$$

which converges to $\mathcal{T}$ under the right conditions.

Subspace Iteration

Given $A \in \mathbb{F}^{n \times n}$, recall that the Power Method is essentially about producing a series of vectors

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

which under suitable scaling converges to a dominant eigenvector of A under suitable conditions.

The scalings will replace x by other vectors in $\mathcal{S} := \mathrm{span}\{x\}$ and depending on their choices the iterations will converge to some vector in the one dimensional invariant eigenspace $\mathcal{T} := \mathrm{span}\{v\}$ where v is a dominant eigenvector of A.

So in Power Method we are doing

$$\mathcal{S}, A(\mathcal{S}), A^2(\mathcal{S}), \ldots$$

which converges to $\mathcal{T}$ under the right conditions.

We could do this for higher dimensional subspaces $\mathcal{S}$!

Subspace iteration, Simultaneous iteration and the QR algorithm

Principal angles and distance between subspaces

Let $\mathcal{S}_1, \mathcal{S}_2$ be two subspaces of $\mathbb{F}^n$ with $\dim \mathcal{S}_1 = l$, $\dim \mathcal{S}_2 = m$ where $l \leq m$.

Principal angles and distance between subspaces

Let $\mathcal{S}_1, \mathcal{S}_2$ be two subspaces of $\mathbb{F}^n$ with $\dim \mathcal{S}_1 = l$, $\dim \mathcal{S}_2 = m$ where $l \leq m$.

Let the columns of $U_l \in \mathbb{F}^{n\times l}$ and $U_2 \in \mathbb{F}^{n\times m}$ form orthonormal bases of $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. If

$$\sigma_1 \geq \cdots \geq \sigma_l \geq 0$$

are the singular values of $U_m^T U_l$ then the **principal angles between $\mathcal{S}_1$ and $\mathcal{S}_2$** are defined as

$$\theta_i = \arccos \sigma_i, \; i = 1, \ldots, l.$$

Principal angles and distance between subspaces

Let $\mathcal{S}_1, \mathcal{S}_2$ be two subspaces of $\mathbb{F}^n$ with $\dim \mathcal{S}_1 = l$, $\dim \mathcal{S}_2 = m$ where $l \leq m$.

Let the columns of $U_l \in \mathbb{F}^{n\times l}$ and $U_2 \in \mathbb{F}^{n\times m}$ form orthonormal bases of $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. If

$$\sigma_1 \geq \cdots \geq \sigma_l \geq 0$$

are the singular values of $U_m^T U_l$ then the **principal angles between** $\mathcal{S}_1$ **and** $\mathcal{S}_2$ are defined as

$$\theta_i = \operatorname{arcos} \sigma_i, \; i = 1, \ldots, l.$$

The **distance** $d(\mathcal{S}_1, \mathcal{S}_2)$ **between** $\mathcal{S}_1$ **and** $\mathcal{S}_2$ is defined by

$$d(\mathcal{S}_1, \mathcal{S}_2) = \sqrt{1 - \sigma_l^2} = \sin\, \theta_l.$$

Principal angles and distance between subspaces

Let $\mathcal{S}_1, \mathcal{S}_2$ be two subspaces of $\mathbb{F}^n$ with $\dim \mathcal{S}_1 = l$, $\dim \mathcal{S}_2 = m$ where $l \leq m$.

Let the columns of $U_l \in \mathbb{F}^{n\times l}$ and $U_2 \in \mathbb{F}^{n\times m}$ form orthonormal bases of $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. If

$$\sigma_1 \geq \cdots \geq \sigma_l \geq 0$$

are the singular values of $U_m^T U_l$ then the **principal angles between** $\mathcal{S}_1$ **and** $\mathcal{S}_2$ are defined as

$$\theta_i = \arccos \sigma_i, \; i = 1, \ldots, l.$$

The **distance** $d(\mathcal{S}_1, \mathcal{S}_2)$ **between** $\mathcal{S}_1$ **and** $\mathcal{S}_2$ is defined by

$$d(\mathcal{S}_1, \mathcal{S}_2) = \sqrt{1 - \sigma_l^2} = \sin\, \theta_l.$$

Given a subspace $\mathcal{T}$ and sequence of subspaces $\{\mathcal{S}_m\}$ of $\mathbb{F}^n$,

$$\lim_{m\to\infty} \mathcal{S}_m = \mathcal{T} \Leftrightarrow d(\mathcal{S}_m, \mathcal{T}) \to 0 \text{ as } m \to \infty.$$

Convergence in Subspace iterations

Let $A \in \mathbb{F}^{n\times n}$ be diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$ such that

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|.$$

Let $\{v_1, \ldots, v_n\}$ be a basis of $\mathbb{F}^n$ where $Av_i = \lambda_i v_i$, $i = 1, \ldots, n$.

Convergence in Subspace iterations

Let $A \in \mathbb{F}^{n\times n}$ be diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$ such that

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|.$$

Let $\{v_1, \ldots, v_n\}$ be a basis of $\mathbb{F}^n$ where $Av_i = \lambda_i v_i$, $i = 1, \ldots, n$.

Suppose $|\lambda_k| > |\lambda_{k+1}|$ for some $1 \leq k \leq n-1$ and let

$$\mathcal{T}_k = \mathrm{span}\{v_1, \ldots, v_k\},\ \mathcal{U}_k = \mathrm{span}\{v_{k+1}, \ldots, v_n\}.$$

Convergence in Subspace iterations

Let $A \in \mathbb{F}^{n\times n}$ be diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$ such that

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|.$$

Let $\{v_1, \ldots, v_n\}$ be a basis of $\mathbb{F}^n$ where $Av_i = \lambda_i v_i$, $i = 1, \ldots, n$.

Suppose $|\lambda_k| > |\lambda_{k+1}|$ for some $1 \leq k \leq n-1$ and let

$$\mathcal{T}_k = \mathrm{span}\{v_1, \ldots, v_k\},\ \mathcal{U}_k = \mathrm{span}\{v_{k+1}, \ldots, v_n\}.$$

Then for *any* subspace $\mathcal{S}$ of $\mathbb{F}^n$ such that $\dim \mathcal{S} = k$, and $\mathcal{S} \cap \mathcal{U}_k = \{0\}$,

$$\lim_{m\to\infty} A^m(\mathcal{S}) = \mathcal{T}_k$$

linearly at the rate $|\lambda_{k+1}|/|\lambda_k|$.

Convergence in Subspace iterations

Let $A \in \mathbb{F}^{n\times n}$ be diagonalizable with eigenvalues $\lambda_1, \dots, \lambda_n$ such that

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|.$$

Let $\{v_1, \dots, v_n\}$ be a basis of $\mathbb{F}^n$ where $Av_i = \lambda_i v_i$, $i = 1, \dots, n$.

Suppose $|\lambda_k| > |\lambda_{k+1}|$ for some $1 \leq k \leq n-1$ and let

$$\mathcal{T}_k = \mathrm{span}\{v_1, \dots, v_k\},\ \mathcal{U}_k = \mathrm{span}\{v_{k+1}, \dots, v_n\}.$$

Then for *any* subspace $\mathcal{S}$ of $\mathbb{F}^n$ such that $\dim \mathcal{S} = k$, and $\mathcal{S} \cap \mathcal{U}_k = \{0\}$,

$$\lim_{m\to\infty} A^m(\mathcal{S}) = \mathcal{T}_k$$

linearly at the rate $|\lambda_{k+1}|/|\lambda_k|$.

Exercise: Since $|\lambda_k| > |\lambda_{k+1}|$, prove the following:

(a) $\mathrm{Null}(A^m) \subseteq \mathcal{U}_k$ for all $m = 1, 2, \dots$.

(b) $\dim A^m(\mathcal{S}) = k$ for all $m = 1, 2, \dots$.

Convergence in Subspace iterations

How do we find $A^m(S)$*?*

Convergence in Subspace iterations

How do we find $A^m(S)$*?*

Find an orthonormal basis of $\mathcal{S}$ *say* $\{q_1, \ldots, q_k\}$ *and produce bases* $\{A^m q_1, \ldots, A^m q_k\}$ *of* $A^m(\mathcal{S})$ *for* $m = 1, 2, \ldots$.

Convergence in Subspace iterations

How do we find $A^m(S)$*?*

Find an orthonormal basis of $\mathcal{S}$ *say* $\{q_1, \ldots, q_k\}$ *and produce bases* $\{A^m q_1, \ldots, A^m q_k\}$ *of* $A^m(\mathcal{S})$ *for* $m = 1, 2, \ldots$.

Not a good idea as there can be underflow and overflow. Also if $|\lambda_1| > |\lambda_2|$, then each of the sequences $\{A^m q_j\}, j = 1, 2, \ldots,$ could converge to some multiple of v_1 as $m \to \infty$ making $\{A^m q_1, \ldots, A^m q_k\}$ an ill conditioned basis for large m.

Convergence in Subspace iterations

How do we find $A^m(S)$?

Find an orthonormal basis of $\mathcal{S}$ say $\{q_1, \ldots, q_k\}$ and produce bases $\{A^m q_1, \ldots, A^m q_k\}$ of $A^m(\mathcal{S})$ for $m = 1, 2, \ldots$.

Not a good idea as there can be underflow and overflow. Also if $|\lambda_1| > |\lambda_2|$, then each of the sequences $\{A^m q_j\}, j = 1, 2, \ldots,$ could converge to some multiple of v_1 as $m \to \infty$ making $\{A^m q_1, \ldots, A^m q_k\}$ an ill conditioned basis for large m.

Instead we start with an orthonormal basis $q_1^{(0)}, \ldots, q_k^{(0)}$ of $\mathcal{S}$ and perform the following iterations.

Convergence in Subspace iterations

How do we find $A^m(S)$?

Find an orthonormal basis of $\mathcal{S}$ say $\{q_1, \ldots, q_k\}$ and produce bases $\{A^m q_1, \ldots, A^m q_k\}$ of $A^m(\mathcal{S})$ for $m = 1, 2, \ldots$.

Not a good idea as there can be underflow and overflow. Also if $|\lambda_1| > |\lambda_2|$, then each of the sequences $\{A^m q_j\}, j = 1, 2, \ldots,$ could converge to some multiple of v_1 as $m \to \infty$ making $\{A^m q_1, \ldots, A^m q_k\}$ an ill conditioned basis for large m.

Instead we start with an orthonormal basis $q_1^{(0)}, \ldots, q_k^{(0)}$ of $\mathcal{S}$ and perform the following iterations.

for $m = 1, 2, \ldots,$

(i) Compute $Aq_1^{(m)}, \ldots, Aq_k^{(m)}$;

(ii) Orthonormalize $Aq_1^{(m)}, \ldots, Aq_k^{(m)}$ to form an orthonormal basis $q_1^{(m+1)}, \ldots, q_k^{(m+1)}$ of $A^{m+1}(\mathcal{S})$.

Convergence in Subspace iterations

How do we know that the iterations have converged?

Convergence in Subspace iterations

How do we know that the iterations have converged?

Let Q_m be a unitary matrix whose first k columns are $q_1^{(m)}, \ldots, q_k^{(m)}$. If the iterations have converged, then $A_m := Q_m^* A Q_m$ is of the form

$$A_m = \left[\begin{array}{c|c} A_{11}^{(m)} & A_{12}^{(m)} \\ \hline & A_{22}^{(m)} \end{array}\right].$$

Convergence in Subspace iterations

How do we know that the iterations have converged?

Let Q_m be a unitary matrix whose first k columns are $q_1^{(m)}, \dots, q_k^{(m)}$. If the iterations have converged, then $A_m := Q_m^* A Q_m$ is of the form

$$A_m = \left[\begin{array}{c|c} A_{11}^{(m)} & A_{12}^{(m)} \\ \hline & A_{22}^{(m)} \end{array}\right].$$

In fact $\|A_{12}^{(m)}\|$ is an alternative measure of the distance between $A^m(\mathcal{S})$ and $\mathcal{T}_k$.

Convergence in Subspace iterations

How do we know that the iterations have converged?

Let Q_m be a unitary matrix whose first k columns are $q_1^{(m)}, \ldots, q_k^{(m)}$. If the iterations have converged, then $A_m := Q_m^* A Q_m$ is of the form

$$A_m = \left[\begin{array}{c|c} A_{11}^{(m)} & A_{12}^{(m)} \\ \hline & A_{22}^{(m)} \end{array}\right].$$

In fact $\|A_{12}^{(m)}\|$ is an alternative measure of the distance between $A^m(\mathcal{S})$ and $\mathcal{T}_k$.

In particular if A is upper Hessenberg and the unitary Q_m are such that A_m are also always upper Hessenberg, then $a_{k+1,k}^{(m)}$ is the only nonzero entry of $A_{12}^{(m)}$ and

$$\lim_{m\to\infty} |a_{k+1,k}^{(m)}| = 0 \Leftrightarrow \lim_{m\to\infty} A^m(\mathcal{S}) = \mathcal{T}_k.$$

For either limits, the convergence is linear at the rate $|\lambda_{k+1}|/|\lambda_k|$.

Simultaneous iterations in subspace iterations

The iterations on the lower dimensional subspaces $\mathrm{span}\{q_1^{(0)}, \ldots, q_j^{(0)}\}, j = 1, 2, \ldots, k-1$ *happen automatically in the process!*

Simultaneous iterations in subspace iterations

The iterations on the lower dimensional subspaces $\mathrm{span}\{q_1^{(0)}, \ldots, q_j^{(0)}\}, j = 1, 2, \ldots, k-1$ *happen automatically in the process!*

Therefore for each $j = 1, \ldots, m-1$, $q_1^{(m)}, \ldots, q_j^{(m)}$ is an orthonormal basis of $A^m(\mathrm{span}\{q_1^{(0)}, \ldots, q_j^{(0)}\})$

Simultaneous iterations in subspace iterations

The iterations on the lower dimensional subspaces $\mathrm{span}\{q_1^{(0)}, \ldots, q_j^{(0)}\}, j = 1, 2, \ldots, k-1$ *happen automatically in the process!*

Therefore for each $j = 1, \ldots, m-1$, $q_1^{(m)}, \ldots, q_j^{(m)}$ is an orthonormal basis of $A^m(\mathrm{span}\{q_1^{(0)}, \ldots, q_j^{(0)}\})$

So if for each $j = 1, \ldots, k$,

$$|\lambda_j| > |\lambda_{j+1}| \text{ and } \mathrm{span}\{q_1^{(0)}, \ldots, q_j^{(0)}\} \cap \mathrm{span}\{v_{j+1}, \ldots, v_n\} = \{0\},$$

then *for large enough m*,

$$\{q_1^{(m)}, \ldots, q_k^{(m)}\},$$

is *nearly* an orthonormal basis of the invariant subspace $\mathcal{T}_k$ with respect to A with the *additional* property that for each $j = 1, \ldots, k-1$,

$$\{q_1^{(m)}, \ldots, q_j^{(m)}\},$$

are *nearly* orthonormal bases of j-dimensional invariant subspaces $\mathrm{span}\{v_1, \ldots, v_j\}$ with respect to A.

QR algorithm and Simultaneous iterations

The goal in the QR Algorithm is to find a Schur form the square matrix $A \in \mathbb{C}^{n \times n}$.

QR algorithm and Simultaneous iterations

The goal in the QR Algorithm is to find a Schur form the square matrix $A \in \mathbb{C}^{n \times n}$.

Due to Schur's Theorem, this is equivalent to finding an orthonormal basis $q_1, \dots, q_n$ of $\mathbb{C}^n$ such that $q_1, \dots, q_j$ is an orthonormal basis of an invariant subspace of $\mathbb{C}^n$ with respect to A for each $j = 1, \dots, n-1$.

QR algorithm and Simultaneous iterations

The goal in the QR Algorithm is to find a Schur form the square matrix $A \in \mathbb{C}^{n\times n}$.

Due to Schur's Theorem, this is equivalent to finding an orthonormal basis $q_1, \ldots, q_n$ of $\mathbb{C}^n$ such that $q_1, \ldots, q_j$ is an orthonormal basis of an invariant subspace of $\mathbb{C}^n$ with respect to A for each $j = 1, \ldots, n-1$.

To find it start with an orthonormal basis $q_1^{(0)}, \ldots, q_n^{(0)}$ of $\mathbb{C}^n$ and perform the following iterations.

QR algorithm and Simultaneous iterations

The goal in the QR Algorithm is to find a Schur form the square matrix $A \in \mathbb{C}^{n\times n}$.

Due to Schur's Theorem, this is equivalent to finding an orthonormal basis $q_1, \ldots, q_n$ of $\mathbb{C}^n$ such that $q_1, \ldots, q_j$ is an orthonormal basis of an invariant subspace of $\mathbb{C}^n$ with respect to A for each $j = 1, \ldots, n-1$.

To find it start with an orthonormal basis $q_1^{(0)}, \ldots, q_n^{(0)}$ of $\mathbb{C}^n$ and perform the following iterations.

for $m = 1, 2, \ldots,$

(i) Compute $Aq_1^{(m)}, \ldots, Aq_n^{(m)}$;

(ii) Orthonormalize $Aq_1^{(m)}, \ldots, Aq_n^{(m)}$ *to form an orthonormal basis* $q_1^{(m+1)}, \ldots, q_n^{(m+1)}$ *of* $\mathbb{C}^n$.

QR algorithm and Simultaneous iterations

The goal in the QR Algorithm is to find a Schur form the square matrix $A \in \mathbb{C}^{n \times n}$.

Due to Schur's Theorem, this is equivalent to finding an orthonormal basis $q_1, \ldots, q_n$ of $\mathbb{C}^n$ such that $q_1, \ldots, q_j$ is an orthonormal basis of an invariant subspace of $\mathbb{C}^n$ with respect to A for each $j = 1, \ldots, n-1$.

To find it start with an orthonormal basis $q_1^{(0)}, \ldots, q_n^{(0)}$ of $\mathbb{C}^n$ and perform the following iterations.

for $m = 1, 2, \ldots,$

(i) Compute $Aq_1^{(m)}, \ldots, Aq_n^{(m)}$;

(ii) Orthonormalize $Aq_1^{(m)}, \ldots, Aq_n^{(m)}$ *to form an orthonormal basis* $q_1^{(m+1)}, \ldots, q_n^{(m+1)}$ *of* $\mathbb{C}^n$.

Suppose $\mathcal{S}_j = \mathrm{span}\{q_1^{(0)}, \ldots, q_j^{(0)}\}, \mathcal{U}_j = \mathrm{span}\{v_{j+1}, \ldots, v_n\}$, for $j = 1, \ldots, n-1$. The required basis will be computed if

$$|\lambda_j| > |\lambda_{j+1}| \text{ and } \mathcal{S}_j \cap \mathcal{U}_j = \{0\}$$

for each $j = 1, 2, \ldots, n-1$.

QR algorithm and Simultaneous iterations

To check the progress of the simultaneous iterations set $Q_m = \begin{bmatrix} q_1^{(m)} & \cdots & q_n^{(m)} \end{bmatrix}$ and check if $A_m := Q_m^* A Q_m$ is getting close to an upper triangular matrix.

QR algorithm and Simultaneous iterations

To check the progress of the simultaneous iterations set $Q_m = \begin{bmatrix} q_1^{(m)} & \cdots & q_n^{(m)} \end{bmatrix}$ and check if $A_m := Q_m^* A Q_m$ is getting close to an upper triangular matrix.

Suppose that the initial orthonormal basis is $e_1, \ldots, e_n$.

QR algorithm and Simultaneous iterations

To check the progress of the simultaneous iterations set $Q_m = \left[\begin{array}{ccc} q_1^{(m)} & \cdots & q_n^{(m)} \end{array}\right]$ and check if $A_m := Q_m^* A Q_m$ is getting close to an upper triangular matrix.

Suppose that the initial orthonormal basis is $e_1, \ldots, e_n$.

In iteration 1

(i) find $Ae_1, \ldots, Ae_n$, that is consider A;

(ii) orthonormalize by finding the Q, say Q_0, of a QR decomposition of A.

QR algorithm and Simultaneous iterations

To check the progress of the simultaneous iterations set $Q_m = \begin{bmatrix} q_1^{(m)} & \cdots & q_n^{(m)} \end{bmatrix}$ and check if $A_m := Q_m^* A Q_m$ is getting close to an upper triangular matrix.

Suppose that the initial orthonormal basis is $e_1, \ldots, e_n$.

In iteration 1

(i) find $Ae_1, \ldots, Ae_n$, that is consider A;

(ii) orthonormalize by finding the Q, say Q_0, of a QR decomposition of A.

Feeling very optimistic about convergence, we find $A_1 := Q_0^* A Q_0$.

QR algorithm and Simultaneous iterations

To check the progress of the simultaneous iterations set $Q_m = \begin{bmatrix} q_1^{(m)} & \cdots & q_n^{(m)} \end{bmatrix}$ and check if $A_m := Q_m^* A Q_m$ is getting close to an upper triangular matrix.

Suppose that the initial orthonormal basis is $e_1, \ldots, e_n$.

In iteration 1

(i) find $Ae_1, \ldots, Ae_n$, that is consider A;

(ii) orthonormalize by finding the Q, say Q_0, of a QR decomposition of A.

Feeling very optimistic about convergence, we find $A_1 := Q_0^* A Q_0$.

By now we have completed the first iteration of the QR algorithm!

QR algorithm and Simultaneous iterations

To check the progress of the simultaneous iterations set $Q_m = \begin{bmatrix} q_1^{(m)} & \cdots & q_n^{(m)} \end{bmatrix}$ and check if $A_m := Q_m^* A Q_m$ is getting close to an upper triangular matrix.

Suppose that the initial orthonormal basis is $e_1, \ldots, e_n$.

In iteration 1

(i) find $Ae_1, \ldots, Ae_n$, that is consider A;

(ii) orthonormalize by finding the Q, say Q_0, of a QR decomposition of A.

Feeling very optimistic about convergence, we find $A_1 := Q_0^* A Q_0$.

By now we have completed the first iteration of the QR algorithm!

At this point we have the choice of performing the next iteration of simultaneous iteration either with respect to the standard basis $e_1, \ldots, e_n$ or the orthonormal basis formed by the columns say $q_1, \ldots, q_n$ of Q_0. If we go with $e_1, \ldots, e_n$, the operations are

QR algorithm and Simultaneous iterations

To check the progress of the simultaneous iterations set $Q_m = \begin{bmatrix} q_1^{(m)} & \cdots & q_n^{(m)} \end{bmatrix}$ and check if $A_m := Q_m^* A Q_m$ is getting close to an upper triangular matrix.

Suppose that the initial orthonormal basis is $e_1, \ldots, e_n$.

In iteration 1

(i) find $Ae_1, \ldots, Ae_n$, that is consider A;

(ii) orthonormalize by finding the Q, say Q_0, of a QR decomposition of A.

Feeling very optimistic about convergence, we find $A_1 := Q_0^* A Q_0$.

By now we have completed the first iteration of the QR algorithm!

At this point we have the choice of performing the next iteration of simultaneous iteration either with respect to the standard basis $e_1, \ldots, e_n$ or the orthonormal basis formed by the columns say $q_1, \ldots, q_n$ of Q_0. If we go with $e_1, \ldots, e_n$, the operations are

(i) find $Aq_1, \ldots, Aq_n$;

(ii) compute the Q of a QR decomposition of AQ_0.

QR algorithm and Simultaneous iterations

If we choose the basis of the columns of Q_0, A and $q_1, \ldots, q_n$ get replaced by their representations in the new basis.

QR algorithm and Simultaneous iterations

If we choose the basis of the columns of Q_0, A and $q_1, \ldots, q_n$ get replaced by their representations in the new basis.

So the iteration is

(i) find $Q_0^* A Q_0 (Q_0^* q_1)[= A_1 e_1], \ldots, Q_0^* A Q_0 (Q_0^* q_n)[= A_1 e_n]$, that is, consider A_1;

(ii) compute the Q, say Q_1, of a QR decomposition of A_1.

QR algorithm and Simultaneous iterations

If we choose the basis of the columns of Q_0, A and $q_1, \ldots, q_n$ get replaced by their representations in the new basis.

So the iteration is

(i) find $Q_0^* A Q_0 (Q_0^* q_1) [= A_1 e_1], \ldots, Q_0^* A Q_0 (Q_0^* q_n) [= A_1 e_n]$, that is, consider A_1;

(ii) compute the Q, say Q_1, of a QR decomposition of A_1.

Feeling very optimistic about the progress, we form $A_2 = Q_1^* A_1 Q_1$.

QR algorithm and Simultaneous iterations

If we choose the basis of the columns of Q_0, A and $q_1, \ldots, q_n$ get replaced by their representations in the new basis.

So the iteration is

(i) find $Q_0^* A Q_0 (Q_0^* q_1)[= A_1 e_1], \ldots, Q_0^* A Q_0 (Q_0^* q_n)[= A_1 e_n]$, that is, consider A_1;

(ii) compute the Q, say Q_1, of a QR decomposition of A_1.

Feeling very optimistic about the progress, we form $A_2 = Q_1^* A_1 Q_1$.

This completes the second iteration of the QR algorithm!

QR algorithm and Simultaneous iterations

If we choose the basis of the columns of Q_0, A and $q_1, \ldots, q_n$ get replaced by their representations in the new basis.

So the iteration is

(i) find $Q_0^* A Q_0 (Q_0^* q_1) [= A_1 e_1], \ldots, Q_0^* A Q_0 (Q_0^* q_n) [= A_1 e_n]$, that is, consider A_1;

(ii) compute the Q, say Q_1, of a QR decomposition of A_1.

Feeling very optimistic about the progress, we form $A_2 = Q_1^* A_1 Q_1$.

This completes the second iteration of the QR algorithm!

Now we have the choice of performing the next iteration in three different bases, viz., $e_1, \ldots, e_n$, the columns of Q_0 or the columns of Q_1. If we choose the last option then the third iteration is

QR algorithm and Simultaneous iterations

If we choose the basis of the columns of Q_0, A and $q_1, \ldots, q_n$ get replaced by their representations in the new basis.

So the iteration is

(i) find $Q_0^* A Q_0 (Q_0^* q_1) [= A_1 e_1], \ldots, Q_0^* A Q_0 (Q_0^* q_n) [= A_1 e_n]$, that is, consider A_1;

(ii) compute the Q, say Q_1, of a QR decomposition of A_1.

Feeling very optimistic about the progress, we form $A_2 = Q_1^* A_1 Q_1$.

This completes the second iteration of the QR algorithm!

Now we have the choice of performing the next iteration in three different bases, viz., $e_1, \ldots, e_n$, the columns of Q_0 or the columns of Q_1. If we choose the last option then the third iteration is

(i) find $Q_1^* A_1 Q_1 (Q_1^* Q_1 e_1), \ldots, Q_1^* A_1 Q_1 (Q_1^* Q_1 e_n)$, i.e., consider A_2;

(ii) compute the Q, say Q_2, of a QR decomposition of A_2.

QR algorithm and Simultaneous iterations

If we choose the basis of the columns of Q_0, A and $q_1, \ldots, q_n$ get replaced by their representations in the new basis.

So the iteration is

(i) find $Q_0^* A Q_0 (Q_0^* q_1) [= A_1 e_1], \ldots, Q_0^* A Q_0 (Q_0^* q_n) [= A_1 e_n]$, that is, consider A_1;

(ii) compute the Q, say Q_1, of a QR decomposition of A_1.

Feeling very optimistic about the progress, we form $A_2 = Q_1^* A_1 Q_1$.

This completes the second iteration of the QR algorithm!

Now we have the choice of performing the next iteration in three different bases, viz., $e_1, \ldots, e_n$, the columns of Q_0 or the columns of Q_1. If we choose the last option then the third iteration is

(i) find $Q_1^* A_1 Q_1 (Q_1^* Q_1 e_1), \ldots, Q_1^* A_1 Q_1 (Q_1^* Q_1 e_n)$, i.e., consider A_2;

(ii) compute the Q, say Q_2, of a QR decomposition of A_2.

To check for convergence form $A_3 := Q_2^* A_2 Q_2$.
This the third iteration of the QR algorithm!

QR algorithm and Simultaneous iterations

If we choose the basis of the columns of Q_0, A and $q_1, \ldots, q_n$ get replaced by their representations in the new basis.

So the iteration is

(i) find $Q_0^* A Q_0 (Q_0^* q_1) [= A_1 e_1], \ldots, Q_0^* A Q_0 (Q_0^* q_n) [= A_1 e_n]$, that is, consider A_1;

(ii) compute the Q, say Q_1, of a QR decomposition of A_1.

Feeling very optimistic about the progress, we form $A_2 = Q_1^* A_1 Q_1$.

This completes the second iteration of the QR algorithm!

Now we have the choice of performing the next iteration in three different bases, viz., $e_1, \ldots, e_n$, the columns of Q_0 or the columns of Q_1. If we choose the last option then the third iteration is

(i) find $Q_1^* A_1 Q_1 (Q_1^* Q_1 e_1), \ldots, Q_1^* A_1 Q_1 (Q_1^* Q_1 e_n)$, i.e., consider A_2;

(ii) compute the Q, say Q_2, of a QR decomposition of A_2.

To check for convergence form $A_3 := Q_2^* A_2 Q_2$.
This the third iteration of the QR algorithm!

Hence QR algorithm executes Simultaneous iteration with suitable change of basis at each iteration.

John Francis

John Francis

Your final task

Your final task

Exercise: Identify David Watkins in this picture taken at IITG in January 2013!

Your final task

Exercise: Identify David Watkins in this picture taken at IITG in January 2013!