Infinite dimensional spectral computations & linear algebra: Extending the QR algorithm to infinite dimensions

Matthew Colbrook

May 30, 2018

Outline

- Background
- Introducing IQR
- Non Normal Operators
- How To Compute
- Numerical Examples
- Conclusion

Background

Background

- Hilbert space $l^2(\mathbb{N})$ with $\|x\|_2 = \sqrt{\sum_{j=1}^{\infty} |x_j|^2}$, $\langle x, y \rangle = \sum_{j=1}^{\infty} x_j \bar{y}_j$
- Bounded linear operator $T : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ realised as matrix

$$\begin{pmatrix} t_{11} & t_{12} & t_{13} & \cdots \\ t_{21} & t_{22} & t_{23} & \cdots \\ t_{31} & t_{32} & t_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

  Denote these by $\mathcal{B}(l^2(\mathbb{N}))$.
- Want to compute spectrum (generalistion of eigenvalues)

$$\sigma(T) := \{z \in \mathbb{C} : T - zI \text{ not invertible}\}.$$

  from the matrix elements. What about eigenvectors etc.?

Well Studied

- Quantum mechanics, quasicrystals

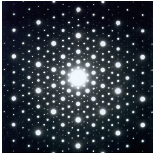



Figure: Left: Dan Shechtman, Nobel Prize in Chemistry 2011. Right: Electron diffraction pattern of quasicrystal.

- Intensely investigated since the 1950s, still very active today.

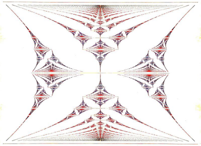


Figure: Left: Artur Avila, Fields Medal 2014. Right: Hofstadter butterfly.

Hierarchy of complexity

Definition (Tower of Algorithms)

A tower of algorithms of height k is a family of sequences of functions

$$\Gamma_{n_k,\ldots,n_1} : \Omega \to \mathcal{M},$$

where $n_k, \ldots, n_1 \in \mathbb{N}$ and $\Gamma_{n_k,\ldots,n_1}$ are "algorithms". Moreover,

$$\sigma(T) = \lim_{n_k \to \infty} \ldots \lim_{n_1 \to \infty} \Gamma_{n_k,\ldots,n_1}(T).$$

Hierarchy of complexity

Definition (Tower of Algorithms)

A tower of algorithms of height k is a family of sequences of functions

$$\Gamma_{n_k,\ldots,n_1} : \Omega \to \mathcal{M},$$

where $n_k, \ldots, n_1 \in \mathbb{N}$ and $\Gamma_{n_k,\ldots,n_1}$ are "algorithms". Moreover,

$$\sigma(T) = \lim_{n_k \to \infty} \ldots \lim_{n_1 \to \infty} \Gamma_{n_k,\ldots,n_1}(T).$$

Definition (Solvability Complexity Index (SCI))

Solvability Complexity Index, $\mathrm{SCI}(\sigma, \Omega)$ is the smallest integer k for which there exists a tower of algorithms of height k. If no such tower exists then $\mathrm{SCI}(\sigma, \Omega) = \infty$.

Example Results

1. General spectral problem has $\mathrm{SCI} = 3$.

Example Results

1. General spectral problem has $\mathrm{SCI} = 3$.
2. Self-adjoint/Normal spectral problem has $\mathrm{SCI} = 2$.

Example Results

1. General spectral problem has $\mathrm{SCI} = 3$.
2. Self-adjoint/Normal spectral problem has $\mathrm{SCI} = 2$.
3. Compact spectral problem has $\mathrm{SCI} = 1$ but no error control.

Example Results

1. General spectral problem has SCI $= 3$.
2. Self-adjoint/Normal spectral problem has SCI $= 2$.
3. Compact spectral problem has SCI $= 1$ but no error control.
4. Splitting the discrete spectrum from essential spectrum (as sets) in generally harder.

Example Results

1. General spectral problem has SCI $= 3$.
2. Self-adjoint/Normal spectral problem has SCI $= 2$.
3. Compact spectral problem has SCI $= 1$ but no error control.
4. Splitting the discrete spectrum from essential spectrum (as sets) in generally harder.

Methods based on approximating pseudospectrum:

$$\sigma_\epsilon(T) = \{z : \left\|(T - zI)^{-1}\right\| \geq \epsilon^{-1}\},$$

where we interpret $\left\|S^{-1}\right\|$ as $+\infty$ if S does not have a bounded inverse.

Motivation

1. Above method can't detect isolated eigenvalues and their multiplicity.

Motivation

1. Above method can't detect isolated eigenvalues and their multiplicity.
2. Due to taking square root, above method can only gain precision $\sqrt{\epsilon_{\text{mach}}}$.

Motivation

1. Above method can't detect isolated eigenvalues and their multiplicity.
2. Due to taking square root, above method can only gain precision $\sqrt{\epsilon_{\text{mach}}}$.
3. Can we generalise staple finite matrix algorithms to infinite dimensions?

Motivation

1. Above method can't detect isolated eigenvalues and their multiplicity.
2. Due to taking square root, above method can only gain precision $\sqrt{\epsilon_{\text{mach}}}$.
3. Can we generalise staple finite matrix algorithms to infinite dimensions?
4. Can we gain error control and classification results in the hierarchy?

Classical QR

$$T = Q_1 R_1$$

Classical QR

$$\begin{aligned} T &= Q_1 R_1 \\ T_1 &= R_1 Q_1 = Q_2 R_2 \end{aligned}$$

Classical QR

$$
\begin{aligned}
T &= Q_1 R_1 \\
T_1 &= R_1 Q_1 = Q_2 R_2 \\
T_2 &= R_2 Q_2 = Q_3 R_3
\end{aligned}
$$

Classical QR

$$
\begin{aligned}
T &= Q_1 R_1 \\
T_1 &= R_1 Q_1 = Q_2 R_2 \\
T_2 &= R_2 Q_2 = Q_3 R_3 \\
&\vdots \\
T_n &= Q_n^* ... Q_1^* T Q_1 ... Q_n
\end{aligned}
$$

Classical Result

Theorem

Let $T \in \mathbb{C}^{N \times N}$ be a normal matrix with eigenvalues satisfying $|\lambda_1| > \ldots > |\lambda_N|$. Let $\{Q_m\}$ be a Q-sequence of unitary operators. Then (up to re-ordering of the basis)

$$Q_m^* T Q_m \longrightarrow \bigoplus_{j=1}^{N} \lambda_j e_j \otimes e_j, \qquad m \to \infty.$$

Classical Result

Theorem

Let $T \in \mathbb{C}^{N\times N}$ be a normal matrix with eigenvalues satisfying $|\lambda_1| > \ldots > |\lambda_N|$. Let $\{Q_m\}$ be a Q-sequence of unitary operators. Then (up to re-ordering of the basis)

$$Q_m^* T Q_m \longrightarrow \bigoplus_{j=1}^{N} \lambda_j e_j \otimes e_j, \qquad m \to \infty.$$

Numerical Example ...

Introducing IQR

The Main Idea

$$
\begin{pmatrix} t_{11} & t_{12} & t_{13} & \cdots \\ t_{21} & t_{22} & t_{23} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \overset{\text{Truncate}}{\Rightarrow} \begin{pmatrix} t_{11} & t_{12} & t_{13} & \cdots & t_{1n} \\ t_{21} & t_{22} & t_{23} & \cdots & t_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & t_{n3} & \cdots & t_{nn} \end{pmatrix} \overset{\text{Iterate QR}}{\Rightarrow} \begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix}
$$

The Main Idea

$$\begin{pmatrix} t_{11} & t_{12} & t_{13} & \cdots \\ t_{21} & t_{22} & t_{23} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \xRightarrow{\text{Truncate}} \begin{pmatrix} t_{11} & t_{12} & t_{13} & \cdots & t_{1n} \\ t_{21} & t_{22} & t_{23} & \cdots & t_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & t_{n3} & \cdots & t_{nn} \end{pmatrix} \xRightarrow{\text{Iterate QR}} \begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix}$$

The Main Idea

$$\begin{pmatrix} t_{11} & t_{12} & t_{13} & \cdots \\ t_{21} & t_{22} & t_{23} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \xRightarrow{\text{Truncate}} \begin{pmatrix} t_{11} & t_{12} & t_{13} & \cdots & t_{1n} \\ t_{21} & t_{22} & t_{23} & \cdots & t_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & t_{n3} & \cdots & t_{nn} \end{pmatrix} \xRightarrow{\text{Iterate QR}} \begin{pmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{pmatrix}$$

$$\begin{pmatrix} t_{11} & t_{12} & t_{13} & \cdots \\ t_{21} & t_{22} & t_{23} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \xRightarrow{\text{Iterate QR?}} \begin{pmatrix} \tilde{d}_{11} & 0 & \cdots \\ 0 & \tilde{d}_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \xRightarrow{\text{Truncate}} \begin{pmatrix} \tilde{d}_{11} & 0 & \cdots & 0 \\ 0 & \tilde{d}_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{d}_{nn} \end{pmatrix}$$

Questions

1. Does the QR algorithm exist in infinite dimensions?

Questions

1. Does the QR algorithm exist in infinite dimensions?
2. When do we gain convergence to a diagonal operator and in what sense?

Questions

1. Does the QR algorithm exist in infinite dimensions?
2. When do we gain convergence to a diagonal operator and in what sense?
3. Can we prove anything for non normal operators in infinite dimensions?

Questions

1. Does the QR algorithm exist in infinite dimensions?
2. When do we gain convergence to a diagonal operator and in what sense?
3. Can we prove anything for non normal operators in infinite dimensions?
4. Can we even compute this beast on a finite machine?

The QR Decomposition

Definition

A Householder reflection is an operator $S \in \mathcal{B}(\mathcal{H})$ of the form

$$S = I - \frac{2}{\|\psi\|^2}\psi \otimes \bar{\psi}, \qquad \psi \in \mathcal{H},$$

where $\bar{\psi}$ denotes the associated functional in $\mathcal{H}^*$ given by $x \to \langle x, \psi \rangle$.

The QR Decomposition

Definition

A Householder reflection is an operator $S \in \mathcal{B}(\mathcal{H})$ of the form

$$S = I - \frac{2}{\|\psi\|^2} \psi \otimes \bar{\psi}, \qquad \psi \in \mathcal{H},$$

where $\bar{\psi}$ denotes the associated functional in $\mathcal{H}^*$ given by $x \to \langle x, \psi \rangle$.In the case where $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ and I_i is the identity on $\mathcal{H}_i$ then

$$U = I_1 \oplus \left(I_2 - \frac{2}{\|\psi\|^2} \psi \otimes \bar{\psi} \right) \qquad \psi \in \mathcal{H}_2,$$

is called a Householder transformation.

The QR Decomposition

Theorem (Hansen 2008)

Let T be a bounded operator on a separable Hilbert space $\mathcal{H}$ and let $\{e_j\}_{j\in\mathbb{N}}$ be an orthonormal basis for $\mathcal{H} \cong l^2(\mathbb{N})$. Then there exist an isometry Q such that $T = QR$ where R is upper triangular with respect to $\{e_j\}$. Moreover,

$$Q = \operatorname*{SOT-lim}_{n\to\infty} V_n$$

where $V_n = U_1 \cdots U_n$ are unitary and each U_j is a Householder transformation.

The QR Decomposition

Theorem (Hansen 2008)

Let T be a bounded operator on a separable Hilbert space $\mathcal{H}$ and let $\{e_j\}_{j\in\mathbb{N}}$ be an orthonormal basis for $\mathcal{H} \cong l^2(\mathbb{N})$. Then there exist an isometry Q such that $T = QR$ where R is upper triangular with respect to $\{e_j\}$. Moreover,

$$Q = \operatorname*{SOT-lim}_{n\to\infty} V_n$$

where $V_n = U_1 \cdots U_n$ are unitary and each U_j is a Householder transformation.

Good for numerics - more stable than Gram-Schmidt...

What's Really Going On?

Assume T invertible...

$$T = Q_1 R_1 = \hat{Q}_1 \hat{R}_1,$$

What's Really Going On?

Assume T invertible...

$$T = Q_1 R_1 = \hat{Q}_1 \hat{R}_1,$$

$$T^2 = Q_1 R_1 Q_1 R_1 = Q_1 Q_2 R_2 R_1 = \hat{Q}_2 \hat{R}_2,$$

What's Really Going On?

Assume T invertible...

$$T = Q_1 R_1 = \hat{Q}_1 \hat{R}_1,$$

$$T^2 = Q_1 R_1 Q_1 R_1 = Q_1 Q_2 R_2 R_1 = \hat{Q}_2 \hat{R}_2,$$

$$T^3 = Q_1 R_1 Q_1 R_1 Q_1 R_1 = Q_1 Q_2 R_2 Q_2 R_2 R_1 = Q_1 Q_2 Q_3 R_3 R_2 R_1 = \hat{Q}_3 \hat{R}_3.$$

What's Really Going On?

Assume T invertible...

$$T = Q_1 R_1 = \hat{Q}_1 \hat{R}_1,$$

$$T^2 = Q_1 R_1 Q_1 R_1 = Q_1 Q_2 R_2 R_1 = \hat{Q}_2 \hat{R}_2,$$

$$T^3 = Q_1 R_1 Q_1 R_1 Q_1 R_1 = Q_1 Q_2 R_2 Q_2 R_2 R_1 = Q_1 Q_2 Q_3 R_3 R_2 R_1 = \hat{Q}_3 \hat{R}_3.$$

An easy induction gives us that

$$T^m = \hat{Q}_m \hat{R}_m.$$

What's Really Going On?

Assume T invertible...

$$T = Q_1 R_1 = \hat{Q}_1 \hat{R}_1,$$

$$T^2 = Q_1 R_1 Q_1 R_1 = Q_1 Q_2 R_2 R_1 = \hat{Q}_2 \hat{R}_2,$$

$$T^3 = Q_1 R_1 Q_1 R_1 Q_1 R_1 = Q_1 Q_2 R_2 Q_2 R_2 R_1 = Q_1 Q_2 Q_3 R_3 R_2 R_1 = \hat{Q}_3 \hat{R}_3.$$

An easy induction gives us that

$$T^m = \hat{Q}_m \hat{R}_m.$$

$\hat{R}_m$ upper triangular since R_j, $j \le m$ are.

What's Really Going On?

Assume T invertible...

$$T = Q_1 R_1 = \hat{Q}_1 \hat{R}_1,$$

$$T^2 = Q_1 R_1 Q_1 R_1 = Q_1 Q_2 R_2 R_1 = \hat{Q}_2 \hat{R}_2,$$

$$T^3 = Q_1 R_1 Q_1 R_1 Q_1 R_1 = Q_1 Q_2 R_2 Q_2 R_2 R_1 = Q_1 Q_2 Q_3 R_3 R_2 R_1 = \hat{Q}_3 \hat{R}_3.$$

An easy induction gives us that

$$T^m = \hat{Q}_m \hat{R}_m.$$

$\hat{R}_m$ upper triangular since R_j, $j \leq m$ are. By invertibility of T, $\langle Re_i, e_i \rangle \neq 0$. Hence

$$\mathrm{span}\{T^m e_j\}_{j=1}^{J} = \mathrm{span}\{\hat{Q}_m e_j\}_{j=1}^{J}, \quad J \in \mathbb{N}.$$

Questions

1. Does the QR algorithm exist in infinite dimensions? ✓
2. When do we gain convergence to a diagonal operator and in what sense?
3. Can we prove anything for non normal operators in infinite dimensions?
4. Can we even compute this beast on a finite machine?

A Result for Normal Operators

Assume the following:

(A1) $T \in \mathcal{B}(\mathcal{H})$ is an invertible normal operator and $\{e_j\}_{j\in\mathbb{N}}$ an orthonormal basis for $\mathcal{H}$. $\{Q_k\}$ and $\{R_k\}$ are Q- and R-sequences of T with respect to the basis $\{e_j\}_{j\in\mathbb{N}}$.

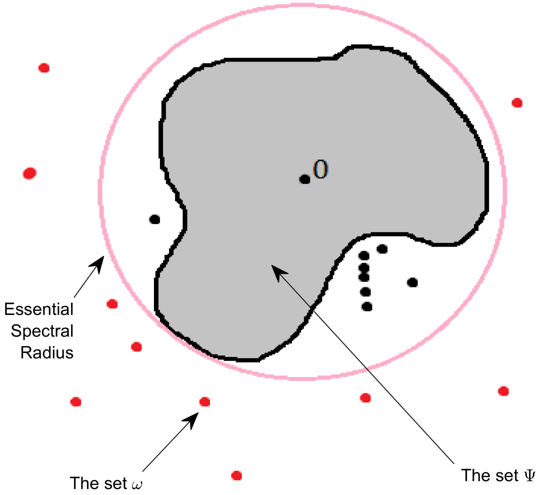
(A2) $\sigma(T) = \omega \cup \Psi$ such that $\omega \cap \Psi = \emptyset$ and $\omega = \{\lambda_i\}_{i=1}^{N}$, where the λ_is are isolated eigenvalues with (possibly infinite) multiplicity m_i. Let $M = m_1 + ... + m_N = \dim(\mathrm{ran}\chi_\omega(T))$ and suppose that $|\lambda_1| > \ldots > |\lambda_N|$. Suppose further that $\sup\{|\theta| : \theta \in \Psi\} < |\lambda_N|$.

Define

$$\rho = \sup\{|z| : z \in \Psi\}, \quad r = \max\{|\lambda_2/\lambda_1|, ..., |\lambda_N/\lambda_{N-1}|, \rho/|\lambda_N|\}$$

then $r < 1$.

What This Really Means!



A Result for Normal Operators

Theorem

There exists $\{\hat{e}_j\}_{j=1}^M \subset \{e_j\}_{j\in\mathbb{N}}$, where $M = m_1 + \ldots + m_N$, so that $\operatorname{span}\{Q_k \hat{e}_j\} \to \operatorname{span}\{\hat{q}_j\}$ where $\{\hat{q}_j\}_{j=1}^M \subset \operatorname{ran}\chi_\omega(T)$ is a collection of orthonormal eigenvectors of T and if $e_j \notin \{\hat{e}_j\}_{j=1}^M$, then $\chi_\omega(T) Q_k e_j \to 0$. Also:

(i) Every subsequence of $\{Q_n^* T Q_n\}_{n\in\mathbb{N}}$ has a convergent subsequence $\{Q_{n_k}^* T Q_{n_k}\}_{k\in\mathbb{N}}$ such that

$$Q_{n_k}^* T Q_{n_k} \xrightarrow{\text{WOT}} \left(\bigoplus_{j=1}^M \langle T\hat{q}_j, \hat{q}_j\rangle \hat{e}_j \otimes \hat{e}_j\right) \oplus \sum_{j\in\Theta} \xi_j \otimes e_j,$$

as $k \to \infty$, where

$$\Theta = \{j : e_j \notin \{\hat{e}_j\}_{j=1}^M\}, \quad \xi_j \in \overline{\operatorname{span}\{e_i\}_{i\in\Theta}}$$

and only $\sum_{j\in\Theta} \xi_j \otimes e_j$ depends on the choice of subsequence.

(ii)

$$\widehat{P}_M Q_n^* T Q_n \widehat{P}_M \xrightarrow{\text{SOT}} \left(\bigoplus_{j=1}^M \langle T\hat{q}_j, \hat{q}_j\rangle \hat{e}_j \otimes \hat{e}_j\right), \qquad \text{as } n \to \infty,$$

where $\widehat{P}_M$ denotes the orthogonal projection onto $\overline{\operatorname{span}\{\hat{e}_j\}_{j=1}^M}$. For any fixed $x \in \operatorname{span}\{\hat{e}_j\}_{j=1}^M$ we have the following rate of convergence

$$\left\|\widehat{P}_M Q_n^* T Q_n \widehat{P}_M x - \left(\bigoplus_{j=1}^M \langle T\hat{q}_j, \hat{q}_j\rangle \hat{e}_j \otimes \hat{e}_j\right) x\right\| \le \mathcal{O}(r^n).$$

Idea of Proof

If we keep applying T to $\text{span}\{e_j\}_{j=1}^{K}$, expect the span of vectors will approach the extreme parts of spectrum (projection sense).

Idea of Proof

If we keep applying T to $\text{span}\{e_j\}_{j=1}^K$, expect the span of vectors will approach the extreme parts of spectrum (projection sense).
Use spectral decomposition:

$$T = \left(\sum_{j=1}^{M} \lambda_j \, \xi_j \otimes \bar{\xi}_j \right) \oplus \chi_\Psi(T) T, \quad \lambda_j \in \omega,$$

where $\{\xi_j\}_{j=1}^M$ is an orthonormal set of eigenvectors of T to make this precise.

Idea of Proof

If we keep applying T to span$\{e_j\}_{j=1}^{K}$, expect the span of vectors will approach the extreme parts of spectrum (projection sense).
Use spectral decomposition:

$$T = \left(\sum_{j=1}^{M} \lambda_j \, \xi_j \otimes \bar{\xi}_j \right) \oplus \chi_{\Psi}(T) T, \quad \lambda_j \in \omega,$$

where $\{\xi_j\}_{j=1}^{M}$ is an orthonormal set of eigenvectors of T to make this precise.
Finally use

$$T^m = \hat{Q}_m \hat{R}_m.$$

Idea of Proof

If we keep applying T to $\mathrm{span}\{e_j\}_{j=1}^{K}$, expect the span of vectors will approach the extreme parts of spectrum (projection sense).
Use spectral decomposition:

$$T = \left(\sum_{j=1}^{M} \lambda_j \, \xi_j \otimes \bar{\xi}_j \right) \oplus \chi_{\Psi}(T)T, \quad \lambda_j \in \omega,$$

where $\{\xi_j\}_{j=1}^{M}$ is an orthonormal set of eigenvectors of T to make this precise.
Finally use

$$T^m = \hat{Q}_m \hat{R}_m.$$

Can upgrade for block convergence (eigenvalues not of distinct magnitude), SOT convergence etc.

Questions

1. Does the QR algorithm exist in infinite dimensions? ✓
2. When do we gain convergence to a diagonal operator and in what sense? ✓
3. Can we prove anything for non normal operators in infinite dimensions?
4. Can we even compute this beast on a finite machine?

Non Normal Operators

Set Up

Assume the following (M now finite):

(A1) $T \in \mathcal{B}(\mathcal{H})$ is an invertible operator and there is an orthogonal projection P of finite rank M with image invariant under T.

(A2) There exist $\alpha > \beta > 0$ such that

$$\|Tx\| \geq \alpha\|x\| \quad \forall x \in \mathrm{ran}(P),$$
$$\|(I-P)T(I-P)\| \leq \beta.$$

(A3) $\{\tilde{P}e_j\}_{j=1}^{M}$ are linearly independent ($\tilde{P}$ canoncal defined later).

Define for orthogonal projections E, F onto $S_1, S_2 \subset \mathcal{H}$ the distance between subspaces

$$\hat{\delta}(S_1, S_2) = \|E - F\| \in [0, 1]$$

and the subspace angle

$$\phi(S_1, S_2) = \sin^{-1}\left(\hat{\delta}(S_1, S_2)\right).$$

Result

Theorem

There exists a canonical M dimensional T^*-invariant subspace S with orthogonal projection $\tilde{P}$ with

(i) The subspace angle $\phi(\text{span}\{e_j\}_{j=1}^M, S) < \pi/2$ and we have

$$\hat{\delta}(\text{span}\{Q_n e_j\}_{j=1}^M, \text{ran}(P)) \leq \frac{\sin\left(\phi(\text{span}\{e_j\}_{j=1}^M, \text{ran}(P))\right)}{\cos\left(\phi(\text{span}\{e_j\}_{j=1}^M, S)\right)} \left(1 + \frac{\|PT(I-P)\|}{\alpha - \beta}\right) \frac{\beta^n}{\alpha^n},$$

(ii) Every subsequence of $\{Q_n^* T Q_n\}_{n \in \mathbb{N}}$ has a convergent subsequence $\{Q_{n_k}^* T Q_{n_k}\}_{k \in \mathbb{N}}$ such that

$$Q_{n_k}^* T Q_{n_k} \xrightarrow{\text{WOT}} \sum_{j=1}^{M} \xi_j \otimes e_j \bigoplus \sum_{i=M+1}^{\infty} \zeta_i \otimes e_i,$$

$$\xi_j \in \overline{\text{span}\{e_j\}_{j=1}^M}, \quad \zeta_i \in \mathcal{H}.$$

Questions

1. Does the QR algorithm exist in infinite dimensions? ✓
2. When do we gain convergence to a diagonal operator and in what sense? ✓
3. Can we prove anything for non normal operators in infinite dimensions? ✓
4. Can we even compute this beast on a finite machine?

How To Compute

Some Definition

Definition

Let T be an infinite matrix acting as a bounded operator on $l^2(\mathbb{N})$ with basis $\{e_j\}_{j\in\mathbb{N}}$. For $f:\mathbb{N}\to\mathbb{N}$ non-decreasing with $f(n)\geq n$ we say that T has quasi-banded subdiagonals with respect to f if $\langle Te_j, e_i\rangle = 0$ when $i > f(j)$.

A Theorem

Theorem

Let $T \in \mathcal{B}(l^2(\mathbb{N}))$ have quasi-banded subdiagonals with respect to f and let T_n be the n-th element in the QR iteration, i.e. $T_n = Q_n^* \cdots Q_1^* T Q_1 \cdots Q_n$, where

$$Q_j = \operatorname*{SOT-lim}_{l \to \infty} \; U_1^j \cdots U_l^j$$

and U_l^j is a Householder transformation. Let P_m be the usual projection onto $\mathrm{span}\{e_j\}_{j=1}^m$ and denote the a-fold iteration of f by $\underbrace{f \circ f \circ \ldots \circ f}_{a \text{ times}} = f_a$. Then

$$\begin{aligned} P_m T_n P_m = {} & P_m U_m^n \cdots U_1^n U_{f_1(m)}^{n-1} \cdots U_1^{n-1} \cdots U_{f_{(n-2)}(m)}^2 \cdots U_1^2 U_{f_{(n-1)}(m)}^1 \cdots U_1^1 \\ & \cdot P_{f_n(m)} T P_{f_n(m)} \\ & \cdot U_1^1 \cdots U_{f_{(n-1)}(m)}^1 U_1^2 \cdots U_{f_{(n-2)}(m)}^2 \cdots U_1^{n-1} \cdots U_{f_1(m)}^{n-1} U_1^n \cdots U_m^n P_m. \end{aligned}$$

Why?

In the subcase of invertibility, a consequence of the fact that if T has quasi-banded subdiagonals with respect to f then

$$P_m T^n P_m = P_m (P_{f_n(m)} T P_{f_n(m)})^n P_m.$$

Why?

In the subcase of invertibility, a consequence of the fact that if T has quasi-banded subdiagonals with respect to f then

$$P_m T^n P_m = P_m (P_{f_n(m)} T P_{f_n(m)})^n P_m.$$

We can apply Gram-Schmidt (or a more stable modified version) to the columns of $P_{f_n(m)} T P_{f_n(m)}$ and truncate the resulting matrix!

Why?

In the subcase of invertibility, a consequence of the fact that if T has quasi-banded subdiagonals with respect to f then

$$P_m T^n P_m = P_m (P_{f_n(m)} T P_{f_n(m)})^n P_m.$$

We can apply Gram-Schmidt (or a more stable modified version) to the columns of $P_{f_n(m)} T P_{f_n(m)}$ and truncate the resulting matrix!

Can also extend to compute the IQR iterates with error control if we can evaluate an increasing family of increasing functions $g^j : \mathbb{N} \to \mathbb{N}$ such that defining the matrix $T_{(j)}$ with columns $\{P_{g^j(n)} T e_n\}$ we have that $T_{(j)}$ is invertible and

$$\left\| (P_{g^j(n)} - I) T e_n \right\| \leq \frac{1}{j}.$$

Questions

1. Does the QR algorithm exist in infinite dimensions? $\checkmark$
2. When do we gain convergence to a diagonal operator and in what sense? $\checkmark$
3. Can we prove anything for non normal operators in infinite dimensions? $\checkmark$
4. Can we even compute this beast on a finite machine? $\checkmark$

Numerical Examples

Example 1

Take unilateral shift $U : e_n \to e_{n+1}$ acting on $l^2(\mathbb{Z})$ (with the natural choice of indexing $\mathbb{Z}$) and perturb by the compact diagonal operator

$$D(e_n) = \frac{5\sin(n)^2}{\sqrt{|n|+1}} e_n.$$

Hence the spectrum of the full operator $T = U + D$ consists of the unit circle and a collection of eigenvalues in the discrete spectrum.

Example 1

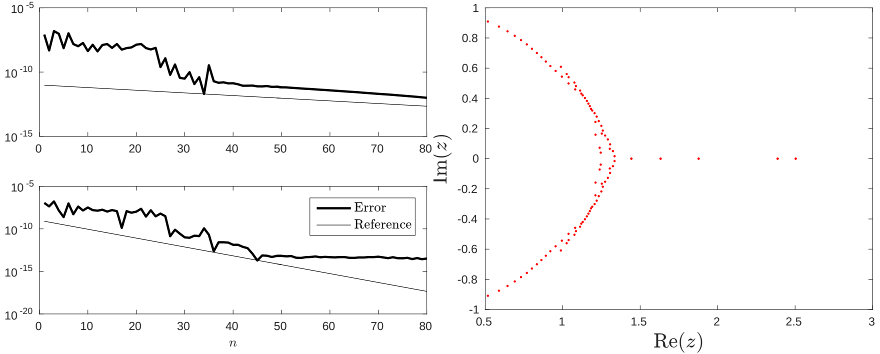


Figure: Left: Error in approximating λ_1 (top) and λ_2 through taking finite sections $P_{100}Q_n^*TQ_n|_{P_{100}\mathcal{H}}$. We have shown the expected rates of convergence as references. Right: The spectrum of $P_{100}Q_{500}^*TQ_{500}|_{P_{100}\mathcal{H}}$ demonstrating convergence to the extremal parts.

Example 2

Almost Mathieu related to a wealth of mathematical/physical problems:

$$(H_1 x)_n = x_{n-1} + x_{n+1} + 2\cos(2\pi n\alpha + \nu)x_n,$$

on $l^2(\mathbb{Z})$. Hamiltonian represents crystal electron in a uniform magnetic field and the spectrum the allowed energies of the system. No discrete spectrum!

Example 2

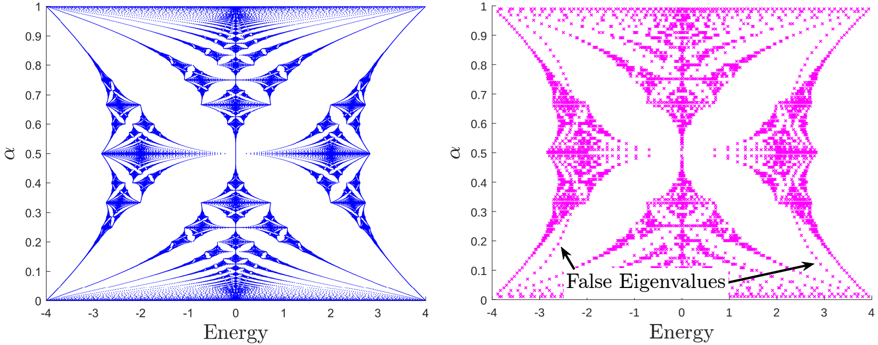


Figure: The spectrum of H_1 calculated analytically for rational α and the output of finite section for $m = 100$. Note that the finite section method causes strong spectral pollution.

Example 2

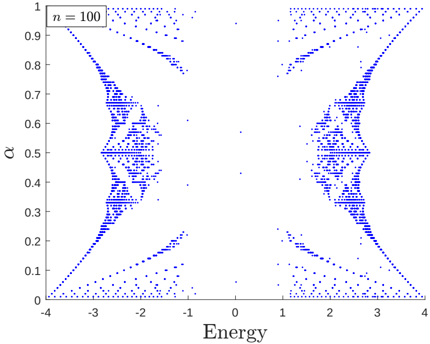


Figure: IQR for $m = 100, n = 100$ and the same for $m = 100, n = 5000$. IQR algorithm is more robust, preserving the structure of the spectrum whilst converging to the boundary of the essential spectrum.

Example 3

Toeplitz operator

$$N = \frac{1}{2}(U_3 + U_{-1})$$

where U_m acts on $l^2(\mathbb{Z})$ by the shift $e_j \to e_{j+m}$ on the standard basis. Not invertible and has no eigenvalues. With no shift, $Q_n^* N Q_n$ appeared

to converge strongly to the operator

$$\tilde{N} = \begin{pmatrix} A & & \\ & A & \\ & & \ddots \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Spectrum equal to $\{\pm 1, \pm 1i\}$ which are the extremal points of $\sigma(N)$. Shift to $Q_n^*(N + I)Q_n$, then we appeared to converge to the diagonal

operator $D = 2I$. BUT curious case of reduced rate of convergence...

Example 3

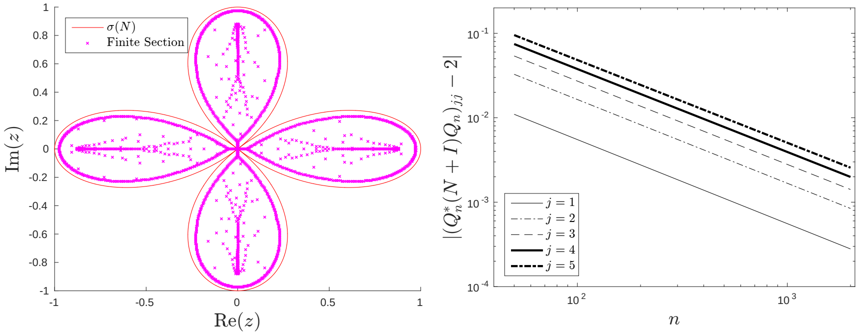


Figure: Left: Spectrum of the normal operator N and finite section approximates. Right: Convergence of first five diagonal entries of $Q_n^*(N+I)Q_n$ to 2. Convergence of the rate $\mathcal{O}(1/n)$ for this part of the essential spectrum as opposed to the linear convergence rate $\mathcal{O}(r^n)$ seen in the first example.

Example 4

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & a_{23} & a_{24} & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & a_{45} & a_{46} & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

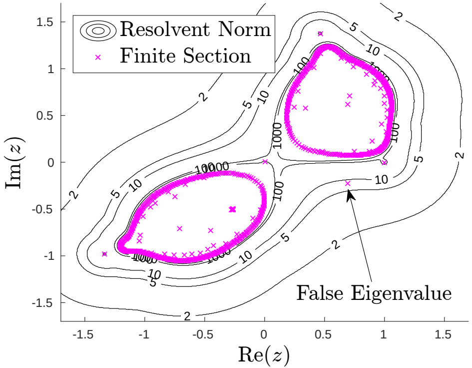
where $a_{2j,2j+1} = i$, and $a_{2j,2j+2} = -i$ if j is prime and $a_{2j,2j+2} = 0$ otherwise.

Example 4

$$T = \begin{pmatrix} 2.5+0.5i & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 1 & 3-0.5i & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 1.7 & 0.05 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0.05 & t_4 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & t_5 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & t_6 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 1 & t_7 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where $t_j = 1 + 0.5(\sin(j) + i\cos(j))$ for $j \geq 4$.

Example 4



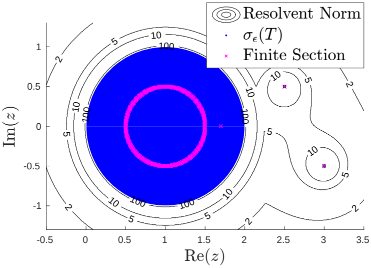


Figure: Left: $\sigma(P_n A|_{P_n\mathcal{H}})$ for $n = 1000$ with the false eigenvalue (recall that $\sigma(A) \subseteq \sigma_\epsilon(A)$). Right: $\sigma(P_n T|_{P_n\mathcal{H}})$ for $n = 500$ along with contours of the resolvent norm and $\sigma_\epsilon(T)$ for $\epsilon = 2\sqrt{\epsilon_{\mathrm{mach}}}$.

Example 4

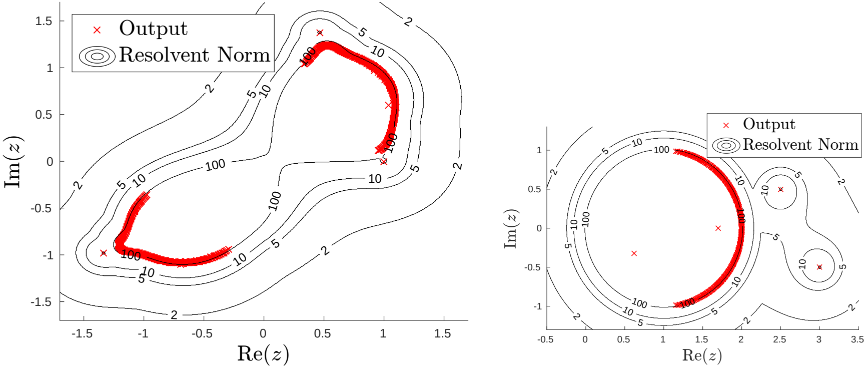


Figure: The figures show $\sigma(P_m Q_n^* A Q_n|_{P_m \mathcal{H}})$ (left), $\sigma(P_m Q_n^* T Q_n|_{P_m \mathcal{H}})$ (right) for $n = 1000, m = 1000$ and $n = 1500, m = 500$ respectively.

Example 4

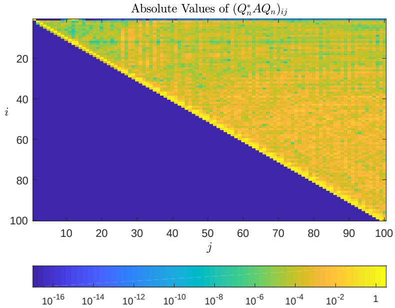


Figure: Left: Absolute values of entries of $Q_{1000}^* A Q_{1000}$. Right: Absolute values of entries of $Q_{1500}^* T Q_{1500}$. We appear to gain convergence to a diagonally dominated upper triangular matrix.

Conclusion

1. We can compute the IQR algorithm on a finite machine.

Conclusion

1. We can compute the IQR algorithm on a finite machine.
2. The IQR algorithm converges to the extremal parts of the spectrum (other convergence theorems not shown).

Conclusion

1. We can compute the IQR algorithm on a finite machine.
2. The IQR algorithm converges to the extremal parts of the spectrum (other convergence theorems not shown).
3. We can converge to dominant invariant subspaces.

Conclusion

1. We can compute the IQR algorithm on a finite machine.
2. The IQR algorithm converges to the extremal parts of the spectrum (other convergence theorems not shown).
3. We can converge to dominant invariant subspaces.
4. These results allow new classification results in SCI hierarchy (not shown).

Conclusion

1. We can compute the IQR algorithm on a finite machine.
2. The IQR algorithm converges to the extremal parts of the spectrum (other convergence theorems not shown).
3. We can converge to dominant invariant subspaces.
4. These results allow new classification results in SCI hierarchy (not shown).
5. Looks like it can be effective at avoiding spectral pollution and appears more robust than finite section.

Conclusion

1. We can compute the IQR algorithm on a finite machine.
2. The IQR algorithm converges to the extremal parts of the spectrum (other convergence theorems not shown).
3. We can converge to dominant invariant subspaces.
4. These results allow new classification results in SCI hierarchy (not shown).
5. Looks like it can be effective at avoiding spectral pollution and appears more robust than finite section.
6. But can't apply shifts.

Conclusion

1. We can compute the IQR algorithm on a finite machine.
2. The IQR algorithm converges to the extremal parts of the spectrum (other convergence theorems not shown).
3. We can converge to dominant invariant subspaces.
4. These results allow new classification results in SCI hierarchy (not shown).
5. Looks like it can be effective at avoiding spectral pollution and appears more robust than finite section.
6. But can't apply shifts.
7. Paper available soon!

Future Work

1. Link with IQL algorithm?

Future Work

1. Link with IQL algorithm?
2. Writing more efficient code.

Future Work

1. Link with IQL algorithm?
2. Writing more efficient code.
3. Many more open computational spectral problems - spectral measures, IQR for unbounded operators etc.

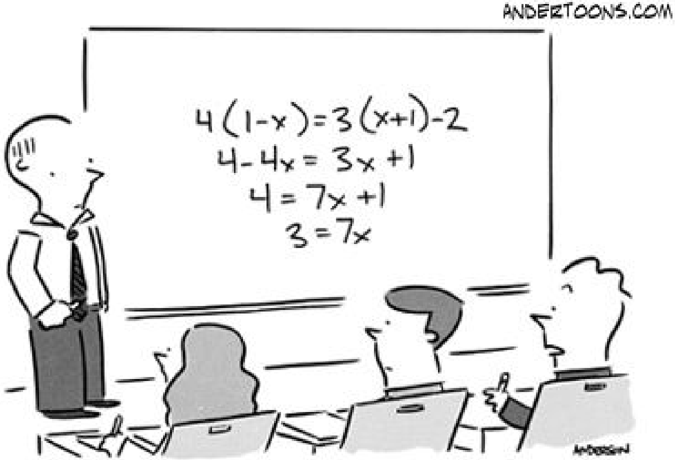
References

Only two main references:

1. First appearance of the algorithm for real symmetric infinite matrices: P Deift, LC Li, and C Tomei. Toda flows with infinitely many variables. Journal of functional analysis, 64(3):358402, 1985.
2. Existence of QR decomposition for non invertible operators and eigenvector convergence theorem for normal operators: AC Hansen. On the approximation of spectra of linear operators on Hilbert spaces. J. Funct. Anal., 254(8):20922126, 2008.

Other SCI results: In progress!

Thanks for listening!
Any questions?



"Wouldn't it be more efficient to just find who's complicating equations and ask them to stop?"