Taylor's Theorem for Matrix Functions and Pseudospectral Bounds on the Condition Number

Samuel Relton

samuel.relton@maths.man.ac.uk
@sdrelton
samrelton.com
blog.samrelton.com

Joint work with Edvin Deadman

edvin.deadman@nag.co.uk

University of Strathclyde
June 23rd, 2015

Manchester Numerical Analysis

Outline

- Taylor's Theorem for Scalar Functions
- Matrix Functions, their Derivatives, and the Condition Number
- Taylor's Theorem for Matrix Functions
- Pseudospectral Bounds on the Condition Number
- Numerical Experiments

Taylor's Theorem - 1

Theorem (Taylor's Theorem)

When $f : \mathbb{R} \to \mathbb{R}$ *is* k *times continuously differentiable at* $a \in \mathbb{R}$ *there exists* $R_k : \mathbb{R} \to \mathbb{R}$ *such that*

$$f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(a)}{j!}(x-a)^j + R_k(x).$$

Different expressions for the remainder term $R_k(x)$ *include*

- *the Lagrange form.*
- *the Cauchy form.*
- *the contour integral form.*

M\crNA
Manchester Numerical Analysis

Taylor's Theorem - 2

We can extend this to complex analytic functions.

If $f(z)$ is complex analytic in an open set $\mathcal{D} \subset \mathbb{C}$ then for any $a \in \mathcal{D}$

$$f(z) = \sum_{j=0}^{k} \frac{f^{(k)}(a)}{j!}(z-a)^j + R_k(z),$$

where

$$R_k(z) = \frac{(z-a)^{k+1}}{2\pi i} \int_{\Gamma} \frac{f(w)dw}{(w-a)^{k+1}(w-z)},$$

and Γ is a closed curve in $\mathcal{D}$ containing a.

Manchester Numerical Analysis

Matrix Functions

We are interested in extending this to matrix functions $f\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$.

For example:

- the matrix exponential

$$e^A = \sum_{j=0}^{\infty} \frac{A^j}{j!}.$$

- the matrix cosine

$$\cos(A) = \sum_{j=0}^{\infty} \frac{(-1)^j A^{2j}}{(2j)!}.$$

Applications include:

- Differential equations: $\frac{du}{dt} = Au(t), \quad u(t) = e^{tA}u(0).$
- Second order ODEs with sine and cosine.
- Ranking importance of nodes in a graph etc...

M\crNA
Manchester Numerical Analysis

Fréchet derivatives

Let $f : \mathbb{C}^{n\times n} \mapsto \mathbb{C}^{n\times n}$ be a matrix function.

Definition (Fréchet derivative)

The Fréchet derivative of f at A is the unique linear function $L_f(A, \cdot) : \mathbb{C}^{n\times n} \mapsto \mathbb{C}^{n\times n}$ such that for all E

$$f(A+E) - f(A) - L_f(A, E) = o(\|E\|).$$

- $L_f(A, E)$ is just a linear approximation to $f(A+E) - f(A)$.
- Higher order derivatives are defined recursively (Higham & R., 2014).

Manchester Numerical Analysis

Condition Numbers

A condition number describes the sensitivity of f at A to small perturbations arising from rounding error etc.

The absolute condition number is given by

$$\text{cond}_{\text{abs}}(f, A) := \lim_{\epsilon \to 0} \sup_{\|E\| \le \epsilon} \frac{\|f(A+E) - f(A)\|}{\epsilon} = \max_{\|E\|=1} \|L_f(A, E)\|,$$

whilst the relative condition number is

$$\text{cond}_{\text{rel}}(f, A) := \text{cond}_{\text{abs}}(f, A) \frac{\|A\|}{\|f(A)\|}.$$

M\crNA Manchester Numerical Analysis

Matrix Functions and Taylor's Theorem - 1

Previous results combining these two ideas include:

- an expansion around αI

$$f(A) = \sum_{j=0}^{\infty} \frac{f^{(j)}(\alpha)}{j!}(A - \alpha I)^j.$$

- an expansion in terms of derivatives

$$f(A+E) = \sum_{j=0}^{\infty} \frac{1}{j!} \left. \frac{d^j}{dt^j} \right|_{t=0} f(A+tE).$$

Note that:

- neither expansion has an explicit remainder term.
- $\left. \frac{d^j}{dt^j} \right|_{t=0} f(A+tE) = L_f(A, E)$ when $j = 1$.

Matrix Functions and Taylor's Theorem - 2

Let us take $D_f^{[j]}(A,E) := \left.\frac{d^j}{dt^j}\right|_{t=0} f(A+tE)$ then we have the following.

Theorem (Taylor's Theorem for Matrix Functions)

Let $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ we analytic in an open set $\mathcal{D} \subset \mathbb{C}$ with A, E satisfying $\Lambda(A), \Lambda(A+E) \subset \mathcal{D}$. Then

$$f(A+E) = T_k(A,E) + R_k(A,E),$$

where

$$T_k(A,E) = \sum_{j=0}^{k} \frac{1}{j!} D^{[j]}(A,E),$$

and

$$R_k(A,E) = \frac{1}{2\pi i} \int_\Gamma f(z)(zI - A - E)^{-1}[E(zI-A)^{-1}]^{k+1} dz,$$

where Γ is a closed contour enclosing $\Lambda(A)$ and $\Lambda(A+E)$.

Matrix Functions and Taylor's Theorem - 3

As an example take $f(z) = z^{-1}$.

$$D^{[1]}_{z^{-1}}(A, E) = -A^{-1}EA^{-1},$$
$$D^{[2]}_{z^{-1}}(A, E) = 2A^{-1}EA^{-1}EA^{-1}.$$

Therefore we have

$$(A+E)^{-1} = \frac{1}{0!}A^{-1} - \frac{1}{1!}A^{-1}EA^{-1} + \frac{2}{2!}A^{-1}EA^{-1}EA^{-1} + \int_{\Gamma} \frac{1}{z}(zI - A - E)^{-1}[E(zI - A)^{-1}]^3 dz.$$

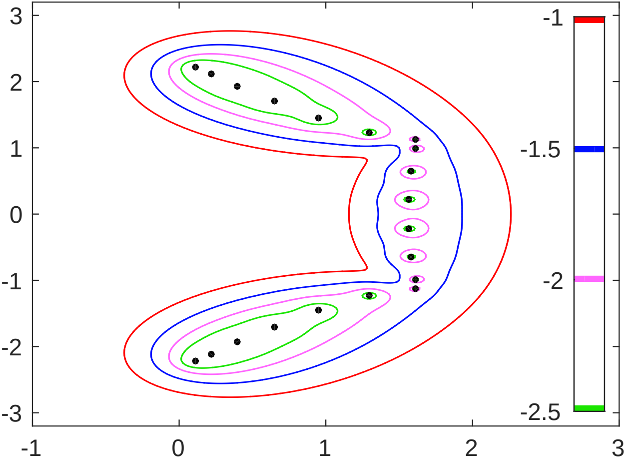
M\crNA
Manchester Numerical Analysis

Applying Pseudospectral Theory - 1

Recall that the ϵ-pseudospectrum of X is the set

$$\Lambda_\epsilon(X) = \{z \in \mathbb{C} : \|(zI - X)^{-1}\| \geq \epsilon^{-1}\}.$$

The ϵ-psuedospectral radius is $\rho_\epsilon = \max |z|$ for $z \in \Lambda_\epsilon(X)$.



Manchester Numerical Analysis

Applying Pseudospectral Theory - 1

Recall that the ϵ-pseudospectrum of X is the set

$$\Lambda_\epsilon(X) = \{z \in \mathbb{C} : \|(zI - X)^{-1}\| \geq \epsilon^{-1}\}.$$

The ϵ-psuedospectral radius is $\rho_\epsilon = \max |z|$ for $z \in \Lambda_\epsilon(X)$.

Using this we can bound the remainder term by

$$\|R_k(A, E)\| \leq \frac{\|E\|^{k+1} \tilde{L}_\epsilon}{2\pi\epsilon^{k+1}} \max_{z \in \tilde{\Gamma}_\epsilon} |f(z)|,$$

where

- $\tilde{\Gamma}_\epsilon$ is a contour enclosing $\Lambda_\epsilon(A)$ and $\Lambda_\epsilon(A + E)$.
- $\tilde{L}_\epsilon$ is the length of the contour $\tilde{\Gamma}_\epsilon$.
- ϵ is a parameter to be chosen.

Applying Pseudospectral Theory - 2

Applying this to $R_0(A, E)$ gives a bound on the condition number.

$$\mathrm{cond}_{\mathrm{abs}}(f, A) \leq \frac{L_\epsilon}{2\pi\epsilon^2} \max_{z \in \Gamma_\epsilon} |f(z)|,$$

where Γ_ϵ encloses $\Lambda_\epsilon(A)$ and has length L_ϵ.

Interesting because:

- Usually only lower bounds on condition number are known.
- Computing (or estimating) this efficiently could be of considerable interest in practice or for algorithm design.

M\crNA
Manchester Numerical Analysis

The Condition Number of A^t - 1

This upper bound is extremely efficient to compute for the matrix function given by $f(x) = x^t$ for $t \in (0, 1)$.

Our experiments will

- determine how tight the upper bound is as ϵ changes.
- see how fast evaluating the upper bound is in comparison to computing it exactly.

Other methods for this problem are:

- "CN Exact" – computes condition number exactly.
- "CN Normest" – lower bound using norm estimator.

M\crNA
Manchester Numerical Analysis

The Condition Number of A^t - 2

This function has a branch cut along the negative real line, meaning we need to choose a keyhole contour. Overall:

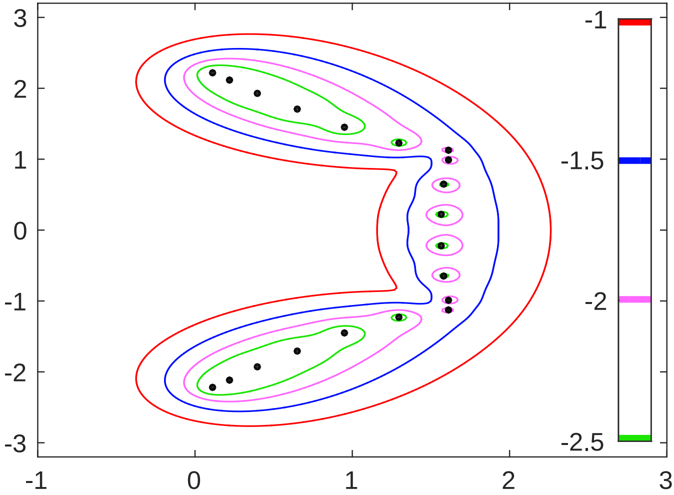
$$\text{cond}_{\text{abs}}(x^t, A) \leq \frac{2(\pi + 1)\rho_{\epsilon\sqrt{n}}^{1+t}}{2\pi\epsilon^2},$$

where ρ_ϵ is the ϵ-pseudospectral radius, computed using code by Gugliemi and Overton.

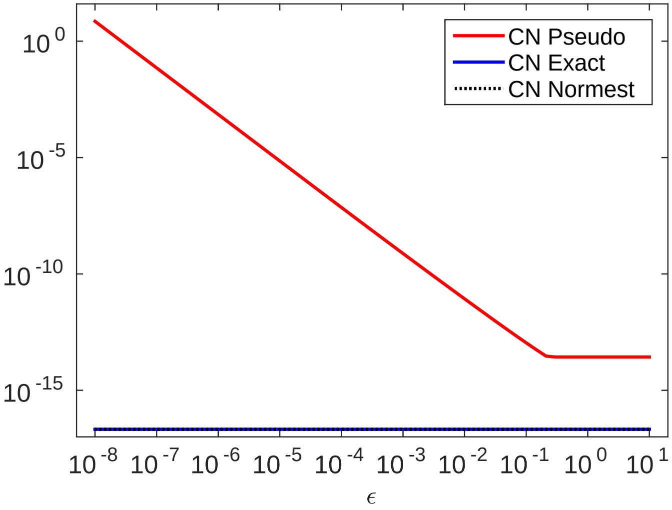
Note: There is an upper limit for ϵ where the pseudospectrum intersects the branch cut. We need to take ϵ smaller than this value.

M\cr NA
Manchester Numerical Analysis

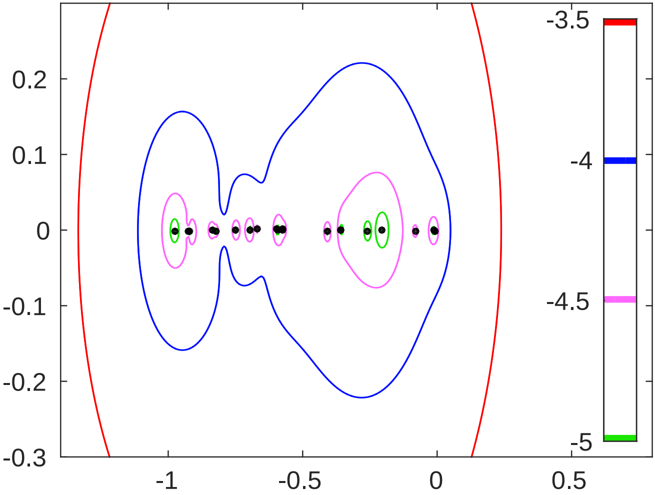
Test matrix - Grcar matrix



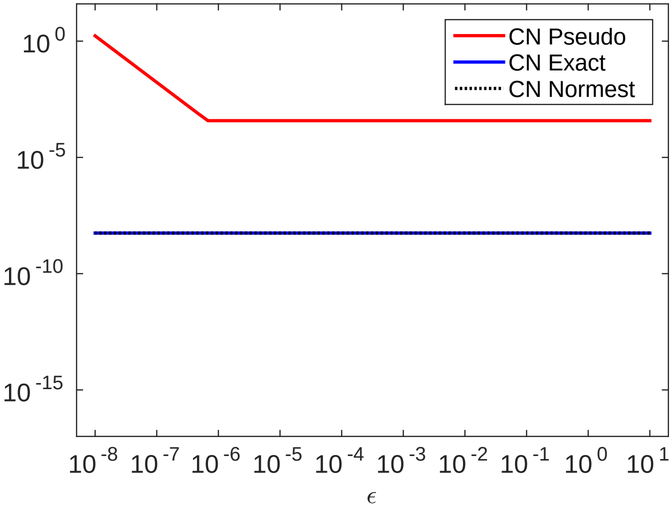
CN Bound as ϵ varies



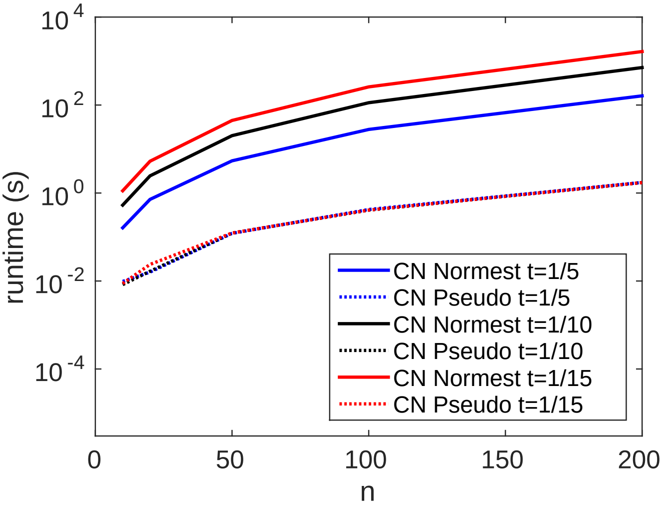
Test matrix - Almost neg. eigenvalues



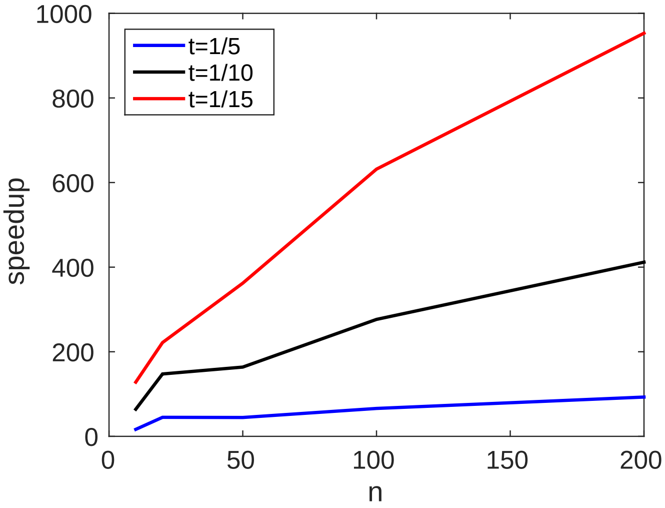
CN Bound ϵ varies



Runtime Comparison - Timings



Runtime Comparison - Speedup



Manchester Numerical Analysis

Conclusions

- Extended Taylor's theorem to matrix functions.
- Applied pseudospectral theory to bound remainder term.
- Bounds are very efficient to compute for A^t.
- If bound is unsatisfactorily large can revert to a more precise method.

Future work:

- Apply to algorithm design.
- Find other classes of functions for which this is efficient.