The infinity norm bounds and characteristic polynomial for high order RK matrices

Gayatri Čaklović

January 2022

Abstract

This paper shows that $t_m \leq \|A\|_\infty \leq \sqrt{t_m}$ holds, when $A \in \mathbb{R}^{m \times m}$ is a Runge-Kutta matrix which nodes originating from the Gaussian quadrature that integrates polynomials of degree $2m - 2$ exactly. It can be shown that this is also true for the Gauss-Lobatto quadrature. Additionally, the characteristic polynomial of $A$, when the matrix is nonsingular, is $p_A(\lambda) = m!t_m^m + (m - 1)!a_{m - 1}t_m^{m - 1} + \cdots + a_0$, where the coefficients $a_i$ are the coefficients of the polynomial of nodes $\omega(t) = (t - t_1) \cdots (t - t_m) = t_m^m + a_{m - 1}t_m^{m - 1} + \cdots + a_0$.

1 The Runge-Kutta method and Gaussian quadrature nodes

The Runge-Kutta methods are a family of iterative methods for approximating the solutions of ordinary differential equations. A method can be represented in a form of a famous Butcher tableau:

$$
\begin{array}{c|c}
  c & A \\
  \hline \\
  b^T \\
\end{array}
$$

where $c, b \in \mathbb{R}^m$ represent the nodes and the weights and the matrix $A \in \mathbb{R}^{m \times m}$ is called the Runge-Kutta matrix of order $m$. More about the origin of the methods can be found in [1].

1.1 Gaussian quadrature nodes

On function spaces, there is a natural scalar product. For integrable functions $f, g : \mathbb{R} \to \mathbb{R}$ a scalar product can be defined as

$$
(f, g) = \int_0^1 f \, g. \quad (1)
$$
This allows us to define orthogonality on function spaces. The family of orthogonal polynomials can be used to construct a high order Gaussian-quadrature rules. Let
\[ \int_0^1 p(s)ds = \sum_{i=1}^m w_i p(t_i) \] (2)
define a quadrature rule, where some points \( t_i \) may be fixed and while other are roots of an orthogonal polynomial. Then, the right sum computes the integral correctly for polynomials of degree \( 2m - 2 \) or less. Depending on the choice of these points, we get different quadrature rules such as Legendre, Radau, Lobatto, etc. The integration weights \( w_i \) are positive and the integration points \( t_i \) are inside the integration interval \([2]\). In our case, without loss of generality, we will assume that \( t_i \in [0, 1] \).

Let \( l_j \) be a Lagrange interpolation polynomial defined in points \( t_i \):
\[ l_j(t) = \prod_{i \neq j} \frac{t - t_i}{t_j - t_i}, \quad j = 1, \ldots, m. \]

If we define the Runge-Kutta matrix as
\[ A_{ij} = \int_0^{t_i} l_j(s)ds, \] (3)
we end up with a high order Runge-Kutta method where the vector \( c = (t_1, \ldots, t_m) \) and \( b = (w_1, \ldots, w_m) \).

### 2 Bounds for \( \|A\|_\infty \)

Let the Runge-Kutta matrix be defined as in (3). In order to compute lower and upper bounds for \( \|A\|_\infty \), we firstly have to understand what does a matrix-vector product \( Ax \) mean. The product behaves as
\[ Ax = \begin{bmatrix} \sum_{j=1}^m x_j \int_{t_0}^{t_i} l_j(s)ds \\ \sum_{j=1}^m x_j \int_{t_0}^{t_2} l_j(s)ds \\ \vdots \\ \sum_{j=1}^m x_j \int_{t_0}^{t_m} l_j(s)ds \end{bmatrix} = \begin{bmatrix} \int_{t_0}^{t_1} \sum_{j=1}^m x_j l_j(s)ds \\ \int_{t_0}^{t_2} \sum_{j=1}^m x_j l_j(s)ds \\ \vdots \\ \int_{t_0}^{t_m} \sum_{j=1}^m x_j l_j(s)ds \end{bmatrix}. \]

This motivates us to construct a one-to-one map from a vector \( x \in \mathbb{R}^m \) to a polynomial \( p \in P_{m-1} \), where \( p := \sum_{j=1}^m x_j l_j \). This is uniquely defined because the polynomial \( p \) passes through \( m \) points \( (t_i, x_i) \) and is of degree \( m - 1 \). Then, the sum \( \int_{t_0}^{t_i} \sum_{j=1}^m x_j l_j(s)ds \) is nothing more than \( \int_{t_0}^{t_i} p(s)ds \) and we can conclude that multiplying a vector with a matrix \( A \) is as computing integrals of a corresponding polynomial.

Now, we are interested in computing the norm \( \|A\|_\infty \). By definition, the norm is
\[ \|A\|_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty. \] (4)
but the infinity norm of a matrix can also be computed as
\[
\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{m} |A_{ij}| = \max_{1 \leq i \leq m} \sum_{j=1}^{m} \left| \int_0^{t_i} l_j(s) ds \right|,
\]
which is the maximum row-sum of absolute values of its entries. From this, we can see that
\[
t_i = \left| \int_0^{t_i} 1 ds \right| = \left| \int_0^{t_i} \sum_{j=1}^{m} l_j(s) ds \right| \leq \sum_{j=1}^{m} \left| \int_0^{t_i} l_j(s) ds \right| \leq \sum_{j=1}^{m} |A_{ij}| \leq \|A\|_\infty.
\]
This yields a lower bound: \(t_m \leq \|A\|_\infty\).

Now we will redefine the definition of the matrix norm (4) in a way that we do not compute the maximum over all vectors \(x\) for which \(\|x\|_\infty = 1\), but over uniquely defined polynomials \(p\) since \(\mathbb{R}^m \ni x \mapsto p \in P_{m-1}\) is a one-to-one map, as discussed above. Proving \(\|A\|_\infty \leq 1\) is then equivalent to proving
\[
\max_{1 \leq i \leq M} \left| \int_0^{t_i} p(s) ds \right| \leq 1
\]
holds, for \(p \in P_{m-1}\) and \(\max_{1 \leq i \leq M} |p(t_i)| = 1\).

2.1 Exact quadrature for degree \(2m - 2\) or higher

First, using the integral Cauchy–Bunyakovsky-Schwarz inequality on the scalar product defined as in (4), we can get an upper bound on \(\int_0^{t_i} p(s) ds\) as
\[
\left| \int_0^{t_i} p(s) ds \right|^2 = \left| \int_0^{t_i} p(s) \cdot 1 ds \right|^2 \leq \int_0^{t_i} p^2(s) ds \left| \int_0^{t_i} 1^2 ds \right| \leq t_i \int_0^{t_i} p^2(s) ds \leq t_i \int_0^{t_i} p^2(s) ds \leq t_i \int_0^{t_i} p^2(s) ds.
\]
Since \( p^2 \in P_{2m-2} \), using the quadrature rule (2) which correctly computes integrals of polynomials of degree \( 2m - 2 \), we get

\[
\int_0^1 p^2(s)ds = \sum_{i=1}^M w_i p^2(t_i) \leq \left( \sum_{i=1}^M w_i \right) \max_{1 \leq i \leq M} p^2(t_i) = \max_{1 \leq i \leq M} p^2(t_i).
\]

Because we are computing (5) over polynomials which satisfy \( \max_{1 \leq i \leq M} |p(t_i)| = 1 \), we know that \( \max_{1 \leq i \leq M} p^2(t_i) = 1 \), thus

\[
\int_0^1 p^2(s)ds \leq \max_{1 \leq i \leq M} p^2(t_i) = 1.
\] (9)

Now (8) + (9) yields

\[
\left| \int_0^{t_i} p(s)ds \right|^2 \leq t_i.
\]

Thus,

\[
t_i \leq \left| \int_0^{t_i} p(s)ds \right| \leq \sqrt{t_i}.
\]

This proves that \( t_m \leq \|A\|_{\infty} \leq \sqrt{t_m} \leq 1 \) holds for nodes originating from the Gaussian quadrature that integrates polynomials of degree \( 2m - 2 \) exactly.

### 2.2 Gauss-Lobatto nodes

The Gauss-Lobatto quadrature integrates polynomials of accuracy \( 2m - 3 \), therefore the same argument why is \( \|A\|_{\infty} \leq 1 \) does not hold. However, we will show that this is still true for \( m \geq 2 \). The Gauss-Lobatto quadrature on \([0, 1]�\) is defined as

\[
\int_0^1 p(s)ds = \sum_{i=1}^M w_i p(t_i) + R_m,
\] (10)

where \( R_m = -cp(2m-2)(\xi), c \geq 0, \xi \in [0, 1]�\) for details. If we manage to bind \( \int_0^1 p^2(s)ds \leq 1 \), then using the same train of thought and arguments as in (8), we are done.

Let \( p^2(s) = as^{2m-2} + \ldots \), where \( a > 0 \). We can write

\[
\int_0^1 p^2(s)ds = \int_0^1 (p^2(s) - as^{2m-2})ds + \int_0^1 as^{2m-2}ds.
\]

The polynomial \( p^2(s) - as^{2m-2} \) is of degree \( 2m - 3 \) and can be computed with quadrature (10) as

\[
\int_0^1 (p^2(s) - as^{2m-2})ds = \sum_{i=1}^m w_i (p^2(t_i) - a t_i^{2m-2}).
\]
This yields
\[ \int_0^1 p^2(s)ds = \sum_{i=1}^m w_i \left( p^2(t_i) - at_i^{2m-2} \right) + \int_0^1 as^{2m-2}ds \]
\[ \leq 1 - a \sum_{i=1}^m w_i t_i^{2m-2} + \int_0^1 as^{2m-2}ds, \]  
where the inequality holds because \( \sum_{i=1}^m w_i p^2(t_i) \leq 1 \) for \( \max_{1 \leq i \leq M} |p(t_i)| \leq 1 \). Now using the quadrature (10) on a function \( t^{2m-2} \) yields
\[ \int_0^1 s^{2m-2} = \sum_{i=1}^m w_i t_i^{2m-2} - c(2m-2)! \]
and combining this with (12) gives
\[ \int_0^1 p^2(s)ds \leq 1 - ac(2m-2)! \leq 1, \]
because \( ac \geq 0 \). This proves that \( t_m \leq \|A\|_\infty \leq \sqrt{t_m} \leq 1 \) is also true for the Gauss-Lobatto quadrature.

3 The characteristic polynomial of A

Let \( \omega(t) = (t-t_1)\ldots(t-t_m) = t^m + a_{m-1}t^{m-1} + \cdots + a_0 \)
be the polynomial of nodes. If \( t_i \) are the roots of an orthogonal polynomial
defining the Gaussian quadrature, then \( \omega \) is exactly that scaled polynomial. We
will show that the characteristic polynomial of \( A \) is
\[ p_A(\lambda) = mt^m + (m-1)!a_{m-1}t^{m-1} + \cdots + a_0. \]
(13)
The eigenvalue of matrix \( A \) is a scalar \( \lambda \in \mathbb{C} \) that satisfies
\[ Ax = \lambda x, \]
for some \( x \neq 0 \). In other words, using a one-on-one map to the polynomial
space \( P_{m-1} \), we can reformulate the eigenproblem as finding a polynomial \( p \neq 0 \)
such that
\[ \int_0^{t_i} p(s)ds = \lambda p(t_i), \quad i = 1, \ldots, m. \]
Substituting \( p \) with \( g' \), where \( g \in P_m \) yields
\[ g(t_i) - g(0) = \lambda g'(t_i), \quad i = 1, \ldots, m. \]
(14)
From here we see that the eigenvector represented as a polynomial \( g \) is not
uniquely defined, because if \( g \) is a solution of (14), then \( d_m g(t) + d_0 \) is also a
solution. Therefore, without loss of generality, we can assume that $g$ has the form

$$g(t) = t^m + d_{m-1}t^{m-1} + \cdots + d_1 t.$$  

Setting $d_m = 1$ is equivalent to imposing that the eigenvector is normalized which yields a uniqueness of the eigenvector. However, here we impose $x$ has a fixed infinity norm, one that may differ from 1. The part where we demand that $d_0 = 0$ comes from substituting $p$ with $g'$. As a consequence we have $g(0) = 0$. With these assumptions, $g$ is uniquely defined since we have $m$ nonlinear equations and $m$ unknowns: $(d_{m-1},\ldots,d_1,\lambda)$ that are described with

$$g(t_i) = \lambda g'(t_i), \quad i = 1,\ldots,m. \quad (15)$$

Let $G(t) := g(t) - \lambda g'(t)$, where $g$ is the solution of (15). The difference $G - \omega$ is a polynomial of degree $m - 1$ and both $\omega$ and $G$ are monic polynomials of degree $m$. Because $G - \omega$ is zero in $m$ different points, we conclude that $G = \omega$. Equations (15) can now be rewritten in a continuous form as

$$g(t) - \lambda g'(t) = \omega(t).$$

Since the polynomials coefficients on the left side should be equal to the coefficients on the right side of the upper equation, we get

$$-\lambda d_1 = a_0$$
$$d_1 - 2\lambda d_2 = a_1$$
$$d_2 - 3\lambda d_3 = a_2$$
$$\vdots$$
$$d_{m-2} - (m-1)\lambda d_{m-1} = a_{m-2}$$
$$d_{m-1} - m\lambda = a_{m-1}.$$  

Substituting the equations from the bottom to the second from the top, we get

$$d_1 = m!\lambda^{m-1} + (m-1)!a_{m-1}\lambda^{m-2} + \cdots + 2\lambda a_2,$$  

whereas from the first one we get $d_1 = -a_0/\lambda$. Substituting this in (16) we end up with a polynomial $p_A$ as in (13) which zeros are the eigenvalues of our matrix $A$. In other words, $p_A$ is exactly the characteristic polynomial of $A$.

It remains to comment why $\lambda = 0$ is not a viable option for an eigenvalue. Matrix $A$ is a mapping in a fashion

$$(t_1^k,\ldots,t_m^k) \rightarrow \frac{1}{k+1}(t_1^{k+1},\ldots,t_m^{k+1}), \quad k = 0,\ldots,m-1,$$

if the quadrature rule defining $A$ integrates polynomials of degree $m - 1$ correctly. The vectors $(t_1^k,\ldots,t_m^k)$ form columns of the Vandermonde matrix $V$ which is known to be nonsingular if the points $t_1,\ldots,t_m$ are distinct. Because of this, we have $AV = D_mV$, where $D_m = \text{diag}(1,1/2,\ldots,1/m)$. From here we see that $A$ is nonsingular as a product of nonsingular matrices.
4 Conclusion

We proved that $t_m \leq \|A\|_\infty \leq \sqrt{t_m} \leq 1$ holds for nodes originating from the Gaussian quadrature that integrates polynomials of degree $2m - 2$ exactly. It is also true for the Gauss-Lobatto quadrature. Additionally, the characteristic polynomial of $A$ can be found, when the matrix is nonsingular. The matrix is nonsingular if it originates from a quadrature that integrates polynomials in an exact way up to degree $m - 1$.

References

[1] John C. Butcher. *Runge–Kutta Methods*, chapter 3, pages 143–331. John Wiley & Sons, Ltd, 2016.

[2] Alfio Quarteroni, Riccardo Sacco, and Fausto Saleri. *Numerical Integration*, pages 379–422. Springer, Berlin, Heidelberg, 2006.