Butcher algebra of the matrix

Sergey Khashin
http://math.ivanovo.ac.ru/dalgebra/Khashin/index.html

Ivanovo State University, Russia

.

New York Group Theory Seminar

Plan

What is Runge-Kutta methods

Let an initial value problem be specified as follows.

$$y' = f(t, y), \ t \in \mathbb{R}, \ y \in \mathbb{R}^n, \ y(t_0) = y_0$$

Now pick a step-size $h > 0$ and define

$$y_1 = y_0 + (k_1 + 2k_2 + 2k_3 + k_4)/6$$

where

$$\begin{aligned}
k_1 &= hf(t_0 &&, y_0), \\
k_2 &= hf(t_0 + \tfrac{h}{2} &&, y_0 + \tfrac{1}{2}k_1), \\
k_3 &= hf(t_0 + \tfrac{h}{2} &&, y_0 + \tfrac{1}{2}k_2), \\
k_4 &= hf(t_0 + h &&, y_0 + k_3),
\end{aligned}$$

Classical Runge-Kutta method is a fourth-order methods with four stages, $RK(4,4)$.

Butcher tableau

All coefficients can be combined into one table(Butcher tableau):

$$
\begin{array}{c|cccc}
c_2 & a_{21} & & & \\
c_3 & a_{31} & a_{32} & & \\
c_4 & a_{41} & a_{42} & a_{43} & \\
\hline
 & b_1 & b_2 & b_3 & b_4
\end{array}
$$

where

$$
\begin{aligned}
c_2 &= a_{21}\,, \\
c_3 &= a_{31} + a_{32}\,, \\
c_4 &= a_{41} + a_{42} + a_{43}\,, \\
1 &= b_1 + b_2 + b_3 + b_4\,.
\end{aligned}
$$

$RK(4,4)$ equations

Coefficients (a_{ij}, b_j) must satisfy the equations (*order conditions* or *Butcher equations*):

$$
\begin{array}{lll}
0) & b_1 + b_2 + b_3 + b_4 & = 1, \\
1) & b_2 c_2 + b_3 c_3 + b_4 c_4 & = 1/2, \\
2) & b_3 a_{32} c_2 + b_4 (a_{42} c_2 + a_{43} c_3) & = 1/6, \\
3) & b_2 {c_2}^2 + b_3 {c_3}^2 + b_4 {c_4}^2 & = 1/3, \\
4) & b_4 a_{43} a_{32} c_2 & = 1/24, \\
5) & b_3 c_3 a_{32} c_2 + b_4 c_4 (a_{42} c_2 + a_{43} c_3) & = 1/8, \\
6) & b_2 c_2^3 + b_3 c_3^3 + b_4 c_4^3 & = 1/4, \\
7) & b_3 a_{32} c_2^2 + b_4 (a_{42} c_2^2 + a_{43} c_3^2) & = 1/12,
\end{array}
$$

Extended matrix

For my purposes it is convenient to use an *extended* $(s+1)\times(s+1)$-matrix A of the $RK(p,s)$-method that is defined as follows.

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ a_{21} & 0 & 0 & 0 & \dots & 0 \\ a_{31} & a_{32} & 0 & 0 & \dots & 0 \\ & \dots & & & & \\ a_{s1} & a_{s2} & \dots & a_{s,s-1} & 0 & 0 \\ \color{red}{b_1} & \color{red}{b_2} & \dots & \color{red}{b_{s-1}} & \color{red}{b_s} & \color{red}{0} \end{pmatrix}$$

where as usual the first column can be expressed in terms of the others:

$$a_{k1} = c_k - a_{k2} - \cdots - a_{k,k-1} \quad \forall k = 2 \dots s \ .$$

John Butcher-1963

For a long time, not only the solution, but also derivation of the *order conditions* in the general case was a big problem. New approaches have been gradually accumulated and the breakthrough came in two articles:

J.C. Butcher. Coefficients for the study of runge-kutta integration processes. *J. Austral. Math. Soc.*, 3:185–201, 1963.

J.C. Butcher. On Runge-Kutta processes of high order*J. Austral. Math. Soc.*, 4:179–194, 1964.

They described the *order conditions* in general: one equation for each rooted tree.

Methods of order 5

In 1964 J.Bucher find the 5-dimensional family of 6-stage methods of order 5.

J. C. Butcher, *On Runge-Kutta processes of high order*, J. Austral. Math. Soc. **4** (1964), 179–194.

In 1969 Cassity showed that the Butcher family is only a subvariety of larger, 6-dimensional family.

C. R. Cassity, *The complete solution of the fifth order Runge-Kutta equations*, SIAM J. Numer. Anal. **6** (1969), 432–436.

What do we mean by "found"? This means that given a certain algorithm, by which bound variables are expressed in terms of free ones.

Methods of order 6

J. Butcher (1966) found the 4-dimensional family of 7-stages methods of order 6.
I found (numerically) a large number of individual methods of type $RK(6,7)$ and define local dimension of the solution variety in these points. It turned out that many of the methods are found not to contain in Butcher family.
Some analytic formulas (Maple-functions) was fond by D.Verner and me:
http://math.ivanovo.ac.ru/dalgebra/Khashin/rk/sh_rk.html

Methods of order ≥ 7

J. Butcher found the 2-dimensional family of 9-stages methods of order 7.

In the works of Curtis, Verner, Cooper, and some other authors some methods of orders 7, 8 and even 10 were found.

J.H. Verner. *Refuge for Runge-Kutta Pairs*,
http://people.math.sfu.ca/~jverner/

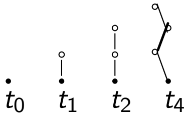
P. Stone. *Peter Stone's Maple Worksheets*.
http://www.peterstone.name/Maplepgs

Sharp P.W., Verner J.H., *Generation of high-order interpolants for explicit Runge-Kutta pairs*, TOMS, 24, 1, 13-29. 1998.

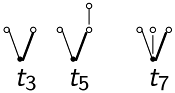
Trees

Following standard Butcher's approach, we use trees. We recall operations from graph theory.

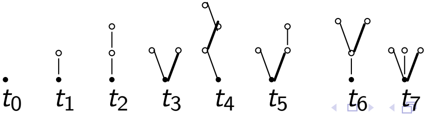
Here t_0 is a tree with only one vertex,
$t_1 = \alpha t_0$ – adding a vertex and an edge to the root,
$t_2 = \alpha^2 t_0$,
$t_4 = \alpha(t_2) = \alpha^3(t_0)$.



Multiplication of trees:
$t_3 = t_1 \cdot t_1$,
$t_5 = t_1 \cdot t_2$,
$t_7 = t_1 \cdot t_1 \cdot t_1$.



So we have the following 8 trees of weight ≤ 3.



Trees semigroup

Definition

We denote the set of all non-isomorphic rooted trees as $\mathcal{T}$.

Theorem

Every tree $t \in \mathcal{T}$ can be obtained from t_0 by combination of operations α and multiplication of trees.

So, $\mathcal{T}$ is a free semigroup, generated by all “one-leg“ trees.

Function $\delta(t)$

Definition

Let $t \in \mathcal{T}$. Then $\delta(t)$ is the product of all orders $(w(t_v) + 1)$, where v denotes a vertex of t and v is not the root:

$$\delta(t) = \prod_{v \neq \text{root}} (w(t_v) + 1) \ .$$

where *weight* $w(t)$ is a number of edges in the tree.

Theorem

The following properties hold:

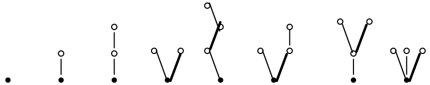
1. $\delta(t_0) = 1$,
2. $\delta(t_1 \cdot t_2) = \delta(t_1)\delta(t_2)$ *for any* $t_1, t_2 \in \mathcal{T}$,
3. $\delta(\alpha t) = \delta(t)(w(t) + 1)$ *for any* $t \in \mathcal{T}$.

Let $e = (1, \dots, 1)^T \in \mathbb{R}^n$ and "$*$" – coordinate-wise multiplication in $\mathbb{R}^n$.

For a given $n \times n$-matrix A we define $\Phi_A : \mathcal{T} \to \mathbb{R}^n$:

$$
\begin{aligned}
\Phi(t_0) &= e, & \delta(t_0) &= 1,\\
\Phi(t_1) &= Ae, & \delta(t_1) &= 1,\\
\Phi(t_2) &= A^2 e, & \delta(t_2) &= 2,\\
\Phi(t_3) &= Ae * Ae, & \delta(t_3) &= 1,\\
\Phi(t_4) &= A^3 e, & \delta(t_4) &= 6,\\
\Phi(t_5) &= Ae * A^2 e, & \delta(t_5) &= 2,\\
\Phi(t_6) &= A(Ae * Ae), & \delta(t_6) &= 2,\\
\Phi(t_7) &= Ae * Ae * Ae, & \delta(t_7) &= 1,
\end{aligned}
$$

where $e = (1, \dots, 1)^t$ and "$*$" – coordinate-wise multiplication in $\mathbb{R}^{s+1}$.



t_0 t_1 t_2 t_3 t_4 t_5 t_6 t_7

Butcher equations

Theorem

Matrix A is a matrix of RK method of order p, if for each rooted tree t of weight $\leq p$ the last coordinate of vector $\Phi_t(A)$ equals $1/\delta(t)$.

$$(\Phi_t(A), e') = 1/\delta(t), \text{ where } e' = (0, \ldots, 0, 1).$$

We will consider this equations only of “one-leg“ trees. It is a very large polynomial systems:

order	1	2	3	4	5	6	7	8	9	10
number of eqs	1	2	4	8	17	37	85	200	486	1205
min. number of stages :				4	6	7	9	11	13	≤ 17

Subspaces L_k and M_k

Consider subspaces generated by $\Phi_t(A)$ with trees of weight k:

$$L_k = < \Phi_t(A) \mid w(t) = k > \subset \mathbb{R}^{s+1} .$$

For example,

$$\begin{aligned} L_0 &= < e > , \\ L_1 &= < Ae > , \\ L_2 &= < A^2e,\ Ae * Ae > , \\ L_3 &= < A^3e,\ A(Ae * Ae),\ A^2e * Ae,\ Ae * Ae * Ae > , \end{aligned}$$

Consider a filtration in $\mathbb{R}^{s+1}$: chain of subspaces $0 \subset M_0 \subset M_1 \subset M_2 \ldots$:

$$\begin{aligned} M_0 &= L_0 , \\ M_1 &= L_0 + L_1 , \\ M_2 &= L_0 + L_1 + L_2 , \\ M_3 &= L_0 + L_1 + L_2 + L_3 , \end{aligned}$$

Theorem This filtration compartible with the multiplication, that is

$$M_i * M_j \subset M_{i+j}, \quad A(M_i) \subset M_{i+1} .$$

Butcher algebra of the matrix

Khashin S.I. 2009:
Let A be an $n \times n$ lower triangular matrix with zero diagonal. Consider subspaces $L_k =< \Phi_t(A) >$ of $\mathbb{R}^n$ where t is a tree of weight k and filtration of the space $\mathbb{R}^n$ for every given matrix A: $M_k = \sum_{i=0}^{k} L_i$.

Definition

We say that the adjoint algebra corresponding to this filtration,

$$B(A) = \bigoplus_{k=0}^{n} B_k(A) = \bigoplus_{k=0}^{n} M_k / M_{k-1}$$

is an *upper Butcher algebra* of matrix A.

Simplifying assumptions

Traditionally, specialists in RK methods, solve their equations by introducing *simplifying assumptions*. These are properties of low order RK methods, which are stated to be true for higher order RK methods, and which lead to a much simpler systems.

It was an art to find good simplifying assumptions.

In my algebraic approach they are natural consequences of dimension restrictions.

Simplifying assumptions via subspaces

Thus,

1. $M_{p-1} = \mathbb{R}^{s+1}$ is the same as $C(2)$;
2. $M_{p-2} = \mathbb{R}^{s+1}$ is the same as $D(1)$;
3. $M_{p-3} = \mathbb{R}^{s+1}$???? (shall we name it $E(0)$???)

Theoretically, we can find further simplifying assumptions as $M_{p-4} = \mathbb{R}^{s+1}, \dots$. However, it turns out that they are not true for many interesting methods.

That is why we suggest further modification of our idea.

Subspaces L'_k

Thus, we change our construction a little (our new subspaces are denoted by primes).

Definition. For an arbitrary tree t, define the vector

$$\Phi'_t(A) = \delta(t)\Phi_t(A) - \underbrace{Ae * \cdots * Ae}_{d} ,$$

where $d = w(t)$ is the weight of the tree, and $\delta(t)$ is some modification of the standard $\gamma(t)$.

Note that the order conditions imply that the last coordinate of this vector is zero for $d < p$.

Definition. For a given matrix A consider subspaces L'_k, $k = 0, 1, \ldots$ generated by vectors $\Phi'_t(A)$ for all trees t of weight k.

$$L'_0 = L'_1 = 0 ,$$

$$L'_2 =< 2A^2e - Ae{*}Ae > ,$$

$$L'_3 =< 6A^3e - Ae{*}Ae{*}Ae,\ 3A(Ae{*}Ae) - Ae{*}Ae{*}Ae,$$

$$2A^2e{*}Ae - Ae{*}Ae{*}Ae >$$

Subspaces M'_k

For given matrix A consider the filtration $0 \subset M'_2 \subset M'_3 \dots$:

$$\begin{aligned} M'_0 &= 0 \,, \\ M'_1 &= 0 \,, \\ M'_2 &= L'_2 \,, \quad (\dim M'_2 = 1) \\ M'_3 &= L'_2 + L'_3 \,, \\ M'_4 &= L'_2 + L'_3 + L'_4 \,, \\ \dots \end{aligned}$$

This filtration is proofed to be compartible with the multiplication, that is

$$M'_i * M'_j \subset M'_{i+j}, \quad A(M'_i) \subset M'_{i+1} \,.$$

Lower Butcher algebra of the matrix

Let A be an $n \times n$ lower triangular matrix with zero diagonal. Consider subspaces $L'_k = < \Phi'_t(A) >$ of $\mathbb{R}^n$ where t is a tree of weight k and filtration of the space $\mathbb{R}^n$ for every given matrix A: $M_k = \sum_{i=0}^{k} L_i$.

Definition

We say that the adjoint algebra corresponding to this filtration,

$$B'(A) = \bigoplus_{k=0}^{n} B'_k(A) = \bigoplus_{k=0}^{n} M'_k / M'_{k-1}$$

is an *lower Butcher algebra* of matrix A.

New simplifying assumptions

We calculate the dimensions of the introduced subspaces $B'_k = M'_k / M'_{k-1}$ for all known RK-methods:

Method, k:	0	1	2	3	4	5	6	7	8
RK(p=3,s= 3) :	0	0	1	1	–	–	–	–	–
RK(p=4,s= 4) :	0	0	1	1	1	–	–	–	–
RK(p=5,s= 6) :	0	0	1	2	1	1	–	–	–
RK(p=6,s= 7) :	0	0	1	1	2	1	1	–	–
RK(p=7,s= 9) :	0	0	1	1	2	2	1	1	–
RK(p=8,s=11) :	0	0	1	1	2	2	2	1	1

Note that the sum of the elements in each row is $s-1$.
We suggest the next new simplifying assumption: $\dim B'_3 = 1$. We see from the table that RK($p=5, s=6$) will not satisfy this condition. However, for all known higher order RK methods it holds.

Vectors w_k

Now more detailed computations.

Definition

For $k \geq 2$ denote by w_k vector

$$w_k = kA(\underbrace{Ae * \cdots * Ae}_{k-1}) - \underbrace{Ae * \cdots * Ae}_{k} \in L'_k.$$

That is

$$\begin{aligned}
w_2 &= 2A^2e - Ae * Ae, \\
w_3 &= 3A(Ae * Ae) - Ae * Ae * Ae, \\
w_4 &= 4A(Ae * Ae * Ae) - Ae * Ae * Ae * Ae, \\
&\dots,
\end{aligned}$$

This vectors w_k allow us to define L'_k recursively (we shall omit the details here, and show only the consequences).

Simplifying assumptions of level $3, 4$

We propose to call

1. $C(2)$ level 1 simplification;
2. $D(1)$ level 2 simplification.

Simplifying assumptions of level 3: $\dim B_3' = 1$, that is $\dim M_3' = 2$.
In other words, the dimension of subspace in $\mathbb{R}^{s+1}$ generated by $w_2, w_3,\ Ae * w_2,\ Aw_2$ equals 2.

Simplifying assumptions of level 4: $\dim B_4' = 2$, that is $\dim M_4' = 4$.
In other words, the dimension of subspace in $\mathbb{R}^{s+1}$ generated by $w_2, w_3,\ Ae * w_2,\ Aw_2,\ w_4,\ Ae{*}w_3,\ Aw_3, w_2 * w_2$ equals 4.

Simplification of level 3

Now more detils on simplification of level 3.
The condition of the linear dependency of the generating vectors implies that everything can be expressed in terms of w_2 and w_3:

$$\begin{aligned} d \cdot Aw_2 &= a_{32}c_2^2(c_2 \cdot w_2 - w_3)\,, \\ d \cdot Ae * w_2 &= (3c_2 - 2c_3)c_2^2 a_{32} \cdot w_2 - (c_2 - c_3)(2a_{32}c_2 - c_3^2) \cdot w_3\,, \end{aligned}$$

where $d = a_{32}c_2^2 + c_3^2(c_2 - c_3)$.

If in addition, the simplifying assumption of level 2 holds and among all the b_i-s, only $b_2 = 0$, then we can simplify further:

$$\begin{aligned} Ae * w_2 &= c_2 w_2\,, \\ Aw_2 &= \frac{c_2}{2c_3}(-c_2 w_2 + w_3)\,. \end{aligned}$$

Conclusion

1. Usual computer algebra do not allow to find higher-order RK methods.
2. Introduction of upper and lower Butcher algebra allows a much better understanding the structure of the order conditions.
3. Using Butcher algebras opens the way to finding the RK methods of arbitrarily high order. In particular, order 9 methods was found, using this approach.
4. Learning of Butcher algebra of an arbitrary square matrix is an independent, interesting mathematical problem, even apart from the RK methods.

Thank you!!!!