Adaptive Recombination of Cubature Formulas

Christian Schmeiser, Alexander Soreff and Josef Teichmann

25. Januar 2006

1 Cubature Formulas

Cubature Formulas – in the classical sense – serve for approximate calculation of integrals.

A cubature formula of order $m \geq 1$ is given through points $x_1, \ldots, x_r \in \mathbb{R}^N$ and weights $\lambda_1, \ldots, \lambda_r > 0$ such that

$$\int_{\mathbb{R}^N} f(y)\mu(dy) = \sum_{i=1}^{r} \lambda_i f(x_i)$$

for polynomials f up to degree m. If – given a measure μ – we are given a Cubature Formula, this usually yields a quick method for the calculation of the integral by approximation of generic function f through polynomials.

2 Chakalov's Theorem

With methods from convex analysis one can easily prove

Theorem 2.1 *Let* μ *be a Borel measure on* $\mathbb{R}^N$ *with finite moments up to order* $m \geq 1$*, then there are points* $x_1, \ldots, x_r \in \operatorname{supp}(\mu)$ *and weights* $\lambda_j > 0$ *for* $j = 1, \ldots, r$ *such that*

$$\int_{\mathbb{R}^N} f(y)\mu(dy) = \sum_{j=1}^{r} f(x_j)\lambda_j$$

for all polynomials on $\mathbb{R}^N$ *up to degree* m*. Furthermore one can bound* r *by the dimension of the vector space of polynomials up to degree* m*.*

The concrete construction of cubature formulas with few points is a difficult, often unsolved problem. In concrete cases there are usually too many points with too small weights.

3 Stochastic Taylor Expansion

Given C^∞-bounded vector fields $V^0, \ldots, V^d$. Let $(Y_t^y)_{t \geq 0}$ be solution of

$$dY_t^y = \sum_{i=0}^{d} V^i(Y_t^y) \circ dB_t^i,$$
$$Y_0^y = y,$$

in $\mathbb{R}^N$, then the following aysymptotic expansion holds:

Denote by $\mathcal{A}$ the set of finite sequences $(i_1, \ldots, i_k)$ with degree-function

$$\deg(i_1, \ldots, i_k) = k + \#\{j,\ i_j = 0\},$$

then for $m \geq 1$ and C^∞-bounded f

$$f(Y_T^y) = \sum_{\text{deg}(i_1,\dots,i_k)\leq m} V^{i_1}\cdots V^{i_k} f(y) \int_{0\leq t_1\leq\cdots\leq t_k\leq T} \circ dB_{t_1}^{i_1}\cdots\circ dB_{t_k}^{i_k} + \\ + R_m(f,T,y)$$

with remainder term estimated via $||R_m(f,t,y)||_2 = \mathcal{O}(T^{\frac{m+1}{2}})$ as $T \to 0$.

4 Cubature Formulas

General diffusions can be described asymptotically by iterated Stratonich integrals (ISI): fix $m \geq 1$, we put all ISI up to degree m in a "polynomial"

$$X_t(B) := \sum_{\deg(i_1,\dots,i_k)\leq m} \int_{0\leq t_1\leq\cdots\leq t_k\leq t} \circ dB^{i_1}_{t_1} \cdots \circ dB^{i_k}_{t_k} \times e_{i_1} \cdots e_{i_k}.$$

By Chakaoloff's Theorem we can find a Cubature Formula of degree 1 for the random variable $X_t(B)$, i.e. we find

$$E(X_t(B)) = \sum_{i=1}^{r} \lambda_i x_i,$$

where $x_i \in \text{supp}(X_t(B))$, $i = 1, \dots, r$.

We define additionally

$$X_t(\omega) := \sum_{\deg(i_1,\dots,i_k) \le m} e_{i_1} \cdots e_{i_k} \times \int_{0 \le t_1 \le \cdots \le t_k \le t} d\omega^{i_1}(t_1) \cdots d\omega^{i_k}(t_k)$$

where we apply $\omega^0(t) := t$. A classical result of Chow tells, that there are H^1-trajectories $\omega_1, \dots, \omega_r : [0, T] \to \mathbb{R}^d$ and $\lambda_1, \dots, \lambda_r > 0$ such that

$$E(X_t(B)) = \sum_{j=1}^{r} \lambda_j X_t(\omega_j).$$

Notice also that ω_i depend on T, they become rougher if $T \to 0$, given ω_i for $T = 1$, we can construct those for arbitrary T through $T^{\frac{1}{2}} \omega_i(\frac{\cdot}{T})$, see Lyons-Victoir (2004).

5 Asymptotic Analysis for expected values

This leads to the following asymptotic formulas,

$$dY_t^y = \sum_{i=0}^{d} V^i(Y_t^y) \circ dB_t^i,$$
$$Y_0^y = y,$$

in $\mathbb{R}^N$:

- fix an order $m \geq 1$.

- there are H^1-trajectories $\omega_1, \ldots, \omega_r : [0, T] \to \mathbb{R}^d$ and weights $\lambda_1, \ldots, \lambda_r > 0$. The trajectories ω_j and their number r depend on d and m, but not on the specific vector fields $V^0, \ldots, V^d$.

- solve r non-autonomous ODEs (with a numerical scheme of order m)

$$\begin{aligned} dY_t^y(\omega_j) &= V^0(Y_t^y(\omega_j))dt + \\ &+ \sum_{i=1}^{d} V^i(Y_t^y(\omega_j))d\omega_j^i(t), \\ Y_0^y(\omega_j) &= y. \end{aligned}$$

Then

$$\begin{aligned} E(f(Y_T^y)) = \sum_{j=1}^{r} \lambda_j f(Y_T^y(\omega_j)) + \\ + \mathcal{O}(T^{\frac{m+1}{2}}). \end{aligned}$$

The constant in the order estimate depends in general on the $(m+1)$st derivative of f.

6 Iteration

We define the following linear operator on C^∞-bounded functions f,

$$Q_s f(y) := \sum_{j=1}^{r} \lambda_j f(Y_s^y(\omega_j)).$$

We obtain,

$$|E(f(Y_t^y)) - Q_{\frac{t}{n}} \circ \cdots \circ Q_{\frac{t}{n}} f(y)| \leq K||f||_{m+1} \frac{1}{n^{\frac{m-1}{2}}},$$

where $||f||_{m+2}$ is a complete norm on $C_b^{m+2}(\mathbb{R}^N)$.

Assuming furthermore uniform hypo-ellipticity, impressive results of Kusuoka (2001) yield the following structure:

- Subdivide $0 = t_0 < t_1 < \cdots < t_l = t$ with $s_i := t_i - t_{i-1}$ for $i = 1, \ldots, l$.

- for a Cubature formulas with degree of accuracy m we obtain the following error bound due to,

$$|E(f(Y_t^y)) - [Q_{s_1} \circ \cdots \circ Q_{s_l} f](y)| \leq K ||f||_{Lip} \left(\sum_{i=1}^{l-1} \frac{s_i^{\frac{m+1}{2}}}{(t - t_i)^{\frac{m}{2}}} + s_l^{\frac{1}{2}} \right),$$

where $||f||_{Lip}$ is the Lipschitz constant of f.

- the tree corresponding to a subdivision $0 = t_0 < t_1 < \cdots < t_l = t$ is encoded by the (Markov) process $(Z_i)_{i=0,\ldots,l}$ taking finitely many values in $\mathbb{R}^N$, i.e.

$$\begin{aligned} E(Z_i | Z_{i-1} = x) &= Z^x_{s_i}(\omega_j) \text{ with probability } \lambda_j,\ j = 1, \dots, r, \\ Z_0 &= y, \end{aligned}$$

for $i = 1, \dots, l$.

This yields a tree, which is in principle exponentially growing and which produces a random variable Z_l approximating – through the previous results – the random variable Y^y_t in a weak sense, i.e.

$$|E(f(Y^y_t)) - E(f(Z_l))| \le K ||f||_{Lip} \left(\sum_{i=1}^{l-1} \frac{s_i^{\frac{m+1}{2}}}{(t - t_i)^{\frac{m}{2}}} + s_l^{\frac{1}{2}} \right).$$

We shall work on the amount of calculations involved to obtain the process $E(f(Z_l))$.

7

8 Advantages of the Method

- the trajectories are universal.

- instead of solving PDEs or simulating random variables and doing MC-simulation, one has to solve ODEs.

- the method respects the geometry, one looks for "typical" trajectories up to time-asymptotics $T^{\frac{m}{2}}$.

- the method yields a deterministic weak approximation of the law with deterministic error bound!

9 Disadvantage of the Method

- the tree grows exponentially.

- too many points with too small weigths at the end of the day...

10 Recombination Structures

Observing the tree yields that – since there are many points – many of them are close! Gronwall's inequality limits the recombinations coming from this consideration.

Take the non-autonomous system for a H^1-curve $\omega : [0,T] \to \mathbb{R}^{d+1}$,

$$\frac{dY_t^y(\omega)}{dt} = \sum_{i=0}^{d} V_i(Y_t^y(\omega))\frac{d\omega^i}{dt},$$
$$Y_0^y(\omega) = y,$$

then there is a constant C (calculated from $V_0, \dots, V_d$) such that

$$||Y_t^y - Y_t^z|| \leq ||y - z|| \exp(C \int_0^t ||\frac{d}{dt}\omega(t)||_{\mathbb{R}^{d+1}} dt).$$

The second recobination is based on algebraic considerations. It can be applied a priori, i.e. telling that the resulting points on the tree at step $i+r$ ($r \geq 1$) of two calculations seen from step i will be near enough, such that one should identify them a priori and perform only one calculation. We shall need – for the formulation of the result, some algebraic preparations done in the following steps.

1. We abbreviate by $\mathcal{A}$ the set of all finite sequences $I := (i_1, \ldots, i_k) \in \{0, \ldots, d\}^k$, $k \geq 0$ and we define a degree function on $\mathcal{A}$ by

$$\deg(i_1, \ldots, i_k) := k + \#\{l \text{ with } i_l = 0\},$$

which simply means that the 0s appearing in $(i_1, \ldots, i_k)$ are counted twice. Additionally, we define $\deg(\emptyset) = 0$. We define a semi-group structure on $\mathcal{A}$ via

$$(i_1, \ldots, i_k) * (j_1, \ldots, j_l) := (i_1, \ldots, i_k, j_1, \ldots, j_l),$$
$$\emptyset * (i_1, \ldots, i_k) = (i_1, \ldots, i_k) * \emptyset = (i_1, \ldots, i_k).$$

2. We denote by $\mathbb{A}^m_{d,1}$the free, nilpotent algebra with $d+1$ generators $e_0, \ldots, e_d$, i.e. the set of all non-commutative polynomials in those variables, such that the following nilpotency relations hold: if $\deg(I) > m$, then $e_{i_1} e_{i_2} \cdots e_{i_k} = 0$ for all $I \in \mathcal{A}$. Hence $\mathbb{A}^m_{d,1}$ is a

finite dimensional, non-commutative, real algebra with unit element 1 and we are given a grading via the degree function, i.e. a monomial $e_{i_1}e_{i_2}\cdots e_{i_k}$ is said to have degree n if $\deg(i_1,\ldots,i_k)=n$. Denote by W_n the linear span of all monomials of degree n, then we obtain

$$\mathbb{A}^m_{d,1}=\oplus_{n=0}^m W_n,$$

furthermore $W_pW_q\subset W_{p+q}$ (where we define $W_p=0$ for $p>m$ due to the given relations), $W_0=\mathbb{R}\cdot 1$, so $\mathbb{A}^m_{d,1}$ is a graded algebra. We denote the canonical projections of x on the subspaces W_n by x_n for $n\geq 0$.

3. On the finite dimensional algebra $\mathbb{A}^m_{d,1}$ we define the exponential series

$$\exp(x):=\sum_{i=0}^{\infty}\frac{x^i}{i!},$$

where the series converges everywhere due to the nilpotency relations. We define the logarithm on elements x with $x_0 \neq 0$. We identify x_0 with a real number, so the series

$$\log(x) = \log(x_0) + \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i} (\frac{x - x_0}{x_0})^i.$$

is well-defined, since $\frac{x-x_0}{x_0} \in \oplus_{n=1}^{m} W_n$ (i.e. the series is finite).

4. We define the Lie algebra $\mathfrak{g}_{d,1}^m$ generated by $e_0, \ldots, e_d$ with respect to the Lie bracket $[x, y] := xy - yx$. The Lie algebra inherits the grading from the algebra $\mathbb{A}_{d,1}^m$ via $U_n := \mathfrak{g}_{d,1}^m \cap W_n$, for $n \geq 0$. Hence

$$\mathfrak{g}_{d,1}^m = \oplus_{n=1}^{m} U_n$$

is a graded Lie algebra. In fact the Lie algebra is free, nilpotent of step m, with d generators of degree 1 and one generator of degree 2.

5. We denote the exponential image of $\mathfrak{g}^m_{d,1}$ by $G^m_{d,1} := \exp(\mathfrak{g}^m_{d,1}) \subset \mathbb{A}^m_{d,1}$ and call it the free, nilpotent Lie group. $G^m_{d,1}$ is indeed a Lie group as closed subgroup of Lie group $1 + \oplus_{n=1}^m W_n$. The tangent space at $x \in G^m_d$ is spanned by the left (or right) translations xw for $w \in \mathfrak{g}^m_{d,1}$.

6. We need the canonical dilatations on Lie algebra and Lie group. We define for $x \in \mathbb{A}^m_{d,1}$ an algebra homomorphism Δ_t via

$$\Delta_t(x_n) = t^n x_n$$

for $x_n \in W_n$, $n \geq 1$ and $t > 0$. The homomorphism is well defined and restricts to a Lie algebra homomorphism Δ_t on $\mathfrak{g}^m_{d,1}$ and and a Lie group homomorphism Δ_t for $t > 0$.

7. Fix $d, m \geq 1$ and elements $\gamma_1, \ldots, \gamma_k \in G^m_{d,1}$. We consider the group Γ generated by $\gamma_1, \ldots, \gamma_k$ and define the word metric on Γ: for any

$g \in \Gamma$ there are $(i_1, \ldots, i_k)$ and integers $n_i \in \mathbb{Z}$ such that

$$g = \gamma_{i_1}^{n_1} \cdots \gamma_{i_m}^{n_m},$$

the word metric is defined through $|g| = \min\{|n_1| + \cdots + |n_m|\}$, where the minimum is taken over all representations of g as above. Hence $|g| = |g^{-1}|$ and we can define $d(g,h) = |g^{-1}h|$ for $g, h \in \Gamma$. The ball of radius r is denoted by $B(r)$, i.e.

$$B(r) = \{g \text{ such that } d(g,e) \leq r\}$$

and we define the volume function $V(r) := \#B(r)$ as the number of elements of $B(r)$ for $r \geq 1$. We call Γ of polynomial growth if there is a polynomial $p(r)$ such that $V(r) \leq p(r)$ for $r \geq 1$. The degree of p is called the growth of Γ.

8. Fix $d, m \geq 1$ and elements $\gamma_1, \ldots, \gamma_k \in G_{d,1}^m$. The central descending

series of of normal subgroups of Γ is defined by

$$\Gamma^1 := \Gamma,$$
$$\Gamma^{j+1} := [\Gamma, \Gamma^j],$$

where $[\Gamma, \Gamma^j]$ denotes the set of all commutators $[g, h] := ghg^{-1}h^{-1}$ of elements $g \in \Gamma$ and $h \in \Gamma^j$, for $j \geq 1$, which is itself a normal subgroup Γ^j.

9. Given $d, m \geq 1$ and elements $\gamma_1, \ldots, \gamma_k \in G^m_{d,1}$, then the growth of the group Γ generated by $\gamma_1, \ldots, \gamma_k$ is polynomial and equals

$$\sum_j \text{rank}(\Gamma^j / \Gamma^{j+1}),$$

which is a finite series, since certainly $\Gamma^{m+1} = \{e\}$.

Fix $m, d \geq 1$, and fix cubature trajectories $\omega_1, \ldots, \omega_r : [0, s] \to \mathbb{R}^{d+1}$. For fixed $i \geq 1$ we define – via $\omega_1, \ldots, \omega_r$ – group elements in $G^m_{d,1}$of the form

$$X_s^{x,j} := \sum_{\substack{I \in \mathcal{A} \\ \deg(I) \leq m}} x e_{i_1} \cdots e_{i_k} \int_{0 \leq t_1 \leq \cdots \leq t_k \leq s} d\omega_j^{i_1}(t_1) \cdots d\omega_j^{i_k}(t_k),$$

for $j = 1, \ldots, r$. We are interested in the sets T_n, which are defined recursively for $n \geq 0$, i.e.

$$T_0 = \{e\},$$
$$T_{n+1} = X_s^{x,j} \text{ for } j = 1, \ldots, r \text{ and } x \in T_n.$$

Given $d, m \geq 1$, then the number of elements in T_n grows polynomially with n.

Obviously $T_n \subset B(n)$ for $n \geq 0$ with respect to the word metric in the group generated by $X_t^{e,j}$ for $j = 1, \ldots, r$. Since this group has a finite central descending series, the growth is a finite number.

Observe that the set T_n corresponds to the points in $G_{d,1}^m$, which are attained by the random walk $(R_m)_{m \geq 0}$ after n steps. This random walk is defined via

$$
\begin{aligned}
R_0 &= e, \\
R_{n+1} &= X_s^{R_n,j} \text{ with probability } \lambda_j \text{ for } j = 1, \ldots, r.
\end{aligned}
$$

Notice that the random variable R_n for $n \geq 0$ can also be understood as the set of points, which are obtained by solving the autonomous equation along the trajectories $\omega_{j_1} * \cdots * \omega_{j_n}$ for some $j_k \in \{1, \ldots, r\}$. Here we apply

$$
\omega * \eta : [0, 2s] \to \mathbb{R}^{d+1},
$$

which is given through gluing of $\omega, \eta : [0, s] \to \mathbb{R}^{d+1}$, i.e.

$$\omega * \eta(t) = \omega(t) \text{ for } t \in [0, s],$$
$$\omega * \eta(t) = \eta(t) + \omega(s) - \eta(s) \text{ for } t \in [s, 2s].$$

Given two trajectories $\omega = \omega_{j_1} * \cdots * \omega_{j_n}$ and $\widetilde{\omega} = \omega_{\widetilde{j_1}} * \cdots * \omega_{\widetilde{j_n}}$ on the time interval $[0, ns]$, such that there is $\widehat{m} \geq m$ with $G^{\widehat{m}}_{d,1}$ with $X^e_{ns}(\omega) = X^e_{ns}(\widetilde{\omega})$ (understood in $G^{\widehat{m}}_{d,1}$!), then

$$||Z_n(\omega) - Z_n(\widetilde{\omega})|| \leq C(ns)^{\widehat{m}+1}[\max_{j,[0,ns]} |\frac{d\omega^j}{dt}|^{\widehat{m}+1} - \max_{j,[0,ns]} |\frac{d\widetilde{\omega}}{dt}^j|^{\widehat{m}+1}] = \mathcal{O}(n^{\widehat{m}+1} s$$