Random time steps geometric integrators of ordinary differential equations for uncertainty quantification of numerical errors

Assyr Abdulle, Giacomo Garegnani

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Swiss Numerics Day, 20 April 2018

Outline

Probabilistic methods – why?

Consider Lorenz equation (atmospheric convection)

$$\begin{aligned} x' &= \sigma(y - x), && x(0) = -10, \\ y' &= x(\rho - z) - y, && y(0) = -1, \\ z' &= xy - \beta z, && z(0) = 40. \end{aligned}$$

For $\rho = 28$, $\sigma = 10$, $\beta = 8/3$ chaotic behaviour.

$\Longrightarrow$ Numerical integration gives unreliable solutions.

Probabilistic methods – why?

Consider Lorenz equation (atmospheric convection)

$$\begin{aligned} x' &= \sigma(y - x), & x(0) &= -10, \\ y' &= x(\rho - z) - y, & y(0) &= -1, \\ z' &= xy - \beta z, & z(0) &= 40. \end{aligned}$$

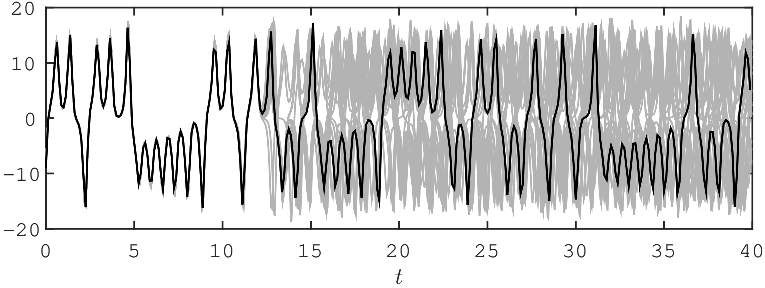
For $\rho = 28$, $\sigma = 10$, $\beta = 8/3$ chaotic behaviour.

$\Longrightarrow$ Numerical integration gives unreliable solutions.

Goal

Establish a probability measure over the numerical solution given by classical methods.

Probabilistic methods – why?

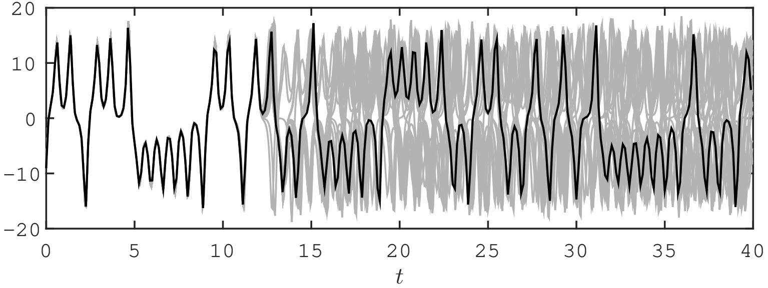


Time evolution of the first component of Lorenz equation

Black line $\rightarrow$ deterministic solution

Gray lines $\rightarrow$ family of perturbed solutions.

Probabilistic methods – why?



Time evolution of the first component of Lorenz equation

Black line $\rightarrow$ deterministic solution

Gray lines $\rightarrow$ family of perturbed solutions.

Chaotic behaviour appears frequently in nonlinear differential equations.

Notation

Autonomous dynamical system, function $f\colon \mathbb{R}^d \to \mathbb{R}^d$ and the ODE

$$y' = f(y), \quad y(0) = y_0.$$

Flow of the equation $\varphi_t\colon \mathbb{R}^d \to \mathbb{R}^d$ such that

$$y(t) = \varphi_t(y_0).$$

One-step method (e.g. Runge Kutta): numerical flow Ψ_h such that

$$y_{n+1} = \Psi_h(y_n).$$

Probabilistic methods for ODEs

Filtering methods for ODEs: fix a prior on $y(t)$ (Gaussian process), update with evaluations of $f(y)$
[Kersting and Hennig, 2016, Chkrebtii et al., 2016]

Randomised methods for ODEs: random perturbation of deterministic numerical solutions $\rightarrow$ sampling

- Additive noise [Conrad et al., 2016],
- Random time steps [Abdulle and Garegnani, 2018].

Probabilistic methods for ODEs

Filtering methods for ODEs: fix a prior on $y(t)$ (Gaussian process), update with evaluations of $f(y)$
[Kersting and Hennig, 2016, Chkrebtii et al., 2016]

Randomised methods for ODEs: random perturbation of deterministic numerical solutions $\rightarrow$ sampling

- Additive noise [Conrad et al., 2016],
- Random time steps [Abdulle and Garegnani, 2018].

Additive noise method

Stochastic process $\{Y_n\}_{n=1,2,\ldots}$ with recurrence

$$Y_{n+1} = \underbrace{\Psi_h(Y_n)}_{\text{deterministic}} + \underbrace{\xi_n(h)}_{\text{random}}.$$

Main assumption: $\{\xi_n\}_{n=0,1,\ldots}$ iid such that for $p > 1$ and $Q \in \mathbb{R}^{d\times d}$

$$\mathbb{E}\,\xi_n(h) = 0, \quad \mathbb{E}\,\xi_n(h)\xi_n(h)^T = Qh^{2p+1}.$$

Additive noise method

Stochastic process $\{Y_n\}_{n=1,2,\dots}$ with recurrence

$$Y_{n+1} = \underbrace{\Psi_h(Y_n)}_{\text{deterministic}} + \underbrace{\xi_n(h)}_{\text{random}}.$$

Main assumption: $\{\xi_n\}_{n=0,1,\dots}$ iid such that for $p > 1$ and $Q \in \mathbb{R}^{d\times d}$

$$\mathbb{E}\,\xi_n(h) = 0, \quad \mathbb{E}\,\xi_n(h)\xi_n(h)^T = Qh^{2p+1}.$$

Properties

If Ψ_h is of order q and for $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ smooth

- Strong convergence: $\mathbb{E}\|y(hn) - Y_n\| \leq Ch^{\min\{p,q\}}$,
- Weak convergence: $|\Phi\big(y(hn)\big) - \mathbb{E}\,\Phi(Y_n)| \leq Ch^{\min\{2p,q\}}$,
- Good qualitative behavior in Bayesian inverse problems.

Additive noise method

Stochastic process $\{Y_n\}_{n=1,2,\ldots}$ with recurrence

$$Y_{n+1} = \underbrace{\Psi_h(Y_n)}_{\text{deterministic}} + \underbrace{\xi_n(h)}_{\text{random}}.$$

Main assumption: $\{\xi_n\}_{n=0,1,\ldots}$ iid such that for $p > 1$ and $Q \in \mathbb{R}^{d\times d}$

$$\mathbb{E}\,\xi_n(h) = 0, \quad \mathbb{E}\,\xi_n(h)\xi_n(h)^T = Qh^{2p+1}.$$

Issues

- Robustness: $\Psi_h(Y_{n-1}) > 0 \;\not\Rightarrow\; \mathbb{P}(Y_n < 0) = 0$,
- Geometric properties are not conserved from Ψ_h. For example if $I(y) = y^T S y$ and $I(\Psi_h(y_0)) = I(y_0)$

$$I(Y_1) = I(y_0) + 2\xi_0(h)^T S \Psi_h(y_0) + \xi_0(h)^T S \xi_0(h).$$

Random time steps

Intrinsic noise: Random time-stepping Runge-Kutta (RTS-RK)

$$Y_{n+1} = \Psi_{H_n}(Y_n),$$

Main assumption: $\{H_n\}_{n=0,1,\dots}$ iid such that for $h, C > 0$ and $p > 1$

$$H_n > 0 \text{ a.s.}, \quad \mathbb{E}\, H_n = h, \quad \operatorname{Var} H_n = Ch^{2p}.$$

Example: $H_n \overset{\text{iid}}{\sim} \mathcal{U}(h - h^p, h + h^p)$.

Random time steps

Intrinsic noise: Random time-stepping Runge-Kutta (RTS-RK)

$$Y_{n+1} = \Psi_{H_n}(Y_n),$$

Main assumption: $\{H_n\}_{n=0,1,\dots}$ iid such that for $h, C > 0$ and $p > 1$

$$H_n > 0 \text{ a.s.}, \quad \mathbb{E}\, H_n = h, \quad \operatorname{Var} H_n = Ch^{2p}.$$

Example: $H_n \overset{\text{iid}}{\sim} \mathcal{U}(h - h^p, h + h^p)$.

Properties

If Ψ_h is of order q and for $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ smooth

- Strong convergence: $\mathbb{E}\|y(hn) - Y_n\| \leq Ch^{\min\{p-1/2, q\}}$,
- Weak convergence: $|\Phi\big(y(hn)\big) - \mathbb{E}\,\Phi(Y_n)| \leq Ch^{\min\{2p-1, q\}}$,
- Good qualitative behavior in Bayesian inverse problems.

Random time steps

Intrinsic noise: Random time-stepping Runge-Kutta (RTS-RK)

$$Y_{n+1} = \Psi_{H_n}(Y_n),$$

Main assumption: $\{H_n\}_{n=0,1,\dots}$ iid such that for $h, C > 0$ and $p > 1$

$$H_n > 0 \text{ a.s.}, \quad \mathbb{E}\, H_n = h, \quad \operatorname{Var} H_n = Ch^{2p}.$$

Example: $H_n \overset{\text{iid}}{\sim} \mathcal{U}(h - h^p, h + h^p)$.

Properties (Geometric)

- Conservation of (polynomial) first integrals is inherited by Ψ_h,
- Flow map is symplectic if Ψ_h is symplectic,
- Long-time conservation of energy in Hamiltonian systems.

Numerical experiment – Geometric properties

Consider the perturbed Kepler equation (model for two-body problem)

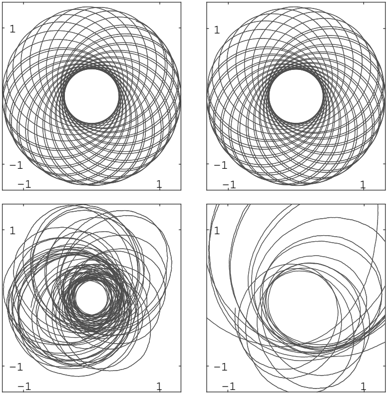
$$q_1' = p_1, \quad p_1' = -\frac{q_1}{\|q\|^3} - \frac{\delta q_1}{\|q\|^5},$$
$$q_2' = p_2, \quad p_2' = -\frac{q_2}{\|q\|^3} - \frac{\delta q_2}{\|q\|^5}.$$

The angular momentum is conserved (quadratic first integral)

$$I(p, q) = q_1 p_2 - q_2 p_1$$

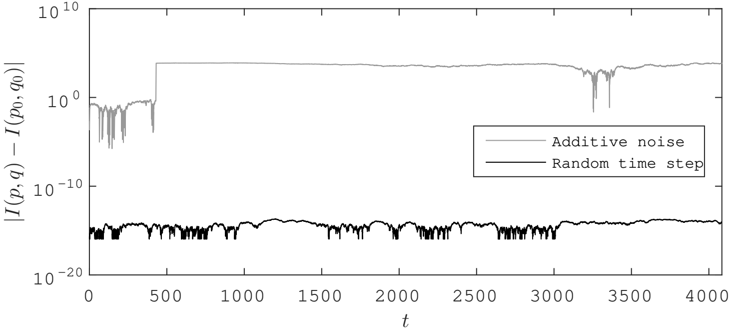
$\rightarrow$ employ a Gauss method (implicit midpoint rule).

Numerical experiment – Geometric properties



RTS-RK (first row), Additive noise (second row). Time $0 \leq t \leq 200$ and $200 \leq t \leq 400$ (left and right)

Numerical experiment – Geometric properties



Conservation of the angular momentum (quadratic first integral)

Bayesian inverse problems

Goal

Given $\vartheta \in \mathbb{R}^n$, $f_\vartheta \colon \mathbb{R}^d \to \mathbb{R}^d$ and the ODE

$$y' = f_\vartheta(y), \quad y(0) = y_{0,\vartheta} \in \mathbb{R}^d,$$

retrieve the true value ϑ^* from observations of $y(t)$, $t > 0$.

Bayesian inverse problems

Goal

Given $\vartheta \in \mathbb{R}^n$, $f_\vartheta \colon \mathbb{R}^d \to \mathbb{R}^d$ and the ODE

$$y' = f_\vartheta(y), \quad y(0) = y_{0,\vartheta} \in \mathbb{R}^d,$$

retrieve the true value ϑ^* from observations of $y(t)$, $t > 0$.

Bayesian setting: fix prior $\pi_{\mathrm{prior}}(\vartheta)$, consider the forward operator $\mathcal{G}$ and model observations as

$$\underbrace{\mathcal{Y}}_{\text{observations}} = \underbrace{\mathcal{G}(\vartheta^*)}_{\text{forward}} + \underbrace{\varepsilon}_{\text{noise}}, \quad \varepsilon \sim \pi_{\mathrm{noise}},$$

then the posterior distribution (density) is

$$\pi(\vartheta \mid \mathcal{Y}) \propto \pi_{\mathrm{prior}}(\vartheta)\pi_{\mathrm{noise}}(\mathcal{Y} - \mathcal{G}(\vartheta)).$$

Bayesian inverse problems

The posterior $\pi(\vartheta \mid \mathcal{Y})$ is not computable, approximate with

$$\pi^h(\vartheta \mid \mathcal{Y}) \propto \pi_{\mathrm{prior}}(\vartheta)\pi_{\mathrm{noise}}(\mathcal{Y} - \mathcal{G}^h(\vartheta)).$$

Bayesian inverse problems

The posterior $\pi(\vartheta \mid \mathcal{Y})$ is not computable, approximate with

$$\pi^h(\vartheta \mid \mathcal{Y}) \propto \pi_{\text{prior}}(\vartheta)\pi_{\text{noise}}(\mathcal{Y} - \mathcal{G}^h(\vartheta)).$$

Properties

$\pi^h \to \pi$ for $h \to 0$ (in the Hellinger distance).

Issue

- π^h concentrated around values "far" from $\vartheta^* \to$ non-predictive posterior

Bayesian inverse problems

The posterior $\pi(\vartheta \mid \mathcal{Y})$ is not computable, approximate with

$$\pi^{\boldsymbol{h},\mathrm{RTS}}(\vartheta \mid \mathcal{Y}) \propto \pi_{\mathrm{prior}}(\vartheta)\mathbb{E}^{\mathbf{H}}\pi_{\mathrm{noise}}(\mathcal{Y} - \mathcal{G}^{\mathbf{H}}(\vartheta)),$$

where $\mathbf{H} = (H_0, H_1, \dots)$ is the vector of all time steps chosen in one run.

Bayesian inverse problems

The posterior $\pi(\vartheta \mid \mathcal{Y})$ is not computable, approximate with

$$\pi^{h,\mathrm{RTS}}(\vartheta \mid \mathcal{Y}) \propto \pi_{\mathrm{prior}}(\vartheta)\mathbb{E}^{\mathbf{H}}\pi_{\mathrm{noise}}(\mathcal{Y} - \mathcal{G}^{\mathbf{H}}(\vartheta)),$$

where $\mathbf{H} = (H_0, H_1, \ldots)$ is the vector of all time steps chosen in one run.

Properties

- $\pi^{h,\mathrm{RTS}} \to \pi$ for $h \to 0$ (in the Hellinger distance). [Lie et al., 2017]
- "correct" the non-predictive behaviour of deterministic approximations

Warning

- Approximation of $\mathbb{E}^{\mathbf{H}}\,\pi_{\mathrm{noise}}(\mathcal{Y} - \mathcal{G}^{\mathbf{H}}(\vartheta))$ is required

Numerical experiment – Bayesian inverse problems

Consider the Hénon-Heiles system (motion of a star around a galactic center), Hamiltonian with energy

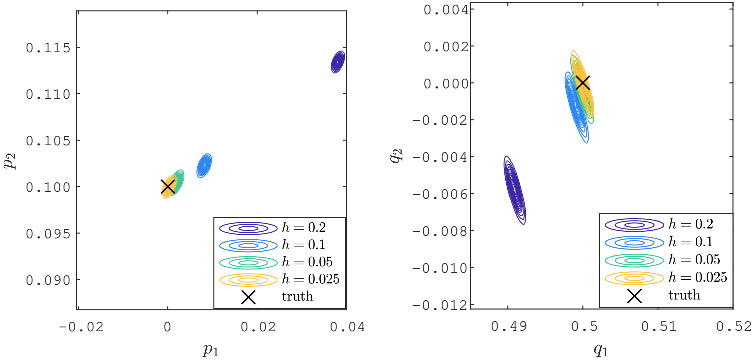
$$E(p, q) = \frac{1}{2}\|p\|^2 + \frac{1}{2}\|q\|^2 + q_1^2 q_2 - \frac{1}{3} q_2^3.$$

Chaotic problem for certain levels of energy.

Goal

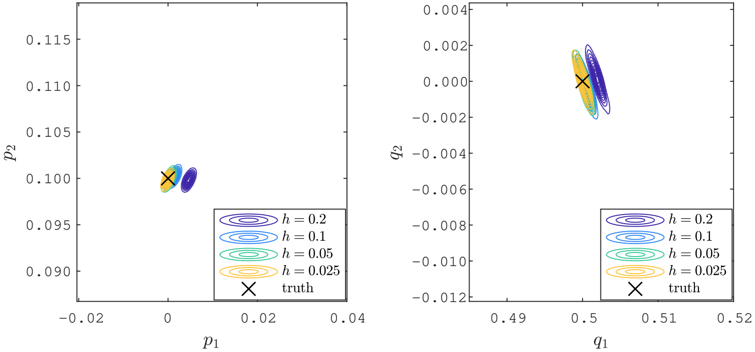
Find posterior $\pi((p_0, q_0) \mid \mathcal{Y})$ over the initial condition from a single observation of $(p(10), q(10))$

Numerical experiment – Bayesian inverse problems



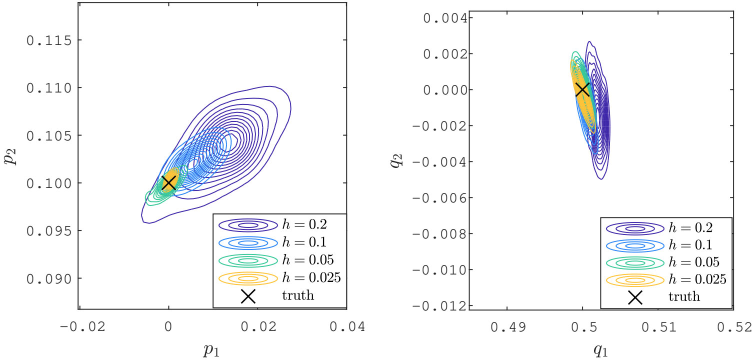
Posterior distributions given by deterministic Heun method.

Numerical experiment – Bayesian inverse problems



Posterior distributions given by deterministic Störmer-Verlet method.

Numerical experiment – Bayesian inverse problems



Posterior distributions given by RTS-RK Störmer-Verlet method.

References

[Abdulle and Garegnani, 2018] A. Abdulle, G. Garegnani (2018).
Random time step probabilistic methods for uncertainty quantification in chaotic and geometric numerical integration.
preprint arXiv:1801.01340

[Chkrebtii et al., 2016] O. A. Chkrebtii, D. A. Campbell, B. Calderhead, M. A. Girolami (2016).
Bayesian solution uncertainty quantification for differential equations.
Bayesian Anal.

[Conrad et al., 2016] Conrad, P. R., Girolami, M., Särkkä, S., Stuart, A., and Zygalakis, K. (2016).
Statistical analysis of differential equations: introducing probability measures on numerical solutions.
Stat. Comput.

[Kersting and Hennig, 2016] Kersting, H. and Hennig, P. (2016).
Active uncertainty calibration in Bayesian ODE solvers.
In *Proceedings of the 32nd Conference on Uncertainty in Artificial Intelligence (UAI 2016)*, pages 309–318. AUAI Press.

[Lie et al., 2017] Lie, H. C., Sullivan, T. J., and Teckentrup, A. L. (2017).
Random forward models and log-likelihoods in Bayesian inverse problems.
preprint arXiv:1712.05717