Stochastic contraction
BACS Workshop
Chamonix, January 14, 2008

Q.-C. Pham N. Tabareau J.-J. Slotine

Nice features of deterministic contraction theory:

- Powerful stability analysis tool for nonlinear systems
- Combination properties (modularity and stability)
- Hybrid, switching systems
- (Concurrent) synchronization in large-scale systems

Why stochastic contraction?

Goal: extend contraction theory to the stochastic case

- Analyze real-life systems, which are typically subject to random perturbations
- Benefit from the nice features of contraction theory

Modelling the random perturbations

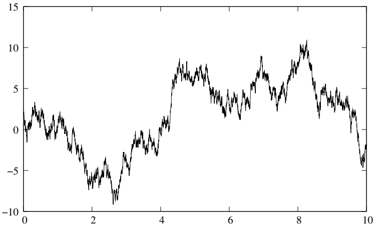
In physics, engineering, neuroscience, finance,... random perturbations are traditionnally modelled with Itô stochastic differential equations (SDE)

$$d\mathbf{x} = \mathbf{f}(\mathbf{x}, t)dt + \sigma(\mathbf{x}, t)dW$$

- $\mathbf{f}$ is the dynamics of the noise-free version of the system
- σ is the noise variance matrix (noise intensity)
- W is a Wiener process ($dW/dt =$ "white noise")

Some notions of stochastic modelling : Random walk and Wiener process

- Random walk (discrete-time): $\mathbf{x}_{t+\Delta t} = \mathbf{x}_t + \xi_t \Delta t$
  where $(\xi_t)_{t\in\mathbb{N}}$ are Gaussian and mutually independent
- If one is interested in very rapidly varying perturbations, Δt has to be very small
- Wiener process (or Brownian motion) (continuous-time): limit of the random walk when $\Delta t \to 0$



Some notions of stochastic modelling : Wiener process and "white noise"

- Problem: a Wiener process is not differentiable (why?), thus it is not the solution of any ordinary differential equation
- Define formally ξ_t ("white noise") = "derivative" of the Wiener process
- Formally: $W(t) - W(0) = \int_0^t \xi_t dt$ or $dW/dt = \xi_t$ or $dW = \xi_t dt$

Some notions of stochastic modelling : Iô SDE

- Stochastic differential equation :

$$d\mathbf{x}/dt = \mathbf{f}(\mathbf{x}, t) + \sigma(\mathbf{x}, t)\xi_t$$

or by mutiplying by dt :

$$d\mathbf{x} = \mathbf{f}(\mathbf{x}, t)dt + \sigma(\mathbf{x}, t)dW$$

- The last equation was made rigourous by K. Itô in 1951

The stochastic contraction theorem

- If the noise-free system is contracting

$$\lambda_{\max}(\mathbf{J}_s) \leq -\lambda$$

The stochastic contraction theorem

- If the noise-free system is contracting

$$\lambda_{\max}(\mathbf{J}_s) \leq -\lambda$$

- and the noise variance is upper-bounded

$$\mathrm{tr}\left(\sigma(\mathbf{x},t)^T\sigma(\mathbf{x},t)\right) \leq C$$

The stochastic contraction theorem

- If the noise-free system is contracting

$$\lambda_{\max}(\mathbf{J}_s) \leq -\lambda$$

- and the noise variance is upper-bounded

$$\mathrm{tr}\left(\sigma(\mathbf{x},t)^T\sigma(\mathbf{x},t)\right) \leq C$$

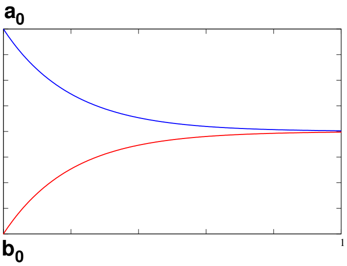
Then after exponential transients, the mean square distance between any two trajectories is upper-bounded by C/λ

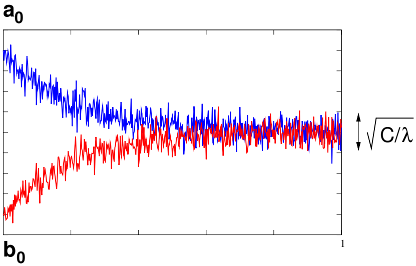
$$\forall t \geq 0 \quad \mathbb{E}\left(\|\mathbf{a}(t)-\mathbf{b}(t)\|^2\right) \leq \frac{C}{\lambda} + \left[\|\mathbf{a}_0-\mathbf{b}_0\|^2 - \frac{C}{\lambda}\right]^+ e^{-2\lambda t}$$

Practical meaning

After exponential transients, we have

$$\mathbb{E}\left(\|\mathbf{a}(t)-\mathbf{b}(t)\|\right) \leq \sqrt{\frac{C}{\lambda}}$$





BACS

Vocabulary and remarks

- We say that a system that verifies the conditions of the stochastic contraction theorem is stochastically contracting with rate λ and bound C
- Discrete and continuous-discrete versions of the theorem are available
- The theorem can be easily generalized to time-varying metrics

"Optimality" of the theorem

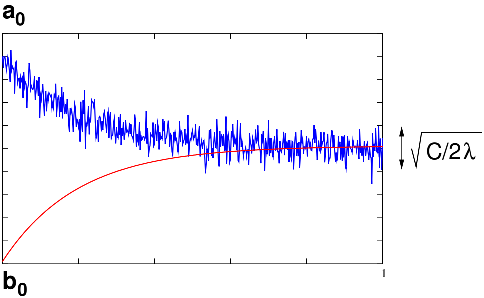
- The mean square bound in the theorem is optimal (consider an Ornstein-Uhlenbeck process $d\mathbf{x} = -\lambda \mathbf{x} dt + \sigma dW$ where the bound is attained)
- In general, one cannot obtain asymptotic almost-sure stability (consider again the Ornstein-Uhlenbeck process)

BACS

Noisy and noise-free trajectories

The theorem can be used to compare noisy and noise-free versions of a (stochastically) contracting system

$$d\mathbf{a} = \mathbf{f}(\mathbf{a}, t)dt + \sigma(\mathbf{a}, t)dW$$
$$d\mathbf{b} = \mathbf{f}(\mathbf{b}, t)dt$$



- Any contracting system is automatically protected against white noise (robustness)
- Very useful in applications (see later)

Combinations of stochastically contracting systems

Combinations results in deterministic contraction can be adapted very naturally for stochastic contraction

- Parallel combinations
- Hierarchical combinations
- Negative feedback combinations
- Small gains

Example: Negative feedback combination

- Two systems coupled by negative feedback gain k

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_1 & -k\mathbf{J}_{21}^T \\ \mathbf{J}_{21} & \mathbf{J}_2 \end{pmatrix}$$

- System 1 stochastically contracting with rate λ_1 and bound C_1
- System 2 stochastically contracting with rate λ_2 and bound C_2
- Then the coupled system is stochastically contracting with rate $\min(\lambda_1, \lambda_2)$ and bound $C_1 + kC_2$

Application: contracting observers and noisy measurements

- Consider the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$ with the measurements $\mathbf{y} = \mathbf{H}(t)\mathbf{x}$
- Typically $\dim(\mathbf{y}) < \dim(\mathbf{x})$
- Recall the deterministic contracting observer

$$\dot{\hat{\mathbf{x}}} = \mathbf{f}(\hat{\mathbf{x}}, t) + \mathbf{K}(t)(\hat{\mathbf{y}} - \mathbf{y})$$

$$\text{i.e.} \quad \dot{\hat{\mathbf{x}}} = \mathbf{f}(\hat{\mathbf{x}}, t) + \mathbf{K}(t)(\mathbf{H}(t)\hat{\mathbf{x}} - \mathbf{H}(t)\mathbf{x})$$

- If $\mathbf{K}$ is chosen such that $\left(\frac{\partial \mathbf{f}(\mathbf{x},t)}{\partial \mathbf{x}} - \mathbf{K}(t)\mathbf{H}(t)\right)$ is negative definite, then the observer system is contracting
- Since actual state of the system $\mathbf{x}$ is a particular solution of the observer system, the state of the observer $\hat{\mathbf{x}}$ will exponentially converge to $\mathbf{x}$

BACS

Application: contracting observers and noisy measurements

- Now, the measurements are corrupted by "white noise"

$$\mathbf{y} = \mathbf{H}(t)\mathbf{x} + \Sigma(t)\xi(t)$$

- Using the formal rule $\xi(t)dt = dW$, the observer equation becomes

$$d\hat{\mathbf{x}} = (\mathbf{f}(\hat{\mathbf{x}}, t) + \mathbf{K}(t)(\mathbf{H}(t)\mathbf{x} - \mathbf{H}(t)\hat{\mathbf{x}}))dt + \mathbf{K}(t)\Sigma(t)dW$$

- Using the same **K** as earlier, the system is stochastically contracting with rate λ and bound C where

$$\lambda = \inf_{\mathbf{x},t} \left| \lambda_{\max}\left(\frac{\partial \mathbf{f}(\mathbf{x}, t)}{\partial \mathbf{x}} - \mathbf{K}(t)\mathbf{H}(t) \right) \right|$$

$$C = \sup_{t \geq 0} \mathrm{tr}(\Sigma(t)^T \mathbf{K}(t)^T \mathbf{K}(t)\Sigma(t))$$

- After exponential transients, $\|\hat{\mathbf{x}} - \mathbf{x}\| \leq \sqrt{C/2\lambda}$

Application: stochastic synchronization

See next talk by Nicolas

Current directions of research

- Extension to space-dependent metrics
- Stochastic contraction analysis of Kalman filters (which are a Bayesian filters) and other Bayesian algorithms

Thank you for your attention!