Coupling and Control

Warwick MIR@AW workshop: Inference and Control for Complex Dynamical Systems

Wilfrid S. Kendall
w.s.kendall@warwick.ac.uk

Department of Statistics, University of Warwick

26th November 2012

Introduction

Aims of this talk:

- to describe the technique of probabilistic coupling,
- to show how it can be viewed as a stochastic control problem,
- and to discuss some recent developments.

Examples

Three examples of probabilistic coupling in action:

1. Proof of equilibrium theorem for finite irreducible aperiodic Markov chains (Doeblin 1938).
2. Competing epidemics (WSK and Saunders 1983).
3. Heating igloos (WSK 1989).

Doeblin's coupling for Markov chains

Theorem (Doeblin coupling)

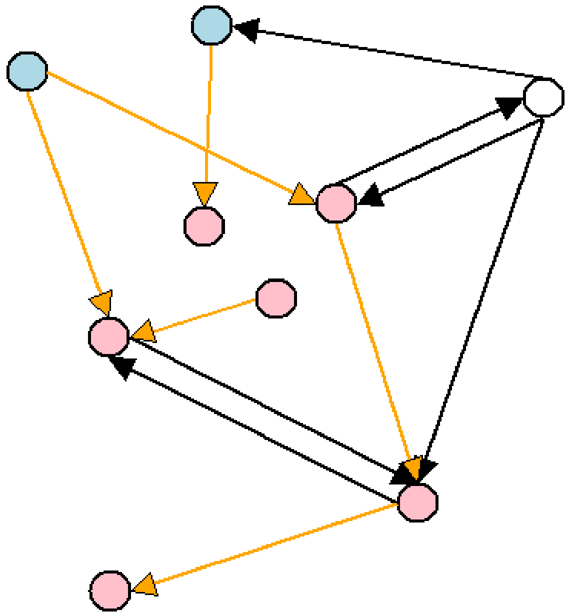
For a finite state space Markov chain, consider two copies; one started in equilibrium ($\mathbb{P}[X = j] = \pi_j$), one at some specified starting state. Run the chains independently till they meet, or couple. Then:

$$\frac{1}{2}\sum_j \left|\pi_j - p_{ij}^{(n)}\right| \quad \leq \quad \mathbb{P}[\text{ no coupling by time } n]$$

We use the "Aldous inequality": suppose X is a Markov chain, with equilibrium distribution π, for which we can produce a coupling between any two points x, y, which succeeds at time $T_{x,y} < \infty$. Then

$$\text{dist}_{tv}\left(\mathcal{L}(X_n), \pi\right) \quad \leq \quad \max_y \left\{\mathbb{P}\left[T_{x,y} > n\right]\right\}.$$

Coupling for rabbits



Heating igloos (I)

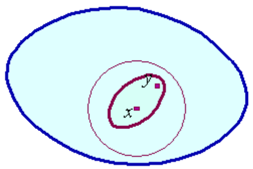
Are larger igloos always colder? (Chavel 1986; WSK 1989; Bass and Burdzy 1993).

Light a match at $t = 0$ at location x. Compare temperatures $p_t(x, y)$, $q_t(x, y)$ at y in well-insulated igloos P and Q, where $P \subseteq Q$. Is it always colder in the larger igloo Q?

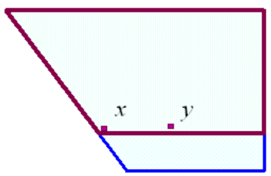
- Yes for large times (convergence to uniformity);
- Not for all times for general igloos;
- Perhaps yes if the igloo is convex?

Heating igloos (II)

Move from Neumann heat kernel to reflecting Brownian motion.



Monotonicity holds if there is a separating circle (WSK 1989):



Monotonicity fails in general:
obtuse angles cause trouble (Bass and Burdzy 1993).

The thematic problem

- Examples indicate a variety of ways in which coupling can be applied.
- Doeblin (1938)'s original coupling is thematic: aim to couple two copies of a process begun at different places, so that they eventually meet.
- Applications: rates of convergence to equilibrium, gradient inequalities, second-eigenvalue estimates.
- The thematic problem has stimulated the theory of probabilistic coupling.
- There is a close relationship with stochastic control.

Diffusions

1. For simplicity, restrict to coupling problems for diffusions (Brownian motion both scalar and vector, Ornstein-Uhlenbeck processes, diffusions with varying coefficients, . . .).
2. Construct two copies X and Y of a diffusion begun at x and y, so as to maximize

$$\mathbb{P}\left[X \text{ and } Y \text{ have met by time } t\right] .$$

3. Can we produce a single construction which maximizes this objective function for all times t?
4. YES: Griffeath (1975), Pitman (1976), Sverchkov and Smirnov (1990), Sverchkov and Smirnov (1990).
5. This maximal coupling is (a) complicated (in general), (b) “anticipating”.

Immersed coupling

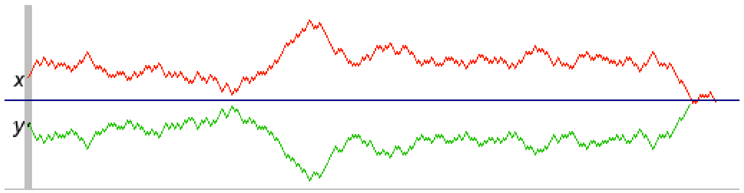
1. Use stochastic calculus to construct co-adapted couplings which do not anticipate each other.
2. If $\mathrm{d}\,X = f(X)\,\mathrm{d}\,B + g(X)\,\mathrm{d}\,t$ then the coupled Y can be given by

$$\mathrm{d}\,Y \quad = \quad f(Y)(J^\top\,\mathrm{d}\,B + K^\top\,\mathrm{d}\,C) + g(Y)\,\mathrm{d}\,t\,,$$

where C is an extra Brownian motion independent of B.
3. We need $J^\top J + K^\top K = \mathbb{I}$. View the infinitesimal correlation J as the control for the problem.
4. Interesting questions:
can we obtain maximal couplings this way?
and if not then can we obtain maximal immersed couplings?

Easy example: scalar Brownian motion

- Reflection Coupling: Make one process meet the other by evolving in a mirror-opposite way! ($J = -1$)



Lindvall (1982)
“On coupling of Brownian motions” (preprint).

- Exceptionally, this coupling is immersed *and* maximal.

Coupling Path Functionals as well

Survey of some known results:

Path and functionals	Couplings	
Brownian motion[1] B	refl	Lindvall (1982)
$B, \int B\,\mathrm{d}\,t$	refl + sync	(Ben Arous et al. 1995)
$B, \int B\,\mathrm{d}\,t, \int\int B\,\mathrm{d}\,s\,\mathrm{d}\,t$	refl + sync	WSK and Price (2004)
$B, \int B\,\mathrm{d}\,t, \ldots, \int \ldots \int B\,\mathrm{d}\,s \ldots \mathrm{d}\,t$	Morse-Thue	WSK and Price (2004)
$BM(\mathbb{R}^2)$, stochastic area[2]	refl + sync	(Ben Arous et al. 1995), WSK (2007)
$BM(\mathbb{R}^n)$, $\binom{n}{2}$ stochastic areas	refl + rotate	WSK (2007)

- WSK (2010) results extend to bounds on speed of coupling for (multiple) stochastic areas.
- Couple all invariant diffusions on nilpotent Lie groups?
  all hypoelliptic diffusions?

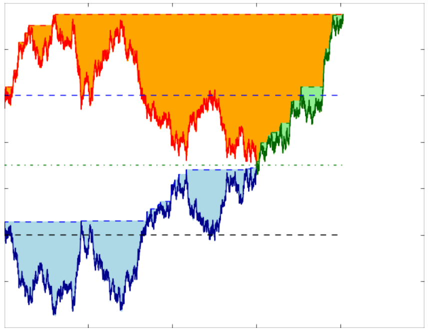
[1] One-dimensional Brownian motion
[2] Stochastic area: $\int B_i\,\mathrm{d}\,B_j - B_j\,\mathrm{d}\,B_i$

Simple example

- Couple Brownian motion X with local time $L^{(0)}$.
- Use Lévy transform to get new Brownian motion B with running supremum S:

$$\begin{aligned} B &= L^{(0)} - |X|\,, \\ S &= L^{(0)}\,. \end{aligned}$$

- Couple 2 copies of (B, S) by reflection / synchronized coupling ($J = -1$, then $J = +1$).



- The resulting coupling is *not* maximal, but *is* maximal amongst all immersed couplings.
- Application to Beneš, Karatzas, and Rishel (1991) diffusion.

Shy coupling

- Can we *avoid* coupling two copies of a random process? ("Shy coupling")
- Consider immersed coupling of reflecting Brownian motion.
- Can't avoid shy coupling in convex domain (Benjamini, Burdzy, and Chen 2007; WSK 2009);

▸ SHY ANIMATION (1)

- Can avoid shy coupling in annulus;

▸ SHY ANIMATIONv (2)

and, amazingly,

- can't avoid shy coupling in simply connected domain (Bramson, Burdzy, and WSK 2011,2012):

▸ SHY ANIMATION (3)

Questions

- Applications for shy coupling?
- Brownian reflection coupling lies at the heart of many coupling algorithms, because it is simple and amenable to exact calculation. Can reflection / synchronized coupling play a similar rôle?
- Link immersion ("co-adapted") coupling to notion of greedy algorithms?
- There are instances where mildly non-co-adapted couplings do substantially better (Gibbs sampler in simplex, per Smith 2011; perfect simulation for super-stable $M/G/c$ queue, per Sigman 2011). How to understand this?

Bass, R. F. and K. Burdzy (1993).
On domain monotonicity of the Neumann heat kernel.
J. Funct. Anal 116, 215–224.

Ben Arous, G., M. Cranston, and WSK (1995).
Coupling constructions for hypoelliptic diffusions: Two examples.
In M. Cranston and M. Pinsky (Eds.), *Proceedings of Symposia in Pure Mathematics*, Volume 57, Providence, RI Providence, pp. 193–212. American Mathematical Society.

Beneš, V. E., I. Karatzas, and R. W. Rishel (1991).
The separation principle for a Bayesian adaptive control problem with no strict-sense optimal law.
In *Applied Stochastic Analysis*, Stochastic Monographs, pp. 121–156. New York: Gordon & Breach.

Benjamini, I., K. Burdzy, and Z.-Q. Chen (2007, March).
Shy couplings.
Probability Theory and Related Fields 137(3-4), 345–377.

Bramson, M., K. Burdzy, and WSK (2011).
Rubber Bands, Pursuit Games and Shy Couplings.

Bramson, M., K. Burdzy, and WSK (2012, July).
Shy Couplings, CAT(0) Spaces, and the Lion and Man.
Annals of Probability, To appear.

Chavel, I. (1986).
Heat diffusion in insulated convex domains.
The Journal of the London Mathematical Society (Second Series) (2) 34(3), 473–478.

Doeblin, W. (1938).
Exposé de la Théorie des Chaînes simples constants de Markoff á un nombre fini d'États.
Revue Math. de l'Union Interbalkanique 2, 77–105.

Griffeath, D. (1975).
A maximal coupling for Markov chains.
Zeitschrift für Wahrscheinlichkeitstheorie und Verwe Gebiete 31, 95–106.

Lindvall, T. (1982).
On Coupling of Brownian Motions.
Technical report 1982:23, Department of Mathematics, Chalmers University of Technology and University of Göteborg.

Pitman, J. W. (1976).
On coupling of Markov chains.
Zeitschrift für Wahrscheinlichkeitstheorie und Verwe Gebiete 35(4), 315–322.

Siegmund, D. (1976).
The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes.
The Annals of Probability 4(6), 914–924.

Sigman, K. (2011).
Exact Simulation of the Stationary Distribution of the FIFO M/G/c Queue.
Journal of Applied Probability 48A, 209–213.

Smith, A. (2011, July).
A Gibbs Sampler on the n-Simplex.
arxiv preprint arxiv:1107, 16.

Sverchkov, M. Y. and S. N. Smirnov (1990).
Maximal coupling for processes in $D[0,\infty]$.
Dokl. Akad. Nauk SSSR 311(5), 1059–1061.

WSK (1989).
Coupled Brownian motions and partial domain monotonicity for the Neumann heat kernel.
Journal of Functional Analysis 86, 226–236.

WSK (2007, May).
Coupling all the Lévy stochastic areas of multidimensional Brownian motion.
The Annals of Probability 35(3), 935–953.

WSK (2009, September).
Brownian couplings, convexity, and shy-ness.
Electronic Communications in Probability 14(Paper 7), 66–80.

WSK (2010).
Coupling time distribution asymptotics for some couplings of the Lévy stochastic area.
In N. H. Bingham and C. M. Goldie (Eds.), *Probability and Mathematical Genetics: Papers in Honour of Sir John Kingman*, Chapter 19, pp. 446–463. Cambridge: Cambridge University Press.

WSK and C. J. Price (2004).
Coupling iterated Kolmogorov diffusions.
Electronic Journal of Probability 9(Paper 13), 382–410.

WSK and I. W. Saunders (1983).
Epidemics in competition II: The general epidemic.
Journal of the Royal Statistical Society (Series B: Methodological) 45, 238–244.