Institute for Applied Mathematics
Winter term 2016/17
Andreas Eberle, Raphael Zimmer

“Markov Processes”, Problem Sheet 10.

Hand in solutions before Monday 9.1., 2 pm

We wish you a merry Christmas and a happy new year!

1. (Gibbs sampler for the Ising model). Consider a finite graph (V, E) with n vertices of maximal degree Δ. The Ising model with inverse temperature $\beta \geq 0$ is the probability measure μ_β on $\{-1, 1\}^V$ with mass function

$$\mu_\beta(\eta) \;=\; \frac{1}{Z(\beta)} \exp\left(\beta \sum_{\{x,y\}\in E} \eta(x)\eta(y)\right),$$

where $Z(\beta)$ is a normalization constant.

a) Show that given $\eta(y)$ for $y \neq x$, $\eta(x) = \pm 1$ with probability $(1 \pm \tanh(\beta m(x, \eta))/2$, where $m(x, \eta) := \sum_{y\sim x} \eta(y)$ is the local magnetization in the neighbourhood of x. Hence determine the transition kernel π for the Gibbs sampler with equilibrium μ_β.

b) Prove that for any $t \in \mathbb{N}$,

$$\mathcal{W}^1(\nu\pi^t, \mu_\beta) \;\leq\; \alpha(n, \beta, \Delta)^t\, \mathcal{W}^1(\nu, \mu_\beta) \;\leq\; \exp\left(-\frac{t}{n}(1 - \Delta \tanh(\beta))\right) \mathcal{W}^1(\nu, \mu_\beta),$$

where $\alpha(n, \beta, \Delta) = 1-(1-\Delta \tanh(\beta))/n$, and $\mathcal{W}^1$ is the transportation metric based on the Hamming distance on $\{-1, 1\}^V$. Conclude that for $\Delta \tanh \beta < 1$, the Gibbs sampler is geometrically ergodic with a rate of order $\Omega(1/n)$. *Hint: You may use the inequality*

$$|\tanh(y + \beta) - \tanh(y - \beta)| \;\leq\; 2 \tanh(\beta) \qquad \text{for any } \beta \geq 0 \text{ and } y \in \mathbb{R}.$$

c) The *mean-field Ising model* with parameter $\alpha \geq 0$ is the Ising model on the complete graph over $V = \{1, \dots, n\}$ with inverse temperature $\beta = \alpha/n$. Show that for $\alpha < 1$, the ϵ-mixing time for the Gibbs sampler on the mean field Ising model is of order $O(n \log n)$ for any $\epsilon \in (0, 1)$.

2. (Successful couplings and TV-convergence to equilibrium). Consider a time-homogeneous Markov process with transition function (p_t) and invariant probability measure μ. A coupling (X_t, Y_t) of versions of the process with initial distributions ν and μ, respectively, is called *successful* if the coupling time

$$T = \inf\{t \geq 0 : X_s = Y_s \text{ for any } s \geq t\}$$

is almost surely finite. Show that a successful coupling exists if and only if $||\nu p_t - \mu||_{TV} \to 0$ as $t \to \infty$.

3. (Bounds for ergodic averages in the non-stationary case). Let (X_n, P_x) be a Markov chain with transition kernel π and invariant probability measure μ, and let

$$A_{b,n}f = \frac{1}{n}\sum_{i=b}^{b+n-1} f(X_i).$$

Assume that there are a distance d on the state space S, and constants $\alpha \in (0,1)$ and $\bar{\sigma} \in \mathbb{R}_+$ such that

(A1) $\mathcal{W}_d^1(\nu\pi, \widetilde{\nu}\pi) \leq \alpha\, \mathcal{W}_d^1(\nu, \widetilde{\nu}) \;\; \forall\, \nu, \widetilde{\nu} \in \mathcal{P}(S)$, and

(A2) $\mathrm{Var}_{\pi(x,\cdot)}(f) \leq \bar{\sigma}^2\, \|f\|^2_{Lip(d)} \;\; \forall\, x \in S,\; f : S \to \mathbb{R}$ Lipschitz.

Prove that under these assumptions the following bounds hold for any $b, n, k \geq 0$, $x \in S$, and for any Lipschitz continuous function $f : S \to \mathbb{R}$:

a) $\mathrm{Var}_x\,[f(X_n)] \;\leq\; \sum_{k=0}^{n-1} \alpha^{2k}\, \bar{\sigma}^2 \|f\|^2_{Lip(d)}$.

b) $|\mathrm{Cov}_x\,[f(X_n), f(X_{n+k})]| \;=\; \left|\mathrm{Cov}_x\left[f(X_n), (\pi^k f)(X_n)\right]\right| \;\leq\; \frac{\alpha^k}{1-\alpha^2}\bar{\sigma}^2 \|f\|^2_{Lip(d)}$.

c) $\mathrm{Var}_x\,[A_{b,n}f] \;\leq\; \frac{1}{n}\frac{\bar{\sigma}^2}{(1-\alpha)^2}\|f\|^2_{Lip(d)}$.

d) $|E_x\,[A_{b,n}f] - \mu(f)| \;\leq\; \frac{1}{n}\frac{\alpha^b}{1-\alpha}\int d(x,y)\,\mu(dy)\,\|f\|_{Lip(d)}$.

e) $E_x\left[|A_{b,n}f - \mu(f)|^2\right] \;\leq\; \frac{1}{n}\frac{1}{(1-\alpha)^2}\left(\bar{\sigma}^2 + \frac{1}{n}\alpha^{2b}(\int d(x,y)\,\mu(dy))^2\right)\|f\|^2_{Lip(d)}$.

4. (Equivalent descriptions for weighted total variation norms). Let $V : S \to (0,\infty)$ be a measurable function, and let $d_V(x,y) := (V(x)+V(y))1_{x\neq y}$. Show that the following identities hold for probability measures μ, ν on $(S, \mathcal{B})$:

$$\begin{aligned}
\|\mu - \nu\|_V &= \left\|\frac{d\mu}{d\lambda} - \frac{d\nu}{d\lambda}\right\|_{L^1(V\cdot\lambda)} \\
&= \sup\{|\mu(f) - \nu(f)| : f \in \mathcal{F}(S) \text{ s.t. } |f| \leq V\} \\
&= \sup\{|\mu(f) - \nu(f)| : f \in \mathcal{F}(S) \text{ s.t. } |f(x) - f(y)| \leq d_V(x,y)\; \forall\, x, y\} \\
&= \inf\{E[d_V(X,Y)] : X \sim \mu, Y \sim \nu\}
\end{aligned}$$