Recent results on hybrid sampling

Gersende Fort *Grenoble, France*

Eric Moulines *Paris, France*

Gareth Roberts *Lancaster, UK*

Jeffrey Rosenthal *Toronto, Canada*

Workshop: MCMC methodology

December 2001, Lancaster

http://www-lmc.imag.fr/lmc-sms/Gersende.Fort

(A-1) Problem

▶ We address a MCMC sampler

- for $\mathbb{R}^d$-valued Markov Chain,
- with known invariant distribution $\pi(dx) =: p(x)\, \mu_d(dx)$. (*target distribution w.r.t. Lebesgue*)

▶ Random Scan Symmetric Random Walk Hastings-Metropolis algorithm henceforth: RSM

$\hookrightarrow$ performs a Metropolis step on just one coordinate at a time.

▶ We propose sufficient conditions for V-geometric ergodicity of the RSM

(A-2) The RSM

$\mathbf{x} = (x_1, \cdots, x_d) \in \mathbb{R}^d$, orthonormal basis $(e_1, \cdots, e_d)$

Given d symmetric proposal distributions on $\mathbb{R}$: $\quad q_1, \cdots, q_d$

Define a RWHM on $\mathbb{R}^d$ acting on direction e_i:

$$P_i\left(\mathbf{x}, A_1 \times \cdots \times A_d\right) := \prod_{k \neq i} \delta_{x_k}(A_k) \left\{ \int_{A_i - x_i} \alpha(x, x + ye_i)\, q_i(y)\mu(dy) \right.$$
$$\left. + \ \delta_{x_i}(A_i) \int \Big(1 - \alpha(x, x + ye_i)\Big)\, q_i(y)\mu(dy) \right\},$$

where $\alpha(z, x) := 1 \wedge [p(z)/p(x)], \qquad (x, z) \in \mathbb{R}^d \times \mathbb{R}^d$

▶ RSM kernel $\qquad P_{\mathrm{RS}} := \dfrac{1}{d} \displaystyle\sum_{i=1}^{d} P_i.$

(A-3) V-geometric ergodicity

Assume

- P is ψ-irreducible, aperiodic,
- there exist a function $1 \leq V < \infty$, a petite set C, some constants $0 < \lambda < 1$, $b < \infty$ s.t.

$$PV \leq \lambda V + b \mathbb{1}_C$$

Then

- there exists an invariant probability measure π and $\pi(V) < \infty$.
- there exist $r > 1$ and $R < \infty$ s.t.

$$\sum_n r^n \, \|P^n(x,\cdot) - \pi(\cdot)\|_V \leq R\, V(x) \qquad \forall x.$$

which implies convergence in Total Variation norm $\|P^n(x,\cdot) - \pi(\cdot)\|_{TV} \leq r^{-n}\, R\, V(x)$

and $|P^n g(x) - \pi(g)| \leq r^{-n}\, R\, V(x)$ for all $|g| \leq V$,

and CLT $\cdots$

(A-4) V-geometric ergodicity of RSM

[Roberts and Rosenthal 1997,1998]

(Main) Assumptions

- Target density p continuously differentiable on $\mathbb{R}^d$.
- Asymptotically exponentially decreasing tails:

$$\lim_{|x|\to\infty} |\nabla \log p(x)| > 0$$

Ex: $p(x) \sim \exp(-|x|^\alpha)$ with $\alpha \geq 1$ Exponential ($\alpha = 1$) and sub-exponential ($\alpha > 1$) target densities

- $\kappa(x)$: The maximum curvature of all geodesic curves through the surface $\{y, p(y) = p(x)\}$ at point x

$$\lim_{|x|\to\infty} \kappa(x) = 0.$$

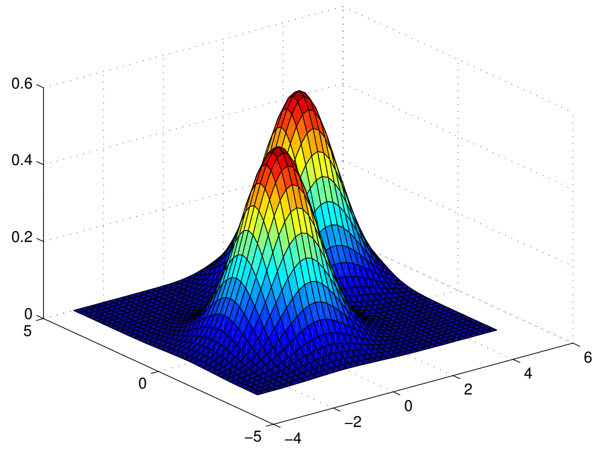
Ex:

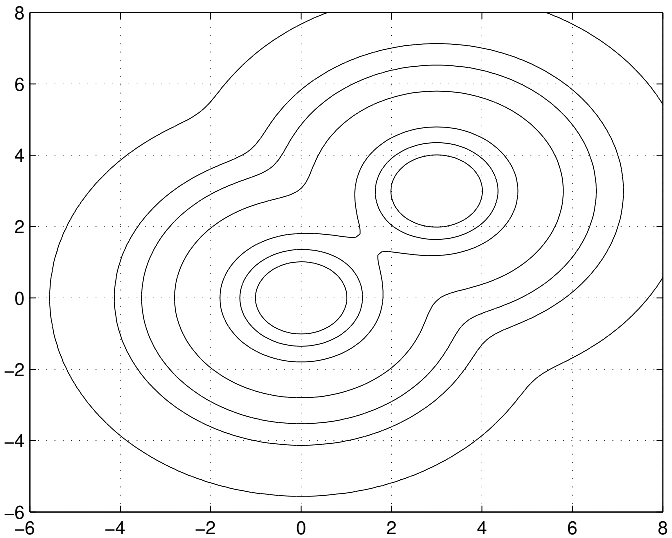
$\{(y_1, y_2), p(y) = p(x)\} = \mathcal{C}(0, |x|)$ $\kappa(x) = 1/|x|$.

$\{(y_1, y_2), p(y) = p(x)\} = \{(t_1, t_2), t_2 = h(t_1)\}$ and $\kappa(x) = |h''(x)| \, (1 + h^{'2}(x))^{-3/2}$

(A-5-a) Example on $\mathbb{R}^2$

$$p(x) = 0.5\exp(-(x_1^2 + x_2^2)/2) + 0.5\exp(-(x_1^2 + x_2^2)/2)$$



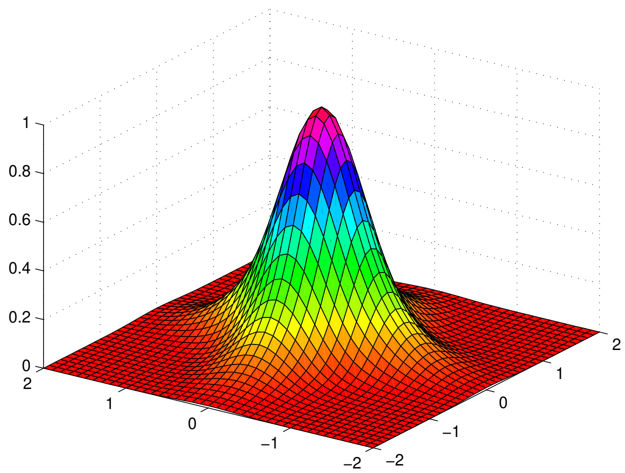


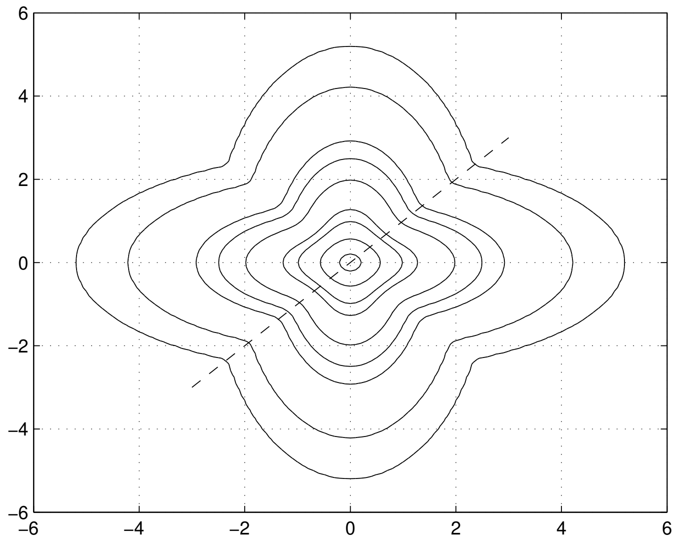
For all x, $\kappa(x) = 1/|x| \longrightarrow 0$

$\hookrightarrow$ RSM is V-geometrically ergodic.

(A-5-b) Example on $\mathbb{R}^2$

$$p(x) = 0.5\exp(-(x_1^2 + 4x_2^2)) + 0.5\exp(-(4x_1^2 + x_2^2))$$





In the first quadrant, on the diagonal (t,t)

$$\kappa((t,t)) = 9\sqrt{2}/5\,t - 1/(\sqrt{2}t) \longrightarrow +\infty$$

$\hookrightarrow$ the results do not apply.

(A-6) Contribution

- V-geometric ergodicity of RSM
  - ψ-irreducibility, aperiodicity, small set
  - A Foster-Lyapunov drift condition for sub-exponential target density
  - A Foster-Lyapunov drift condition for exponential target density

- A necessary condition for V-geometric ergodicity of RSM

- Examples:
  - RSM is V-geometrically ergodic when the full HM is not.
  - RSM can not be V-geometrically ergodic and our results do not apply $\cdots$

(B-1) Assumptions

A1 the target density p is continuous on $\mathbb{R}^d$.

A2 symmetric proposal distributions $q_1, \cdots, q_d$ on $\mathbb{R}$, and

$$|y| \leq \delta_i \Longrightarrow q_i(y) \geq \epsilon_i \qquad \text{for all } i \in \{1, \cdots, d\}.$$

Result *[Roberts and Rosenthal 1998]*

- P_{RS} is Lebesgue-irreducible, aperiodic,
- with invariant distribution probability measure $\pi(dx) := p(x)\mu_d(dx)$
- any bounded set is small

(B-2) A key property *[Roberts and Rosenthal 1998]*

$P_{\mathrm{RS}} := d^{-1} \sum_{i=1}^{d} P_i$

Define: $V_s(x) := p(x)^{-s}, \quad 0 < s < 1.$

Result For all $i \in \{1, \cdots, d\}, \ \ x \in \mathbb{R}^d$,

$P_i V_s(x) \leq r(s) V_s(x)$ $\qquad r(s) > 1$ $\qquad r(s) \downarrow 1$ as $s \to 0$

Idea Show that: $\exists\, \zeta < 1, \ \forall\, \{x^j\}, |x^j| \to \infty, \ \exists\, i \ \forall\, s, \ \limsup_j P_i V_s(x^j)/V_s(x^j) < \zeta$

Let s $\quad (d-1) r(s) + \zeta < d$.

If no $\quad$ "$\limsup_{|x| \to \infty} P_{\mathrm{RS}} V_s(x)/V_s(x) < 1$"

Then $\quad \exists \{x^j\}, |x^j| \to \infty, \limsup_j P_{\mathrm{RS}} V_s(x^j)/V_s(x^j) \geq 1$

But $\quad P_{\mathrm{RS}} V_s(x^j)/V_s(x^j) = d^{-1} \sum_{i=1}^{d} P_i V_s(x^j)/V_s(x^j) \leq d^{-1}\left((d-1) r(s) + \zeta\right) < 1$

Set: : $\mathcal{A}(x,i) - x = \{y, p(x+ye_i) \geq p(x)\}$ $\mathcal{R}(x,i) - x = \{y, p(x+ye_i) \leq p(x)\}$

$$\frac{P_i V_s(x)}{V_s(x)} = \int_{\mathcal{A}(x,i)-x} \left[\frac{p(x+ye_i)}{p(x)}\right]^{-s} q_i(y)\mu(dy)$$
$$+\int_{\mathcal{R}(x,i)-x} \left\{\left[\frac{p(x+ye_i)}{p(x)}\right]^{1-s} - \frac{p(x+ye_i)}{p(x)}\right\} q_i(y)\mu(dy) + \int_{\mathcal{R}(x,i)-x} q_i(y)\mu(dy)$$

Ex. $|x^j| \to \infty$, and $x_i^j \to +\infty$ $\mathcal{A}(x^j,i) - x^j \to \mathbb{R}_-$ $\mathcal{R}(x^j,i) - x^j \to \mathbb{R}_+$

$$\lim_j \frac{P_i V_s(x^j)}{V_s(x^j)} = \int_{\mathbb{R}_-} \lim_j \left[\frac{p(x^j)}{p(x^j+ye_i)}\right]^{s} q_i(y)\mu(dy)$$
$$+\int_{\mathbb{R}_+} \lim_j \left\{\left[\frac{p(x^j+ye_i)}{p(x^j)}\right]^{1-s} - \frac{p(x^j+ye_i)}{p(x^j)}\right\} q_i(y)\mu(dy) + \int_{\mathbb{R}_+} q_i(y)\mu(dy) \longrightarrow 1/2$$

- $\forall\, y > 0, \qquad \lim_j \frac{p(x^j)}{p(x^j - ye_i)} = 0 \qquad \lim_j \frac{p(x^j+ye_i)}{p(x^j)} = 0$

(B-3-a) RSM for sub-exponential density

A3 For some $0 \leq \delta < \Delta \leq +\infty$, $\xi := \inf_{1 \leq i \leq d} \int_\delta^\Delta q_i(y)\mu(dy) > 0$

For any sequence $\{x^j\}$ with $\lim_j |x^j| = +\infty$, one may extract a subsequence $\{\tilde{x}^j\}$, s.t. $\exists\, i \in \{1, \ldots, d\}, \forall\, y \in [\delta, \Delta]$

$$\lim_j \frac{p(\tilde{x}^j)}{p(\tilde{x}^j - \mathrm{sign}(\tilde{x}_i^j)\, y\, e_i)} = 0, \quad \text{and} \quad \lim_j \frac{p(\tilde{x}^j + \mathrm{sign}(\tilde{x}_i^j)\, y\, e_i)}{p(\tilde{x}^j)} = 0.$$

Theorem Let $0 < s < 1$ s.t. $r(s) < 1 + \xi/(d - 2\xi)$.

There exist $0 < \lambda < 1, b < \infty, C$ compact

$$P_{\mathrm{RS}} V_s(x) \leq \lambda V_s(x) + b \mathbb{1}_C(x) \qquad V_s(x) := p^{-s}(x).$$

$\hookrightarrow$ V-geometric ergodicity of the RSM

(B-3-b) A sufficient condition for A3

- p is continuously differentiable.
- There exists $\Delta > 0$ s.t. $\forall i \in \{1, \cdots, d\}, 0 < y < \Delta, \epsilon \in \{-1, 1\}$,

$$\lim_{x_i \to \epsilon\infty} \sup_{\{x_{-i}: |x_j| \leq |x_i|, j \neq i\}} \sup_{\{t;\, |t| \leq y\}} \epsilon\, \nabla_i \log p(x + te_i) = -\infty.$$

OR for some functions ϕ_k s.t. $\lim_n \phi(t_n) = +\infty$ if and only if $\lim_n t_n = +\infty$,

$$\lim_{x_i \to \epsilon\infty} \sup_{\{x_{-i}: \phi_j(|x_j|) \leq \phi_i(|x_i|), j \neq i\}} \sup_{\{t;\, |t| \leq y\}} \epsilon\, \nabla_i \log p(x + te_i) = -\infty.$$

Ex: On $\mathbb{R}^d$, $\log p(x) \propto -|x|^l$

if $l > 1$, A3 applies — A3 covers sub-exponential density

if $l = 1$, the ratios tend to $\exp(\alpha y)$ — A3 is not for exponential density.

(B-4) RSM for exponential density

A3' There is $\alpha > 0$ and $1/\alpha \leq \delta < \Delta \leq \infty$ s.t. for any sequence $\{x^j\}$ with $\lim_j |x^j| = +\infty$, one may extract a subsequence $\{\tilde{x}^j\}$, s.t. for some $i \in \{1, \cdots, d\}$, we have, for all $y \in [\delta, \Delta]$,

$$\lim_j \frac{p(\tilde{x}^j)}{p(\tilde{x}^j - \mathrm{sign}(x_i^j)\ y\ e_i)} \leq \exp(-\alpha y) \quad \text{and} \quad \lim_j \frac{p(\tilde{x}^j + \mathrm{sign}(\tilde{x}_i^j)\ y\ e_i)}{p(\tilde{x}^j)} \leq \exp(-\alpha y).$$

A4 $\inf_{i \in \{1,\ldots,d\}} \int_\delta^\Delta\ y\ q_i(y) \mu(dy) \geq d\ \alpha^{-1}\ (e-1)^{-1}$.

Theorem For some $0 < s < 1$, there exist $0 < \lambda < 1, b < \infty, C$ compact

$$P_{\mathrm{RS}} V_s(x) \leq \lambda V_s(x) + b \mathbb{1}_C(x) \qquad V_s(x) := p^{-s}(x).$$

$\hookrightarrow$ V-geometric ergodicity of RSM

(B-5) A necessary condition for V-geometric ergodicity

Remind:

- RW Hastings-Metropolis on $\mathbb{R}$: *[Mengersen, Tweedie 1996]*

$$\int |y| \, q(y) \mu(dy) < \infty \qquad \hookrightarrow \exists s > 0 \quad \int_{\mathbb{R}} \exp(s|x|) \, p(x) \, \mu(dx) < \infty.$$

- RW Hastings-Metropolis on $\mathbb{R}^d$: *[Jarner, Hansen 2000]*

$$\int |y| \, q(y) \mu_d(dy) < \infty \qquad \hookrightarrow \exists s > 0 \quad \int_{\mathbb{R}^d} \exp(s|x|) \, p(x) \, \mu_d(dx) < \infty.$$

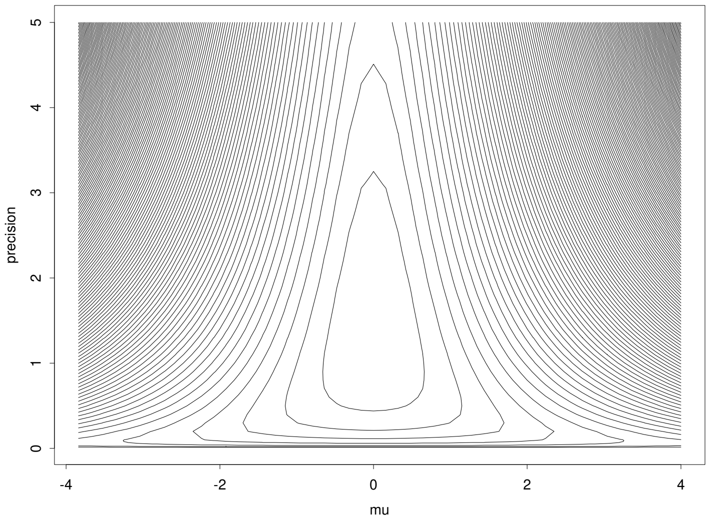
Result:

- RSM on $\mathbb{R}^d$:

$$\int |y| \, q_i(y) \mu(dy) < \infty \qquad \hookrightarrow \exists s_i > 0 \quad \int_{\mathbb{R}^d} \exp\left(\sum_{i=1}^{d} s_i |x_i| \right) \, p(x) \, \mu_d(dx) < \infty.$$

(C-1-a) Example: Normal-Inverse Gamma model

$p((\mu,\tau)) \propto \tau^{(n+1)/2} \prod_{i=1}^{n} \exp(-\tau(y_i - n)^2/2)$ on $\mathbb{R} \times \mathbb{R}_+$



Posterior distribution in a two dimensional Bayesian analyses of an i.i.d. Gaussian model

Data $y_k \sim \mathcal{N}(\mu, \tau^{-1})$

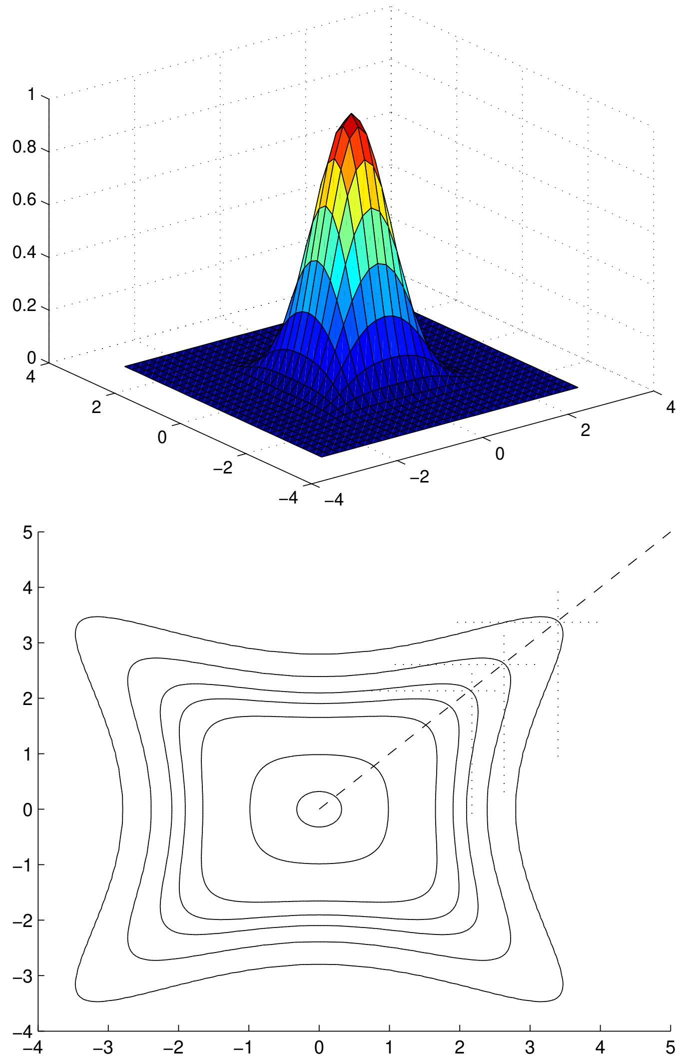
Prior: flat on μ (constant density), and $\tau^{-1/2}$

- RSM is not V-geometrically-ergodic

  μ has heavy tailed distribution $\sim \mu^{-(n+3)}$

  necessary condition unverified or capacitance argument

- our conditions do not apply



(C-1-b) Example: sub-geometrical density

$$p(x) \propto \exp\left(-(x_1^2 + (x_1^2 - x_2^2)^2/4 + x_2^2)\right)$$

- RSM is not V-geometrically ergodic

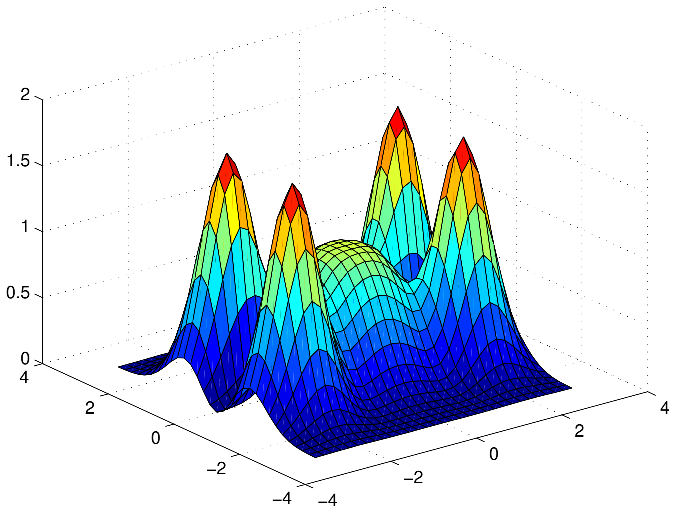
  Ess sup argument

- our conditions do not apply

  Along the sequence $x^j = (j, j)$,

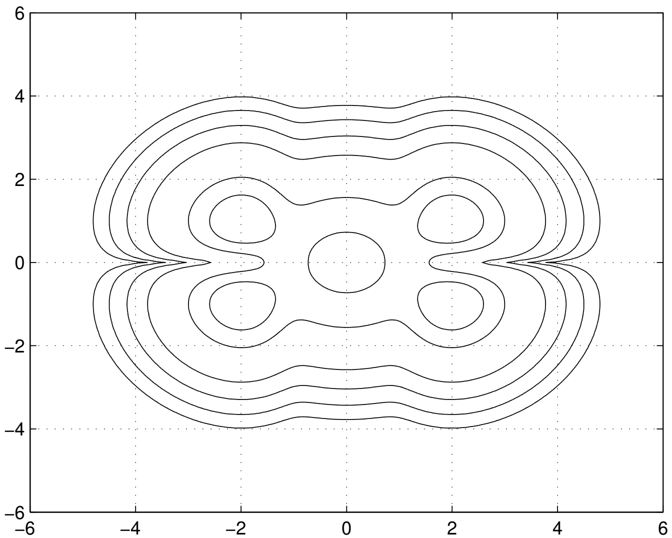
  $P_i V_s / V_s$ goes to 1,

  whatever i



(C-2-a) Example: sub-geometrical density

$$p(x) \propto \left(1 + x_1^2 + x_2^2 + x_1^8 x_2^2\right) \exp\left(-(x_1^2 + x_2^2)\right)$$



- The curvature condition is not verified

  Along the x-axis,

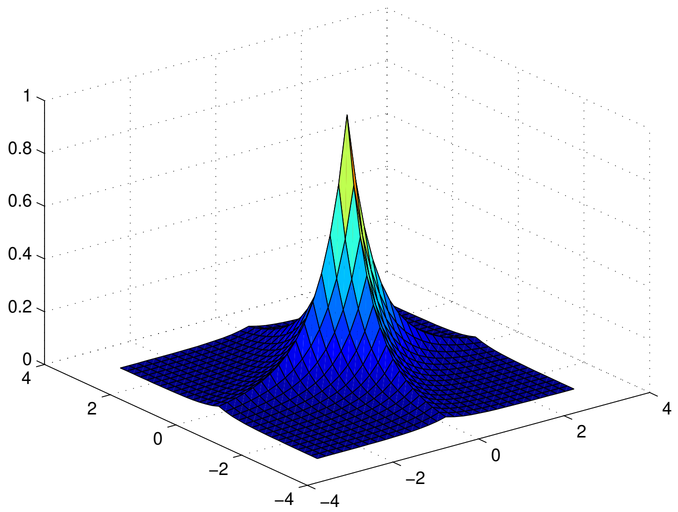
  $\kappa(x) \to \infty$

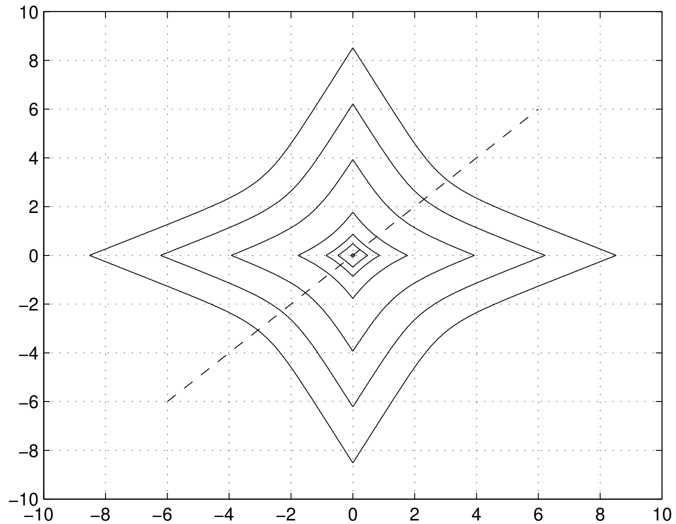
- Our conditions appply

$\hookrightarrow$ RSM is V-geometrically ergodic

(C-2-b) Example: geometrical density

$$p(x) \propto 0.5\ \exp\left(-\left(|x_1| + 2|x_2|\right)\right) + 0.5\ \exp\left(-\left(2|x_1| + |x_2|\right)\right)$$





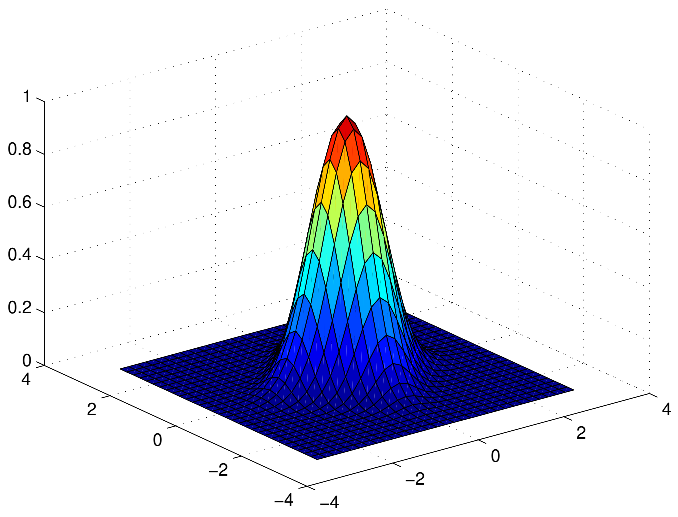
- The curvature condition is not verified

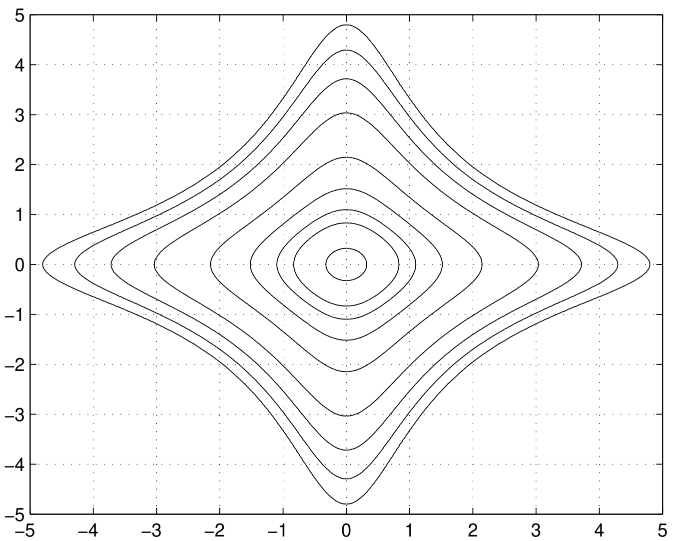
  Along the diagonal $x = (t, t)$

  $\kappa(x) = 1/(3\sqrt{2}) > 0$

- Our conditions appply

$\hookrightarrow$ RSM is V-geometrically ergodic





(C-3-a) Example: sub-geometrical density

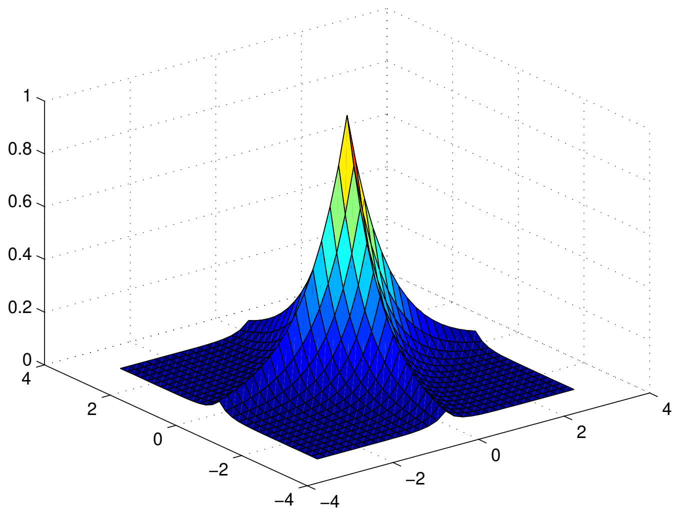
$$p(x) \propto \exp\left(-(x_1^2 + x_1^2 x_2^2 + x_2^2)\right)$$

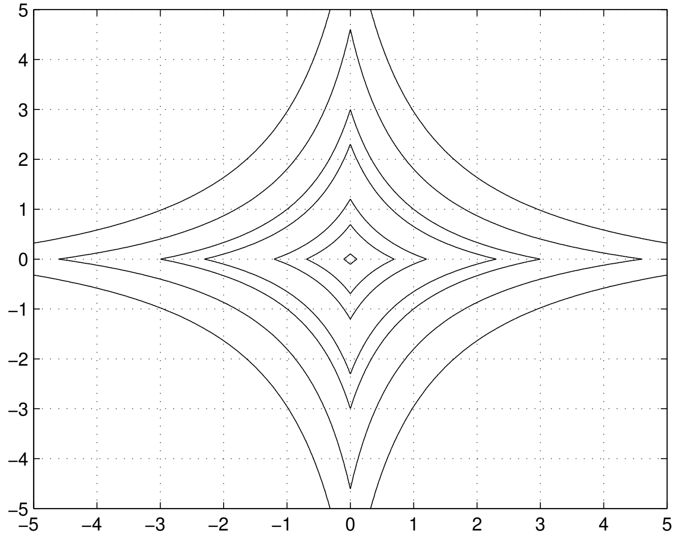
- The (full) RW is not geometrically ergodic

  [Jarner, Hansen 2000]

- Our conditions appply

$\hookrightarrow$ RSM is V-geometrically ergodic in situations where the full-dimensional Metropolis algorithm is not geometrically ergodic





(C-3-b) Example: geometrical density

$$p(x) \propto \exp\left(-\left(|x_1| + |x_1||x_2| + |x_2|\right)\right)$$

- The (full) RW is not geometrically ergodic

  ess sup argument

- Our conditions appply

$\hookrightarrow$ RSM is V-geometrically ergodic in situations where the full-dimensional Metropolis algorithm is not geometrically ergodic