Magnolia, Magnolia, Magnolia L.

HCMC
Turning an algorithm into a dynamical system

Soumyadib Ghosh   Yingdong Lu   Tomasz Nowicki

IBM Research
tnowicki@us.ibm.com

MM 70



Table of contents

MCMC

- Markov Chain Monte Carlo methods create samples with *probability density proportional to a known function*.
- These samples can be used to evaluate an integral over that variable, as its expected value or variance.
- MCMC address multi-dimensional problems better than simple Monte Carlo algorithms.
- Metropolis-Hastings algorithm.

https://en.wikipedia.org/wiki/Markov_chain_Monte_Carlo

(H)MC

Goal: Find the density proportional to $\mathfrak{f} : \mathbb{Q} \to \mathbb{R}$.

double the dimension of $\mathbb{Q}$ introducing an (isomorphic) $\mathbb{P}$,

chose a density $\mathfrak{g}$ acting on $\mathbb{P}$, can be $\mathfrak{g}(p)$ or $\mathfrak{g}(p|q)$,

Choose an initial density. Procedure, one step:

from a density $h(q)$ take a sample q,

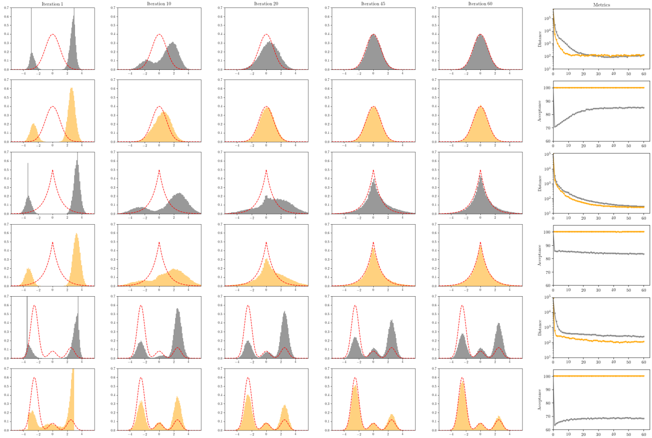
from $\mathfrak{g}(p)$ take a sample p,

move the (sampled) point (q, p) to (Q, P)
according to a motion $H : \mathbb{Q} \times \mathbb{P} \circlearrowleft$, depending on $\mathfrak{f}$ and $\mathfrak{g}$,

collect Q, this will represent a sample from $\hat{h} = \mathcal{T}h$.

Rinse and repeat until satisfied.

HMC vs MC



Plots of iterate distributions for the MHMC method (grey) and HMC method (orange). Top two rows target $\mathfrak{f}$ is normal, the next two (symmetric) exponential and last two are a multimodal

(H)MC, abstracted

Given $\mathfrak{f}$ choose $\mathfrak{g}$ and H with **invariance properties** and define an action $\mathcal{T}$ on the densities h:

Spread: $\mathfrak{h}(q,p) = h(q)\cdot\mathfrak{g}(p)$ on $\mathbb{Q}\times\mathbb{P}$,

Move: $\hat{\mathfrak{h}}(q,p) = (\mathfrak{h}\circ H)(q,p) = \mathfrak{h}(Q,P)$ on $\mathbb{Q}\times\mathbb{P}$,

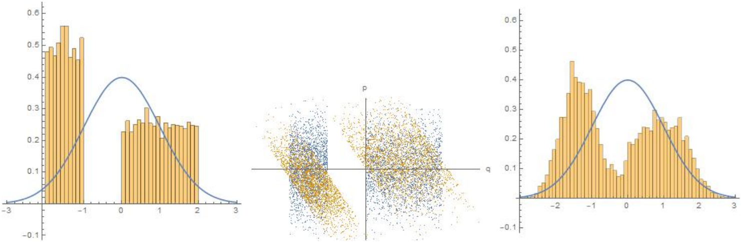
Collect: $\mathcal{T}h(q) = \hat{h}(q) = \int_{\mathbb{P}} \hat{\mathfrak{h}}(q,p)\mathfrak{g}(P(q,p))$

Iterate $\mathcal{T}$ until satisfied.

Invariance properties:

$\mathfrak{f}\cdot\mathfrak{g} = (\mathfrak{f}\cdot\mathfrak{g})\circ H$
$\iint_{\mathbb{Q}\times\mathbb{P}} A\circ H = \iint_{\mathbb{Q}\times\mathbb{P}} A$.

HMC, how it works



(H)MC, the definition

Let $\mathfrak{f} : \mathbb{Q} \to \mathbb{R}$, $\mathfrak{g} : \mathbb{P} \to \mathbb{R}$ and
$H : \mathbb{Q} \times \mathbb{P} \to \mathbb{Q} \times \mathbb{P}, \qquad H(q, p) = (Q, P)$ satisfy:

Invariance properties:

$\mathfrak{f} \cdot \mathfrak{g} = (\mathfrak{f} \cdot \mathfrak{g}) \circ H$
$\iint_{\mathbb{Q}\times\mathbb{P}} A \circ H = \iint_{\mathbb{Q}\times\mathbb{P}} A.$

Define a function $\mathcal{T}h : \mathbb{Q} \to \mathbb{R}$:

(H)MC

$$\begin{aligned} \mathcal{T}h &= \int_{\mathbb{P}} (h \cdot \mathfrak{g}) \circ H \\ (\mathcal{T}h)(q) &= \int_{\mathbb{P}} (h \cdot \mathfrak{g}) \circ H(q, p)\, dp = \int_{\mathbb{P}} h(Q(q, p)) \cdot \mathfrak{g}(P(q, p))\, dp\,. \end{aligned}$$

An additional assumption

Coverage condition

For any q the function $p \mapsto Q(q,p) = \mathcal{Q}_q(p)$ maps $\mathbb{P}$ onto $\mathbb{Q}$

This condition can be weakened.

It corresponds to the irreducibility or ergodicity (of sorts) of the movement.

Hamiltionian MC

set $\mathcal{U}(q) = -\log \mathfrak{f}(q)$, $\mathcal{V}(p) = -\log \mathfrak{g}(p)$, $\mathcal{H} = \mathcal{U} + \mathcal{V}$

$$\dot{Q} = \frac{\partial Q}{\partial t} = \quad \frac{\partial \mathcal{H}}{\partial p}(Q, P), \quad \dot{P} = \frac{\partial P}{\partial t} = -\frac{\partial \mathcal{H}}{\partial q}(Q, P).$$

pick t, set $H(q, p) = (Q_t, P_t)$ a solution after time t starting at (q, p).

Due to the properites of the Hamiltionian motion

Both **invariance properties** are satisfied,
$\mathfrak{f} \cdot \mathfrak{g}$ equals $\exp(-\mathcal{H})$ up to a constant
Hamiltonian flows preserve the Lebesgue measure.
The normalizing constant of $\mathfrak{f}$ is irrelevant for H.

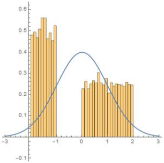
Example, $\mathcal{T}h = \int_{\mathbb{P}} (h \cdot \mathfrak{g}) \circ H, \quad (Q, P) = H(q, p)$

$$\mathfrak{f}(q) = \exp(-q^2/2),\ \mathfrak{g}(p) = \exp(-p^2/2)/\sqrt{2\pi}.$$
$$\mathcal{H}(q, p) = q^2/2 + p^2/2$$
$$Q_t(p, q) = q\cos t + p\sin t,\ P_t(q, p) = -q\sin t + p\cos t.$$

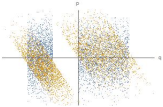
$$\mathcal{T}h(q) = \int_{\mathbb{P}} h(q\cos t + p\sin t)\mathfrak{g}(-q\sin t + p\cos t)\, dp$$

Consider $t = 0, \pi/2, \pi$. Usually $t > 0$ is small.

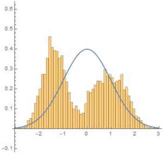
$t = \pi/8$



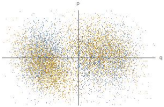
(0)



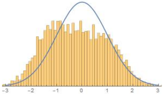
(0-1)



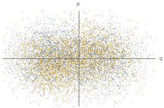
(1)



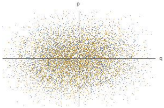
(1-2)



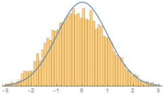
(5)



(5-6)



(9-10)



(10)

$\mathcal{L}^2_{\mathfrak{f}}$

Suppose $\mathfrak{f} > 0$ and $\int_{\mathbb{Q}} \mathfrak{f} < \infty$.

Inner product for the functions on $\mathbb{Q}$

$$\langle a, b\rangle = \langle a, b\rangle_{\mathfrak{f}} = \int_{\mathbb{Q}} \frac{a \cdot b}{\mathfrak{f}} \, .$$

Let $V(h) = \langle h, h\rangle$ and $\mathcal{L}^2 = \mathcal{L}^2_{\mathfrak{f}}(\mathbb{Q}) = \{h : \mathbb{Q} \to \mathbb{R} : V(h) < \infty\}$.

Clearly $\mathfrak{f} \in \mathcal{L}^2_{\mathfrak{f}}$.

Alternative space

Instead of h one can consider "likelihoods" $\tilde{h} = h/\mathfrak{f}$, then
$\langle \tilde{a}, \tilde{b}\rangle = \int_{\mathbb{Q}} \tilde{a}\tilde{b} \cdot \mathfrak{f}$.
$\tilde{\mathcal{T}}\tilde{h} = \int_{\mathbb{P}} \tilde{h} \circ H \cdot \mathfrak{g}$.

Properties of $\mathcal{T}$, $\mathcal{T}h = \int_{\mathbb{P}} (h \cdot \mathfrak{g}) \circ H$

Invariance

$\int_{\mathbb{Q}} \mathcal{T}h = \int_{\mathbb{Q}} h$
$\mathcal{T}\mathfrak{f} = \mathfrak{f}$.

$\int_{\mathbb{Q}} \mathcal{T}h = \int_{\mathbb{Q}} \int_{\mathbb{P}} (h \cdot \mathfrak{g}) \circ H = \int_{\mathbb{Q}} \int_{\mathbb{P}} (h \cdot \mathfrak{g}) = \int_{\mathbb{Q}} h \cdot \int_{\mathbb{P}} \mathfrak{g} = \int_{\mathbb{Q}} h.$

$\int_{\mathbb{P}} (\mathfrak{f} \cdot \mathfrak{g}) \circ H = \int_{\mathbb{P}} \mathfrak{f} \cdot \mathfrak{g} = \mathfrak{f} \cdot \int_{\mathbb{P}} \mathfrak{g} = \mathfrak{f}.$

Adjoint Operator, $\mathcal{T}_H^* = \mathcal{T}_{H^{-1}}$

Suppose that H is invertible. If we denote by $\mathcal{T}^\#$ the operator $\int_{\mathbb{P}} (h \cdot \mathfrak{g}) \circ H^{-1}$ then $\mathcal{T}^* = \mathcal{T}^\#$.

$\langle \mathcal{T}a, b \rangle = \int_{\mathbb{Q}} (\int_{\mathbb{P}} (a \cdot \mathfrak{g}) \circ H) \cdot \frac{b}{\mathfrak{f}} = \iint_{\mathbb{Q} \times \mathbb{P}} (a \cdot \mathfrak{g}) \cdot \frac{b}{\mathfrak{f}} \circ H^{-1} = \iint_{\mathbb{Q} \times \mathbb{P}} a \cdot \frac{\mathfrak{g}\,\mathfrak{f}}{\mathfrak{f}} \cdot \frac{b}{\mathfrak{f}} \circ H^{-1} =$
$\iint_{\mathbb{Q} \times \mathbb{P}} a \cdot \frac{(\mathfrak{g}\,\mathfrak{f}) \circ H^{-1}}{\mathfrak{f}} \cdot \frac{b}{\mathfrak{f}} \circ H^{-1} = \int_{\mathbb{Q}} \frac{a}{\mathfrak{f}} \cdot \int_{\mathbb{P}} (b \cdot \mathfrak{g}) \circ H^{-1} = \int_{\mathbb{Q}} \frac{a}{\mathfrak{f}} \cdot \mathcal{T}^\# b = \langle a, \mathcal{T}^\# b \rangle$

The self-adjoint operator

For an involution σ on $\mathbb{P}$, $(\sigma \circ \sigma = \mathrm{Id})$
define an involution Σ on $\mathbb{Q} \times \mathbb{P}$ by
$\Sigma(q, p) = (q, \sigma(p))$.

Symmetric $\mathfrak{g}$

If $\mathfrak{g} \circ \sigma = \sigma \circ \mathfrak{g}$ and $\Sigma \circ H = H^{-1} \circ \Sigma$ and $\left|\frac{\partial \sigma(p)}{\partial p}\right| = 1$ then $\mathcal{T} = \mathcal{T}^*$.

$\mathcal{T} h = \int_{\mathbb{P}} (h\mathfrak{g}) \circ H = \int_{\mathbb{P}} h(Q)\mathfrak{g}(P) = \int_{\mathbb{P}} h(Q)\mathfrak{g}(\sigma P) = \int_{\mathbb{P}} (h\mathfrak{g}) \circ (\Sigma \circ H) = \int_{\mathbb{P}} (h\mathfrak{g}) \circ H^{-1} \circ \Sigma = \int_{\mathbb{P}} (h\mathfrak{g}) \circ H^{-1} = \mathcal{T}^* h$

Hamiltonian with even $\mathfrak{g}$

HMC with even $\mathfrak{g}$

If $\mathfrak{g}(p) = \mathfrak{g}(-p)$ then $\mathcal{T}^* = \mathcal{T}^\# = \mathcal{T}$.

The HMC operator derived from the Hamiltonan motion is self-adjoint.

Going back in time

By changing the sign of time and the sign of the variables p and P_t we obtain exactly the same equations.
For the projection on $\mathbb{Q}$ the signs of p and P are irrelevant.

$$\dot{Q} = \frac{\partial Q}{\partial t} = \quad \frac{\partial \mathcal{H}}{\partial p}(Q, P), \quad \dot{P} = \frac{\partial P}{\partial t} = -\frac{\partial \mathcal{H}}{\partial q}(Q, P).$$

The norm in $\mathcal{L}^2_{\mathfrak{f}}$, under $\mathcal{T}$

Define $\mathcal{D}^2 : \mathcal{L}^2_{\mathfrak{f}} \circlearrowleft$ to be (the variance of $(h/\mathfrak{f}) \circ H$ w.r.t. $\mathfrak{g}(p)$):

$$\mathcal{D}h = \mathfrak{f} \cdot \left(\int_{\mathbb{P}} \left(\frac{h}{\mathfrak{f}} \circ H - \frac{\mathcal{T}h}{\mathfrak{f}} \right)^2 \mathfrak{g} \right)^{1/2} .$$

$\int_{\mathbb{P}} \frac{h}{\mathfrak{f}} \circ H \cdot \mathfrak{g} = \frac{\mathcal{T}h}{\mathfrak{f}}$ $\int_{\mathbb{P}} \frac{h}{\mathfrak{f}} \circ H \cdot \frac{\mathfrak{f}}{\mathfrak{f}} \mathfrak{g} = \int_{\mathbb{P}} h \circ H \cdot \frac{\mathfrak{g} \circ H}{\mathfrak{f}}$.

If $\mathcal{D}h(q) = 0$ then $\frac{h}{\mathfrak{f}} \circ H$ is equal to $\frac{\mathcal{T}h}{\mathfrak{f}}(q)$, $\mathfrak{g}(p)$-a.e.

by the coverage condition.

If $V(\mathcal{D}h) = 0$ then $\frac{h}{\mathfrak{f}}$ is constant $(\mathfrak{f} \cdot \mathfrak{g})(q, p)$ a.e.

$V(\mathcal{D}h) = V(h) - V(\mathcal{T}h)$

$V(\mathcal{D}h) = \int_{\mathbb{Q}} \mathfrak{f} \cdot \left(\frac{\mathcal{D}h}{\mathfrak{f}} \right)^2 = V(h) - V(\mathcal{T}h)$

$\mathcal{D}h = \mathfrak{f} \cdot \left(\int_{\mathbb{P}} \left((h/\mathfrak{f}) \circ H - (\mathcal{T}h/\mathfrak{f})\right)^2 \mathfrak{g}\right)^{1/2}$

With $X_q(p) = \frac{h}{\mathfrak{f}} \circ H(q,p)$

$\mathbb{E}_{\mathfrak{g}} X_q = \frac{\mathcal{T}h}{\mathfrak{f}}(q).$ $\quad \int_{\mathbb{P}} X_q \cdot \mathfrak{g} = \int_{\mathbb{P}} \frac{h}{\mathfrak{f}} \circ H \cdot \frac{\mathfrak{f}}{\mathfrak{f}} \mathfrak{g} = \int_{\mathbb{P}} h \circ H \cdot \frac{\mathfrak{g} \circ H}{\mathfrak{f}} = \frac{1}{\mathfrak{f}} \cdot \int_{\mathbb{P}} (h \cdot \mathfrak{g}) \circ H$

$\mathbb{D}^2_{\mathfrak{g}}(X_q) = \left(\frac{\mathcal{D}h}{\mathfrak{f}}\right)^2 (q).$

$\mathbb{D}^2_{\mathfrak{g}}(X_q) = \mathbb{E}_{\mathfrak{g}}(X_q^2) - (\mathbb{E}_{\mathfrak{g}}(X_q))^2$

$\mathbb{E}_{\mathfrak{g}}(X_q^2) = \int_{\mathbb{P}} \left(\frac{h}{\mathfrak{f}}\right)^2 \circ H \cdot \mathfrak{g}$

$\int_{\mathbb{Q}} \mathfrak{f} \cdot \mathbb{E}_{\mathfrak{g}}(X_q^2) = V(h)$ $\quad \iint_{\mathbb{Q}\times\mathbb{P}} \left(\frac{h}{\mathfrak{f}}\right)^2 \circ H \cdot (\mathfrak{g} \cdot \mathfrak{f}) \circ H = \iint_{\mathbb{Q}\times\mathbb{P}} \frac{h^2}{\mathfrak{f}} \cdot \mathfrak{g}$

$\int_{\mathbb{Q}} \mathfrak{f} \cdot (\mathbb{E}_{\mathfrak{g}} X_q)^2 = \int_{\mathbb{Q}} \frac{(\mathcal{T}h)^2}{\mathfrak{f}} = V(\mathcal{T}h)$

$V(\mathcal{D}h) = V(h) - V(\mathcal{T}h)$

$V(\mathcal{D}h) = \int_{\mathbb{Q}} \mathfrak{f} \cdot \left(\frac{\mathcal{D}h}{\mathfrak{f}}\right)^2 = \int_{\mathbb{Q}} \mathfrak{f} \cdot \mathbb{D}^2_{\mathfrak{g}}(X_q) = \int_{\mathbb{Q}} \mathfrak{f} \cdot \left(\mathbb{E}_{\mathfrak{g}}(X_q^2) - \mathbb{E}_{\mathfrak{g}}{}^2(X_q)\right)$

Monotonicity of V

Then norm V decreases and converges under iterations

$$V(\mathcal{T}h) = V(h) - V(\mathcal{D}h)$$
$$V(\mathcal{T}^n h) \searrow V_\infty(h).$$

$$\begin{aligned}
V(\mathcal{D}h) &= \int_{\mathbb{Q}} \mathfrak{f} \int_{\mathbb{P}} \left(\frac{h}{\mathfrak{f}} \circ H - \frac{\mathcal{T}h}{f}\right)^2 \mathfrak{g} \\
&= \iint_{\mathbb{Q}\times\mathbb{P}} \left(\frac{h}{\mathfrak{f}} \circ H - \frac{\mathcal{T}h}{f}\right)^2 \mathfrak{f}\,\mathfrak{g} \\
&= \iint_{\mathbb{Q}\times\mathbb{P}} \left(\left(\frac{h}{\mathfrak{f}} \circ H\right)^2 \mathfrak{f}\,\mathfrak{g} - 2\left(\frac{h}{f} \circ H \cdot \frac{\mathcal{T}h}{\mathfrak{f}}\right) \mathfrak{f}\,\mathfrak{g} + \left(\frac{\mathcal{T}h}{\mathfrak{f}}\right)^2 \mathfrak{f}\,\mathfrak{g}\right) \\
&= \iint_{\mathbb{Q}\times\mathbb{P}} \left(\frac{h}{\mathfrak{f}}\right)^2 \mathfrak{f}\,\mathfrak{g} - 2\int_{\mathbb{Q}} \frac{\mathcal{T}h}{\mathfrak{f}} \int_{\mathbb{P}} \frac{h}{f} \circ H \cdot \mathfrak{f}\,\mathfrak{g} + \int_{\mathbb{Q}} \frac{(\mathcal{T}h)^2}{\mathfrak{f}} \int_{\mathbb{P}} \mathfrak{g} \\
&= V(h) - 2V(\mathcal{T}h) + V(\mathcal{T}h)
\end{aligned}$$

By invariance properties

$\iint A \circ H \cdot \mathfrak{f}\,\mathfrak{g} = \iint A \circ H \cdot (\mathfrak{f}\,\mathfrak{g}) \circ H = \iint A \cdot \mathfrak{f}\,\mathfrak{g}$ and
$\int_{\mathbb{P}} (h/\mathfrak{f}) \circ H \cdot \mathfrak{f}\,\mathfrak{g} = \int_{\mathbb{P}} (h/\mathfrak{f}) \circ H \cdot (\mathfrak{f}\,\mathfrak{g}) \circ H = \int_{\mathbb{P}} h \circ H \cdot \mathfrak{g} \circ H = \mathcal{T}h$

Convergence

Convergence Theorem

If $\mathfrak{f}$, $\mathfrak{g}$ and H have **invariance properties**
and satisfy the **coverage condition**
and $\mathcal{T}$ is **self adjoint**
then the iterates of any initial density will align with $\mathfrak{f}$:

$$\mathcal{T}^n h \to \alpha \cdot \mathfrak{f},$$

where $\alpha = \int h / \int \mathfrak{f}$ *is the normalizing constant.*

Assumptions

$\mathcal{T} h = \int_{\mathbb{P}} (h \cdot \mathfrak{g}) \circ H$

$(\mathcal{T} h)(q) = \int_{\mathbb{P}} (h \cdot \mathfrak{g}) \circ H(q, p)\, dp = \int_{\mathbb{P}} h(Q(q, p)) \cdot \mathfrak{g}(P(q, p))\, dp$

For (almost) any q the function $p \mapsto Q(q, p)$ maps $\mathbb{P}$ onto $\mathbb{Q}$.

But convergence in what sense, one might ask...

In case of not self adjoint operator

Without the self adjoint property use the operator $\mathcal{S} = \mathcal{T}^* \circ \mathcal{T}$.

Convergence Theorem

If $\mathfrak{f}$, $\mathfrak{g}$ and H have **invariance properties**
and either H or H^{-1} satisfies the **coverage condition**
then the iterates of any initial density will align with $\mathfrak{f}$:

$$\mathcal{S}^n h \to \alpha \cdot \mathfrak{f},$$

where $\alpha = \int h / \int \mathfrak{f}$ *is the normalizing constant.*

Remark that the limit is invariant under both $\mathcal{T}$ and $\mathcal{T}^*$.

Assumptions

$\mathcal{T}h = \int_{\mathbb{P}} (h \cdot \mathfrak{g}) \circ H$

$(\mathcal{T}h)(q) = \int_{\mathbb{P}} (h \cdot \mathfrak{g}) \circ H(q, p)\, dp = \int_{\mathbb{P}} h(Q(q, p)) \cdot \mathfrak{g}(P(q, p))\, dp$

For (almost) any q the function $p \mapsto Q(q, p)$ maps $\mathbb{P}$ onto $\mathbb{Q}$.

$V(\mathcal{T}^n h) \searrow V_\infty(h)$ then $h_\infty = \alpha\mathfrak{f}$ (a.e.), with $\alpha = \int h / \int \mathfrak{f}$.

If $V(\mathcal{T}^n h - h_\infty) \to 0$

then $\mathcal{T} h_\infty = h_\infty$, $V(\mathcal{D} h_\infty) = 0$, $\mathcal{D} h_\infty = 0$.
The **coverage condition** implies $h_\infty / \mathfrak{f} = \alpha$

When a (sub) sequence $\mathcal{T}^n h$ converges weakly to h_w

then there is an m large enough and a subsequence n_k such that $\mathcal{T}^{2n_k}(\mathcal{T}^m h) \rightharpoonup h_w$. *Abuse notation and denote $\mathcal{T}^m h$ by h and n_k by n.*

$V(h_w) \leq V_\infty$ as (usual)
$V(h_w) = \langle h_w, h_w \rangle \leftarrow \langle h_w, \mathcal{T}^{2n} h \rangle \leq \sqrt{V(h_w)(V_\infty + \epsilon)}$.

$V(h_w) \geq V_\infty$ as (usual for self-adjoint)
$V_\infty - \epsilon \leq \langle \mathcal{T}^n h, \mathcal{T}^n h \rangle = \langle h, \mathcal{T}^{2n} h \rangle \to \langle h, h_w \rangle \leq \sqrt{(V_\infty + \epsilon) V(h_w)}$

$V(h_w) = V_\infty$ implies $V(\mathcal{T}^{2n} h - h_w) \to 0$, as (usual)
If $a_m \rightharpoonup a$ and $||a_m|| \to ||a||$ then $||a_m - a||^2 = ||a_m||^2 - 2\langle a_m, a \rangle + ||a||^2 \to 0$

Each sequence

has a subsequence converging to the same limit, *i.e.* the sequence converges to this limit.

Rate of convergence

Kernel operator approach (same as dominant convergence),

Curvature approach.

Both seem to provide exponential convergence in case of bounded, strongly positive $U''(q)$ and fail if the extra assumption is skipped.

Kernel approach

Recall **coverage assumption**: for (almost) every q the map $p \mapsto Q(q, p)$, denoted by $\mathcal{Q}_q(p)$ covers $\mathbb{Q}$. Suppose it is smoothly invertible.

$$(\mathcal{T}h)(q) = \int_{\mathbb{P}} h \circ H \cdot \mathfrak{g} \circ H = \int_{\mathbb{Q}} \frac{h(Q)K(q,Q)}{f(Q)}\, dQ$$

$$= \int_{\mathbb{P}} \frac{h \circ H \mathfrak{g} \circ H \mathfrak{f} \circ H}{\mathfrak{f} \circ H} = \int_{\mathbb{P}} \frac{h(Q) \cdot \mathfrak{g}(p) \cdot \mathfrak{f}(q)}{\mathfrak{f}(Q)}\, dp = \int_{\mathbb{Q}} \frac{h(Q) \cdot (\mathfrak{g}(\mathcal{Q}_q^{-1}(Q))\mathfrak{f}(q) |\frac{\partial \mathcal{Q}}{\partial p}|^{-1})}{\mathfrak{f}(Q)}\, dQ$$

Kernel $K(q, Q) = \mathfrak{g}(\mathcal{Q}_q^{-1}(Q))\mathfrak{f}(q)|\frac{\partial \mathcal{Q}}{\partial p}|^{-1}$.

If $K(q, Q) \in \mathcal{L}_{\mathfrak{f}}^{22}$ then the operator $\mathcal{T}$ is compact.

$$||K(q, Q)||_{\mathfrak{f}}^2 = \int_q \frac{\int_Q \frac{K(q,Q)^2}{\mathfrak{f}(Q)}\, dQ}{\mathfrak{f}(q)}\, dq$$

$$||K(q, Q)||_{\mathfrak{f}}^2 = \iint_{q,p} |\frac{\partial \mathcal{Q}_q}{\partial p}|^{-1} \mathfrak{g}(p)\mathfrak{g}(P)\, dp\, dq$$

How to deal with $\mathcal{Q}_q(p)$

$$\left(\frac{\partial \dot{Q}}{\partial q}\right) = \frac{\partial}{\partial q}\frac{\partial \mathcal{H}}{\partial P} = \frac{\partial^2 \mathcal{H}}{\partial P^2}\frac{\partial P}{\partial q},$$

$$\left(\frac{\partial \dot{Q}}{\partial p}\right) = \frac{\partial}{\partial p}\frac{\partial \mathcal{H}}{\partial P} = \frac{\partial^2 \mathcal{H}}{\partial P^2}\frac{\partial P}{\partial p},$$

$$\left(\frac{\partial \dot{P}}{\partial q}\right) = -\frac{\partial}{\partial q}\frac{\partial \mathcal{H}}{\partial Q} = -\frac{\partial^2 \mathcal{H}}{\partial Q^2}\frac{\partial Q}{\partial q},$$

$$\left(\frac{\partial \dot{P}}{\partial p}\right) = -\frac{\partial}{\partial p}\frac{\partial \mathcal{H}}{\partial Q} = -\frac{\partial^2 \mathcal{H}}{\partial Q^2}\frac{\partial Q}{\partial p}.$$

with initial condition,

$$\frac{\partial Q}{\partial q}(0) = I, \; \frac{\partial Q}{\partial p} = 0, \; \frac{\partial P}{\partial q}(0) = 0, \; \frac{\partial P}{\partial q}(0) = I.$$

The solution bears the form,

$$\begin{pmatrix} \frac{\partial Q}{\partial q} \\ \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} \\ \frac{\partial P}{\partial p} \end{pmatrix} = \begin{pmatrix} \cos(\sqrt{UV}) \\ \sin(\sqrt{UV})\sqrt{UV}^{-1}V^{-1} \\ -\sin(\sqrt{VU})\sqrt{VU}^{-1}U^{-1} \\ \cos(\sqrt{VU}) \end{pmatrix} \tag{1}$$

with $U = \int_0^t \frac{\partial^2 \mathcal{H}}{\partial P^2} ds$ and $V = \int_0^t \frac{\partial^2 \mathcal{H}}{\partial Q^2} ds$, $sinA = (e^{iA} - e^{-iA})/(2i)$ and $cosA = (e^{iA} + e^{-iA})/2$ and the exponential and square root functions of (positive) matrices are well defined.

Jaśmin, Jasmine, Jasminum L.