بسم الله الرحمن الرحیم

جلسه ۱۰،۵، ۱۴۰۰

* Hamiltonian Monte Carlo (HMC)

① Markov chain Monte Carlo (MCMC)

Recall that a typical classification of stochastic processes can be represented as follows:
فرآیندهای تصادفی

Independent Processes

Ⓐ Completely Random (کاملاً تصادفی (هیچ ارتباطی با یکدیگر ندارند))

Dependent parameter

$$\underbrace{\{x\} : \{x_1(t_1), x_2(t_2), \ldots x_N(t_N)\}}_{\text{Time - Series}}$$

Independent parameter

$$p(x_N, t_N; x_{N-1}, t_{N-1}; x_{N-2}, t_{N-2}; \ldots ; x_1, t_1)$$

$$= p(x_N, t_N)\, p(x_{N-1}, t_{N-1}) \ldots \times p(x_1, t_1)$$

No - Correlation $\longleftrightarrow \langle x(t)\, x(t')\rangle = \delta_D(t-t')$

Ⓑ Dependent Stochastic Processes

Correlation time scale (فرآیندهای همبسته (وابسته))

Angle
Lenght

$$p(x_N, t_N; x_{N-1}, t_{N-1}; \ldots ; x_1, t_1)$$
N-Joint PDF

$$= p(x_N, t_N | x_{N-1}, t_{N-1}; x_{N-2}, t_{N-2}; \ldots ; x_1, t_1)$$

$p(x_N, t_N)$

$$\underbrace{p(x_{N-1}; t_{N-1} \mid x_{N-2}, t_{N-2}, \ldots ; x_1, t_1)}_{p(x_{N-1}, t_{N-1})}$$

$$\vdots$$

$$\underbrace{p(x_2, t_2 \mid x_1, t_1)}_{p(x_2, t_2)} \, p(x_1, t_1)$$

(c) Markov Processes فرآیندهای مارکوف

$$p(x_N, t_N; x_{N-1}, t_{N-1}; \ldots ; x_1, t_1) = p(x_N, t_N \mid x_{N-1}, t_{N-1})$$

$$\times \, p(x_{N-1}, t_{N-1} \mid x_{N-2}, t_{N-2}) \times \ldots \times p(x_2, t_2 \mid x_1, t_1)$$

$$p(x_1, t_1)$$

$$p(x_N, t_N \mid x_{N-1}, t_{N-1}, x_{N-2}, t_{N-2}, \ldots ; x_2, t_2; x_1, t_1)$$

$$\boxed{= p(x_N, t_N \mid x_{N-1}, t_{N-1})}$$ — از جمله ویژگی های فرآیندهای مارکوف

old

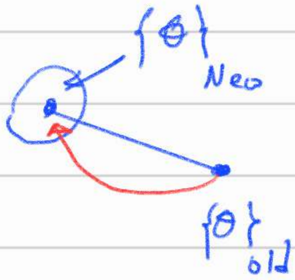
New Posit

MCMC algorithm for optimization

Import Data $\longrightarrow \{D\} : \{(x_i, y_i)\}$ $i=1, N$

$\{\theta\} . \{\theta_j\},$ $j=1, M$

Select $\{\theta\}_{old}$, $\mathcal{L}_{old} \sim \exp(-\chi^2_{old}(\{\theta\}_{old}))$



loop on MCMC $\longrightarrow \mathcal{O}(M_{MCMC})$

Select $\{\theta\}_{New}$ from $\{\theta\}_{old}$

Compute $\chi^2_{New}(\{\theta\}_{New})$

$$\Delta\chi^2 \equiv \chi^2_{New}(\{\theta\}_{New}) - \chi^2_{old}(\{\theta\}_{old})$$

Check acceptance Rate

$$AC = \min\{1, e^{-\Delta\chi^2/2}\}$$

Metropolis algorithm

$\xi =$ Call Random number

if $\xi \leq e^{-\Delta\chi^2/2}$ then

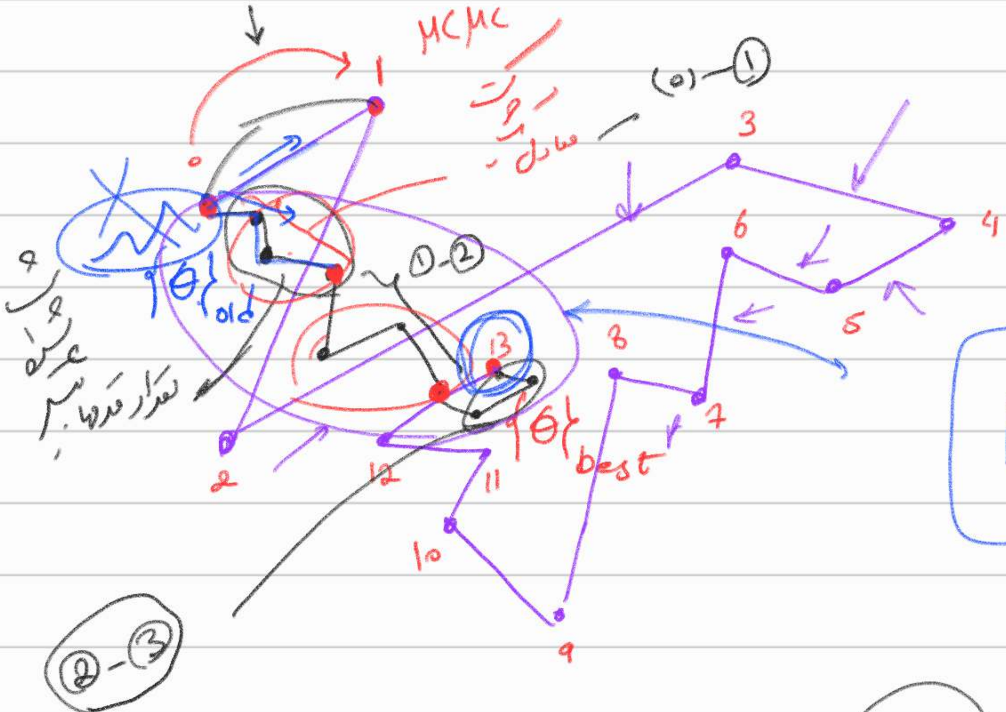
$\{\theta\}_{old} = \{\theta\}_{New}$

$\chi^2_{old} = \chi^2_{New}$

End if

write $\{\theta\}_{old}$, χ^2_{old}.

End loop



$M_{MCMC} \sim 500000$

10^6.

0 — 1 ۳ قدم تصادفی

1 — 2 ۳ قدم تصادفی

2 — 3 ۳ قدم تصادفی → Best location

هدفمند رو هوشمند استراتژی حرکت انتخاب شود

$\theta_{old} \to \theta_{new}$ معادلات تحول مبتنی بر هامیلتون استفاده می‌شود

الگوریتم؟ ؟

راه‌کار چیست؟ → HMC

② Hamiltonian Monte Carlo (HMC)

هامیلتونی

نسبت به MCMC به لحاظ الگوریتمی پیچیده‌تر است

more complex than MCMC

To Reduce time consuming

کاهش زمان محاسبات شود.

$$\mathcal{O}(M_{MCMC}) \gg \mathcal{O}(M_{HMC})$$

$$M_{HMC} = \left(M_{MCMC} \times \underbrace{M_{\text{Intermediat Jump}}}_{\text{کاهش می‌یابد}}\right)$$

$$\{\theta\}_{old} \xrightarrow{HMC} \{\theta\}_{New}$$

$$\{\theta\}_{Initial} \longrightarrow \{\theta\}_{Best}$$

ایده اصلی این موضوع این است که به جای حرکت کاملاً تصادفی (مارکوفی) از دینامیک مستخرج شده از یک هامیلتونی مؤثر استفاده کنیم.

Effective Hamiltonian

Phenomenological Suggestion

$$\mathcal{H}(\{q\},\{p\}) \longrightarrow \mathcal{H}(\{\theta\},\{u\})$$

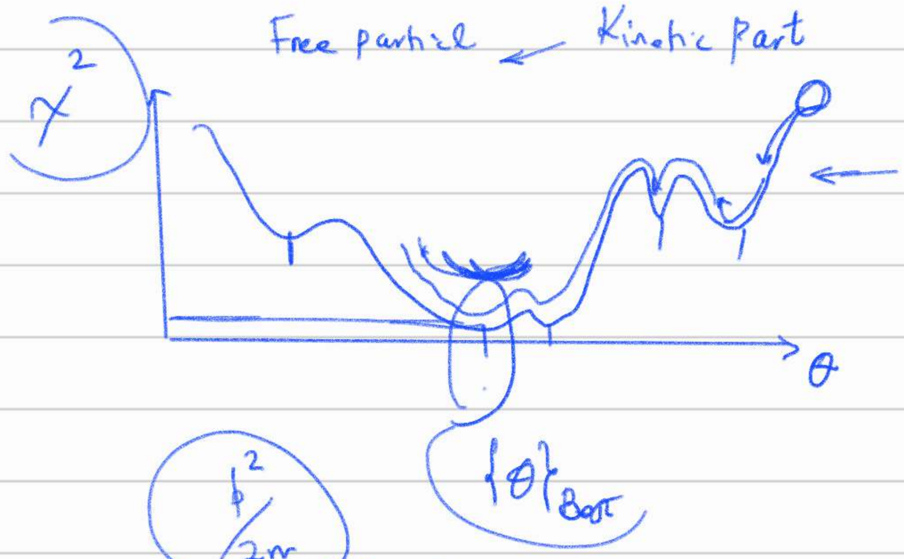
✓ در مدل‌سازی خطی وجود داشت

?

$$\mathcal{H}(\{\theta\},\{u\}) = K(\{u\}) + V(\{\theta\})$$

Free particle ← Kinetic part

Potential part



$$\frac{p^2}{2m}$$

$$\mathcal{H}(\{\theta\},\{u\}) = \frac{u^T u}{2} + \frac{\chi^2(\{\theta\})}{2}$$

انرژی جنبشی

انرژی پتانسیل را بررسی می‌کند

کمینه کردن این هامیلتونی یعنی

$\{\theta\}_{Best}$

finding min of χ^2 → finding min of $\mathcal{H}(u,\theta)$

$$\mathcal{L}_{MCMC} = e^{-\frac{\chi^2}{2}} \longrightarrow \mathcal{L}_{HMC} = e^{-\mathcal{H}}$$

$$\min\{\mathcal{H}(\{u\},\{\theta\})\} = \mathcal{H}(\{u\},\{\theta\}_{Best})$$

کمیت اضافی والتیمکی ظاهر شود

$$\mathcal{L}_{HMC} \sim e^{-\mathcal{H}} \sim \mathcal{L}(\{D\}|\{\theta\})\,\underbrace{\mathcal{N}(u,0,1)}_{e^{-\frac{u^T u}{2}}}$$

$$\{\theta\}_{old} \longrightarrow \{\theta\}_{New} \longleftarrow \text{MCMC}$$

$$\left\{ \begin{array}{l} \underline{\{\theta\}_{old}} \rightarrow \{\theta\}^{HMC}_{old} \\ \underline{\{u\}_{old}} - \{u\}^{HMC}_{old} \end{array} \right\}$$

Some Interesting Properties

$$\mathcal{H}(\{u\},\{\theta\}) = K(\{u\}) + V(\{\theta\})$$

$$\boxed{\mathcal{H} = \frac{u^T u}{2} + \frac{\chi^2(\{\theta\})}{2}}$$

$$\{\theta\} \longrightarrow \{q\}$$

$$\{u\} \longrightarrow \{p\}$$

$$\boxed{\begin{array}{l} \dot{p} = -\dfrac{\partial \mathcal{H}}{\partial q} \\ \dot{q} = \dfrac{\partial \mathcal{H}}{\partial p} \end{array}}$$

$$\boxed{\dot{u} = -\frac{\partial \mathcal{H}}{\partial \theta}}, \quad \boxed{\dot{\theta} = \frac{\partial \mathcal{H}}{\partial u}}$$

Recall
$$\begin{cases}\dfrac{dp}{dt} = -\dfrac{\partial H}{\partial q}\\[2mm] \dfrac{dq}{dt} = \dfrac{\partial H}{\partial p}\end{cases} \quad , \quad H = \underbrace{\frac{p^2}{2} + \frac{1}{2}q^2}_{\text{نوسانگر هماهنگ}}$$

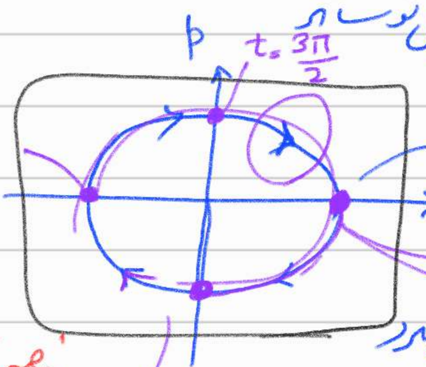
$\dot p = -\dfrac{\partial H}{\partial q} = -q$

$\dot q = p \longrightarrow \ddot q = \dot p$ $\quad \ddot q = -q \longrightarrow \underbrace{\ddot q + q = 0}$

معادله دیفرانسیل نوسانگر

$q(t) = \cos(t)$

$p(t) = -\sin(t)$

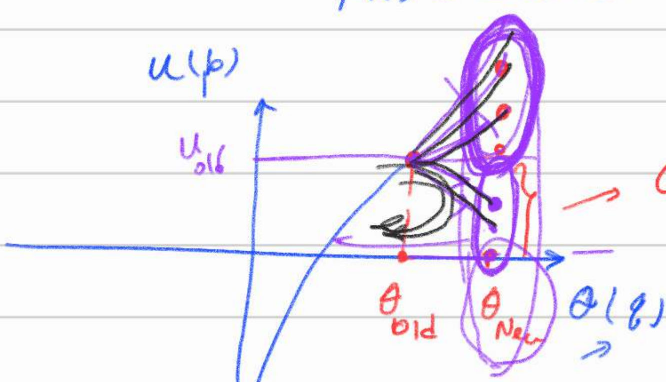


$q(t=0) = \cos(0) = 1$

$p(t=0) = \sin(0) = 0$

$t = \pi/2 \to q(t=\pi/2) = \cos(\pi/2) = 0$

$p(t=\pi/2) = -\sin(\pi/2) = -1$



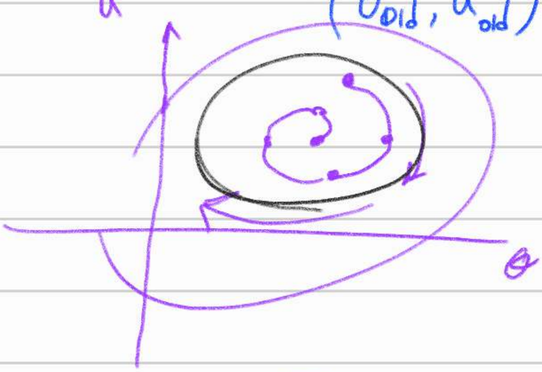
$\chi(\theta_{old}) < \chi(\theta_{New})$

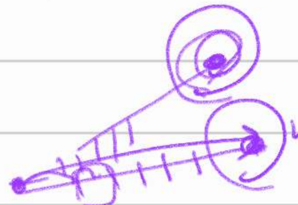
$\dot p = -\dfrac{\partial H}{\partial q}$

$\downarrow$

$\dot u = -\dfrac{\partial H}{\partial \theta} = -\dfrac{\partial}{\partial \theta}\left\{\dfrac{u^T u}{2} + \dfrac{\chi^2}{2}\right\}$

$$\boxed{\dot u = -\frac{1}{2}\frac{\partial \chi^2}{\partial \theta}}$$

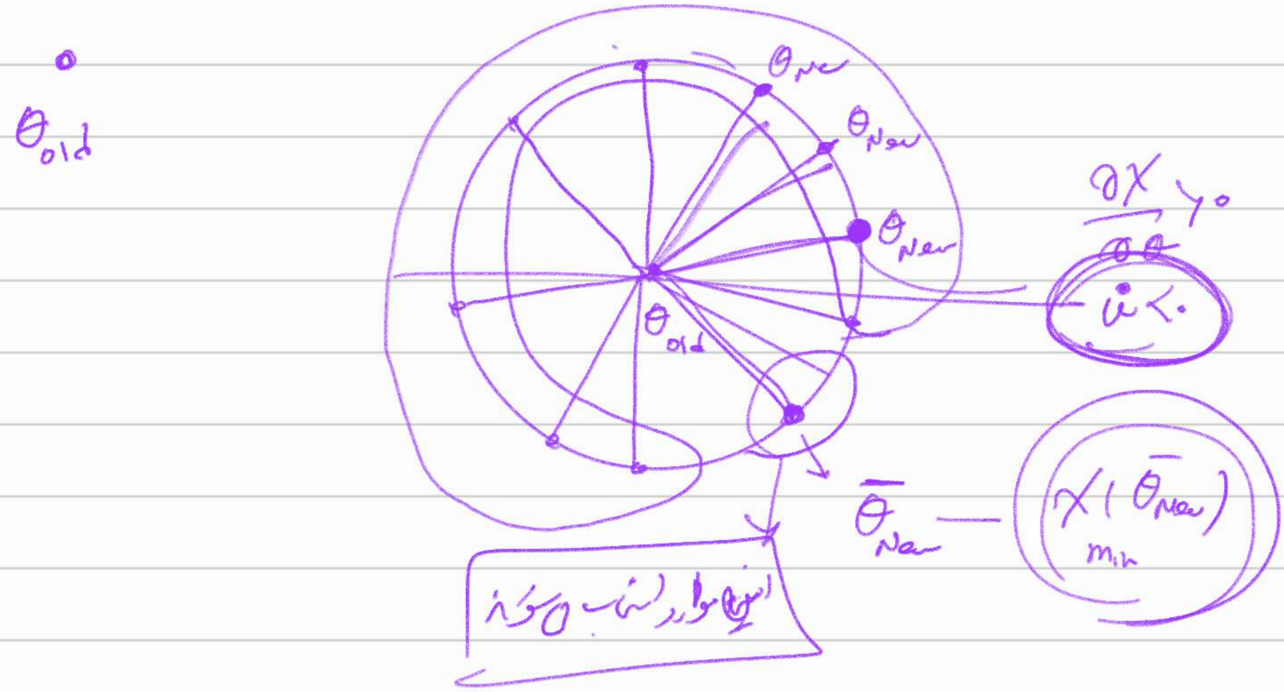




$\chi^2(\theta_1) < \chi^2(\theta_2)$

$\dfrac{\partial \chi^2}{\partial \theta} > 0$ $\quad \boxed{\dot u < 0}$

$$\boxed{\underbrace{u(i+1)}_{u_{New}} - \underbrace{u(i)}_{u_{old}} < 0}$$



$$H(\theta, u) = \frac{u^2}{2} + \frac{\chi^2(\theta)}{2} \longrightarrow \frac{p^2}{2m} + V(q)$$

$$q^2/2$$

$$\frac{\chi^2}{2} \triangleq \frac{\Delta\theta^T \cdot F \cdot \Delta\theta}{2}$$

$$\simeq \theta^2 + \cdots$$

$$\simeq q^2$$

HMC Algorithm