Wyklad5

26 października 2012
13:53

4. Kwantowy efekt Zenona

Rozważmy układ w stanie początkowym $|0\rangle$ ewoluujący zgodnie z hamiltonianem H

$$|\psi(t)\rangle = e^{-iH\cdot t}|0\rangle =$$

$$= a_0(t)|0\rangle + \sum_{i>0} a_i(t)|i\rangle$$

Mierzymy układ w bazie $|0\rangle, |1\rangle, |2\rangle, \ldots$
prawdop. że pozostanie w stanie początkowym:

$$p_0 = |\langle 0|\psi(t)\rangle|^2 = \left|\langle 0| e^{-iHt}|0\rangle\right|^2$$

Dla krótkich t możemy rozwinąć:

$$p_0 = \langle 0| 1 - iHt - \frac{H^2t^2}{2}|0\rangle \langle 0| 1 + iHt - \frac{H^2t^2}{2}|0\rangle$$

$$= 1 - it\langle H\rangle + it\langle H\rangle + \langle H\rangle^2 t^2 - t^2\langle H^2\rangle + O(t^4)$$

$$= 1 - t^2 \Delta^2 H + O(t^4)$$

Robimy pomiar N krotnie w czasie t
Jakie jest prawdop. że pozostanie w stanie $|0\rangle$

$$p_0^{(N)} \approx \left(1 - \left(\frac{t}{N}\right)^2 \Delta^2 H + O\left(\left(\frac{t}{N}\right)^4\right)\right)^N \approx$$

$$\approx 1 - \frac{t^2}{N}\Delta^2 H \xrightarrow{N\to\infty} 1$$

Stan nie ewoluuje, częsty pomiar „zamraża" go w stanie początkowym
„Ruch jest złudzeniem" ...

II Klasyczna teoria estymacji

1. Estymacja parametru (podejście klasyczne – nie Bayesowskie)

$$\Theta \xrightarrow{p_\Theta(x)} X$$

Θ – nieznany (ale określony) parametr (y)

X – obserwowane zdarzenie (a)

$p_\Theta(x)$ – rodzina rozkładów prawdopodobieństwa opisująca zmienną(e) losową X

Chcemy wywnioskować Θ na podstawie X

$\tilde{\Theta}(X)$ – estymator, ma nam dać informację o Θ

Przykład

$X = (x_1, \ldots, x_N)$

$x_i = \Theta + w_i$ niezależne zmienne losowe

$w_i \sim N(0, \sigma^2)$

rozkład gaussowski, średnia, wariancja

$x_i \sim N(\Theta, \sigma^2)$

$$p_\Theta(X) = p_\Theta(x_1) \cdot \ldots \cdot p_\Theta(x_N)$$

$$p_\Theta(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\Theta)^2}{2\sigma^2}}$$

Mierzymy $(x_1, \ldots, x_N)$

Jeśli byłby „dobry" estymator Θ

- $\tilde{\Theta}(x) = x_1$
- $\tilde{\Theta}(x) = \frac{\sum_i x_i}{N}$

1.1 Optymalne estymatory

Nieobciążoność estymatora

$$\langle \tilde{\Theta} \rangle = \Theta \quad \text{tzn.} \quad \forall_{\Theta} \int \tilde{\Theta}(x)\, p_{\Theta}(x)\, dx = \Theta$$

średnio daje prawdziwą wartość

- $\tilde{\Theta}(x) = x_1$ $\qquad \langle \tilde{\Theta} \rangle = \langle x_1 \rangle = \Theta$ ok
- $\tilde{\Theta}(x) = \frac{\sum_i x_i}{N}$ $\qquad \langle \tilde{\Theta} \rangle = \frac{1}{N} \sum_i \langle x_i \rangle = \Theta$ ok

Optymalność estymatora

Wariancja estymatora

$$\Delta^2 \tilde{\Theta} = \langle (\tilde{\Theta} - \Theta)^2 \rangle = \int (\tilde{\Theta}(x) - \Theta)^2\, p_{\Theta}(x)\, dx$$

Chcemy aby $\Delta^2 \tilde{\Theta}$ była jak najmniejsza

Uwaga: można napisać estymator, który będzie dobry dla pewnych wartości Θ a gorszy dla innych.
trywialny przykład $\tilde{\Theta} = \Theta_0$ nie zależne od danych
– to jest przykład estymatora obciążonego

Zazwyczaj bardziej interesujący jest:

„Minimum variance unbiased estimator" (MVU)

Nieobciążony estymator z minimalną wariancją

- minimalna wariancja dla wszystkich θ

Przykład

$$x_i \sim N(\theta, \sigma^2)$$

$$\tilde{\theta}(x_1, \dots, x_N) = \frac{1}{N} \sum_i x_i$$

$$s^2_{\tilde{\theta}} = \left\langle \left(\frac{1}{N} \sum_i x_i - \theta \right)^2 \right\rangle = \frac{1}{N^2} \left\langle (x_1 - \theta + x_2 - \theta + \dots x_N - \theta)^2 \right\rangle =$$

$$= \frac{1}{N^2} N \sigma^2 = \frac{\sigma^2}{N}$$ jest MVU (dowód później)

{ gdyby wziąć $\tilde{\theta}(x_1 \dots x_N) = x_1$ $\quad s^2_{\hat{\theta}} = \sigma^2$

Przykład: (MVU może nie istnieć)

$$x_1 = N(\theta, 1)$$

$$x_2 = \begin{cases} N(\theta, 1) & \theta \geq 0 \\ N(\theta, 2) & \theta < 0 \end{cases}$$

Dla $\theta \geq 0$ najlepszy estymator to $\tilde{\theta} = \frac{1}{2}(x_1 + x_2)$,

Dla $\theta < 0$ najlepszy estymator to $\tilde{\theta} = \frac{2}{3} x_1 + \frac{1}{3} x_2$.

Są nieobciążone ale nie ma jednego estymatora dającego minimalną wariancję dla wszystkich θ

Jak szukać optymalnych estymatorów?

1.2 Ograniczenie Cramera-Rao (CR)

Chcemy mieć ograniczenie na najlepszy możliwy estymator nieobciążony wtedy jak znajdziemy estymator wysycający ograniczenie wiemy że jest optymalny

Intuicja: $p_\theta(x)$ im bardziej różnią $p_\theta(x)$

zmienia się z θ tym większa precyzja estymacji
w ograniczeniu pojawia się $\frac{d}{d\theta} p_\theta(x)$

Założenia

$$\langle\tilde{\theta}\rangle = \int dx\, \tilde{\theta}(x)\, p_\theta(x) = \theta$$ – nieobciążoność

w zasadzie będziemy potrzebować:

$$(*)\quad \frac{d\langle\tilde{\theta}\rangle}{d\theta} = \int dx\, \tilde{\theta}(x) \frac{dp_\theta(x)}{d\theta} = 1$$ – lokalna nieobciążoność

$$(**)\quad \int dx \frac{dp_\theta(x)}{d\theta} = 0$$ – warunek regularności

$\parallel$ jeśli można zamienić kolejność całkowania i różniczkowania

$$\frac{d}{d\theta}\int dx\, p_\theta(x) = \frac{d}{d\theta} 1 = 0$$ spełniony

Wyprowadzenie:

$$\underbrace{\int dx\, p_\theta(x)\left(\tilde{\theta}(x)-\theta\right)^2}_{\Delta^2\tilde{\theta}} \cdot \underbrace{\int dx \frac{1}{p_\theta(x)}\left(\frac{dp_\theta(x)}{d\theta}\right)^2}_{F}$$

$$= \int dx\left[\sqrt{p_\theta(x)}\left(\tilde{\theta}(x)-\theta\right)\right]^2 \cdot \int dx\left(\frac{1}{\sqrt{p_\theta(x)}}\frac{dp_\theta(x)}{d\theta}\right)^2$$

$$\overset{C-S}{\geq}\left(\int dx\left(\tilde{\theta}(x)-\theta\right)\frac{dp_\theta(x)}{dx}\right)^2 = \Big(\underbrace{\int dx\, \hat{\theta}(x)\frac{dp_\theta(x)}{d\theta}}_{\overset{(*)}{=}1} - \theta\underbrace{\int dx \frac{dp_\theta(x)}{d\theta}}_{\overset{(**)}{=}0}\Big)^2$$

$$\Delta^2\tilde{\theta} \geq \frac{1}{F}, \qquad F = \int dx \frac{1}{p_\theta(x)}\left(\frac{dp_\theta(x)}{d\theta}\right)^2$$

$$\Delta^2\tilde{\theta} \geq \frac{1}{F}, \quad \underbrace{F = \int dx \, \frac{1}{p_\theta(x)} \left(\frac{dp_\theta(x)}{d\theta}\right)^2}_{\text{Informacja Fishera}}$$

F w ogólności zależy od θ

Uwaga:

Możemy spotkać inne równoważne postacie F

$$F = \left\langle \left(\frac{d}{d\theta} \log p_\theta(x)\right)^2 \right\rangle = \int dx \, p_\theta(x) \frac{1}{p_\theta(x)^2} \left(\frac{dp_\theta(x)}{d\theta}\right)^2 = \int \frac{1}{p_\theta(x)} \left(\frac{dp_\theta(x)}{d\theta}\right)^2$$

$$F = -\left\langle \frac{d^2}{d\theta^2} \log p_\theta(x) \right\rangle$$

jawnie widać addytywność

Jeśli mamy estymator nieobciążony dla którego $\Delta^2\tilde{\theta} = \frac{1}{F}$ wiemy że jest optymalny. Mówimy wtedy $\tilde{\theta}$ jest „efficient".

Addytywność F

Jeśli $p^{(12)}_\theta(x_1, x_2) = p^{(1)}_\theta(x_1) \cdot p^{(2)}_\theta(x_2)$ to

$$F^{(12)} = F^{(1)} + F^{(2)}$$

Wniosek: N niezależnych realizacji zmiennej losowej X: $F^{(N)} = N \cdot F$

$$\Delta^2\tilde{\theta} \geq \frac{1}{N \cdot F}$$

Przykład

$$x_i \sim N(\theta, \sigma^2) \qquad i = 1, \ldots, N$$

$$\ldots \; 1 \; \ldots \; e^{-\frac{(x_i-\theta)^2}{2\sigma^2}}$$

$$p_\theta(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\quad}{2\sigma^2}}$$

$$F = \left\langle \left(\frac{d}{d\theta} \log p_\theta(x_i)\right)^2 \right\rangle = \int dx\, p_\theta(x) \left(\frac{x_i-\theta}{\sigma^2}\right)^2 = \frac{\sigma^2}{\sigma^4} = \frac{1}{\sigma^2}$$

Mając N realizacji:

$$\delta^2 \tilde{\theta} \geq \frac{1}{NF} = \frac{\sigma^2}{N}$$

Czyli tyle ile wychodzi z estymatora $\tilde{\theta} = \frac{\sum_i x_i}{N}$

efficient estymator

Wysycanie nierówności CR

Patrząc na wyprowadzenie nierówności C-S wysycam iff:

$$\lambda(\theta)\sqrt{p_\theta(x)}\left(\hat{\theta}(x) - \theta\right) = \frac{1}{\sqrt{p_\theta(x)}} \frac{dp_\theta(x)}{d\theta}$$

$$\boxed{\frac{d}{d\theta} \log p_\theta(x) = \lambda(\theta)\left(\hat{\theta}(x) - \theta\right)}$$

Jeśli znajdziemy $\lambda(\theta)$, $\hat{\theta}(x)$ t. że powyższe równanie spełnione wysycamy C-R

Przykład

$$x_i = N(\theta, \sigma^2)$$

$$p_\theta(x) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\theta)^2}{2\sigma^2}}$$

$$\log p_\theta(x) = -\frac{N}{2}\log 2\pi\sigma^2 - \sum_{i=1}^{N} \frac{(x_i-\theta)^2}{2\sigma^2}$$

$$\frac{d}{d\theta}\log p_\theta(x) = \sum_i \frac{(x_i-\theta)}{\sigma^2} = \lambda(\theta)\left(\hat{\theta}(x) - \theta\right)$$

jeśli weźmiemy $\hat{\theta}(x) = \sum_i \frac{x_i}{N}$, $\lambda(\theta) = \frac{N}{\sigma^2}$

jeśli weźmiemy $\hat{\theta}(x) = \sum_i \frac{x_i}{N}$, $\lambda(\theta) = \frac{N}{\sigma^2}$

<u>Ogólnie:</u>

$$\log p_\theta(x) = \Lambda(\theta)\,\hat{\theta}(x) + a(\theta) + b(x) \quad , \quad a' = -\lambda \cdot \theta, \quad \Lambda' = \lambda$$

$$p_\theta(x) = e^{a(\theta)}\, e^{b(x)}\, e^{\Lambda(\theta)\,\hat{\theta}(x)}$$

Szczególny przypadek tzw. exponential family

$$p_\theta(x) = e^{a(\theta) + b(x) + c(\theta)\,d(x)}$$

gdzie żeby estymacja θ być efficient $\theta = -\frac{\frac{da}{d\theta}}{\frac{dc}{d\theta}}$

Jeśli $-\frac{\frac{da}{d\theta}}{\frac{dc}{d\theta}} \neq \theta$ wtedy nie istnieje estymator θ

który jest efficient ale istnieje dla funkcji od parametru $g(\theta) = -\frac{\frac{da}{d\theta}}{\frac{dc}{d\theta}}$

estymacja θ^2 jeśli $x_i = N(\theta, \sigma^2)$