Monte Carlo quadrature

Taichi Kato

① Empirical mean

Given X_i $(i=1,2,\dots M)$

[sequence of independent random variables
(PDF : π_x)

And, $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{pmatrix} = \begin{pmatrix} X_1(w) \\ X_2(w) \\ \vdots \\ X_3(w) \end{pmatrix}$ ← independent samples

Then, Empirical mean is

$$\overline{x}_M = \frac{1}{M}\sum_{i=1}^{M} x_i$$

(※ $\overline{x}_M$ is realisation of random valuable $\overline{X}_M = \frac{1}{M}\sum_{i=1}^{M} X_i$)

This consist Monte Carlo approximation to

$$\int_{\mathbb{R}} x \, \mu_x \, dx$$

Also, $\overline{x}_M$ provides an approximation to $\overline{x}$ as $M \to \infty$

cf which estimator is better?

1) $\hat{P}_M = \frac{1}{M}\sum_{i=1}^{M}(X_i - \overline{x}_M)(x_i - \overline{x}_M)^T$

2) $\hat{P}_M = \frac{1}{M-1}\sum_{i=1}^{M}(x_i - \overline{x}_M)(x_i - \overline{x}_M)^T$

In both cases, the value of $\hat{P}_M$ is the limit of $M \to \infty$. However, in case 1), $\mathbb{E}[\hat{P}_M] < P$

(P is coverience matrix of X_i)

So, 2) is better.

② Convergence of sequences of random variables

- X_m converges like

$$\mathbb{P}\big(\lim_{m\to\infty} X_m = \underset{\uparrow \text{ random variable}}{X}\big) = 1$$

- Convergence in probability is defined like

$$\forall \varepsilon > 0, \ \lim_{m\to\infty} P(|X_m - X| > \varepsilon) = 0$$

- Weak convergence (or distribution) is defined like

$\forall f$, f is bounded and continuous

$$\lim_{m\to\infty} \mathbb{E}[f(X_m)] = \mathbb{E}[f(X)]$$

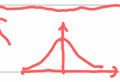
③ Central limit theorem

Given, random variable $\overline{X}_M$, PDF π_x, mean $\overline{x}$, finite variance σ^2, then random valuable Y_M is defined as

$$Y_M = \sqrt{\frac{M}{\sigma^2}}\,(\overline{X}_M - \overline{x})$$ ← random valuable of error

⇓

Y_M converges weakly to Gaussian random variable (mean: 0, variance: 1) as $M \to \infty$



④ Confidence interval

For quadrature rules, expectations converge as number of quadrature point goes to ∞.

⇓ For emprical means

Under repeated sampling of emprical mean $\overline{x}_M$, a confidence interval is constructed.

True mean $\overline{x}$ is contained in the interval with some chosen probability P (Usually $P = 0.95$)

⑤ chebychev's inequality

Given random valuable $\bar{X}_m$ (mean: $\bar{x}$, variable: σ^2),
then chebychev's inequality states that

$$\mathbb{P}(|Z_m| \geq k\sigma_m^2) \leq \frac{1}{k^2}, \quad k>0$$

$$\left(Z_m = \bar{X}_m - \bar{x}, \ \sigma_m^2 = \frac{\sigma^2}{m} \right)$$

(proof)

$$\chi_I \overset{def}{\equiv} \begin{cases} 1 & \text{if } x \in I \quad (I \subset \mathbb{R}) \\ 0 & \text{otherwise} \end{cases}$$

And choose $I = \{x \in \mathbb{R} : |x| \geq k\sigma_m^2\}$. Then,

$$\mathbb{P}(|Z_m| \geq k\sigma_m^2) = \mathbb{E}[\chi_I(Z_m)]$$

$$\leq \mathbb{E}\left[\left(\frac{Z_m}{k\sigma_m^2}\right)^2\right]$$

$$= \frac{1}{(k\sigma_m)^2}\mathbb{E}[Z_m^2] = \frac{1}{k^2} \quad \square$$

(cf)

- Strong law of large number

$$\mathbb{P}\left(\lim_{m\to\infty} \bar{X}_m = \bar{x}\right) = 1 \quad (※ \ \mathbb{E}[|X|] < \infty)$$

- Weak law of large number

$$\forall \varepsilon > 0, \ \lim_{m\to\infty} \mathbb{P}(|\bar{X}_m - \bar{x}| > \varepsilon) = 0$$

Both laws imply that Monte Carlo approximations $\bar{X}_m$ converge to the true mean $\bar{x}$ as $m \to \infty$.
However, they don't imply how $\bar{x}_m$ converge to $\bar{x}$. ($\leftrightarrow$ Chebychev's inequality do)

⑥ Monte Carlo approximation

For given PDF π_X, the Monte Carlo approximation to $\mathbb{E}[f(x)]$ is given by

$$\bar{f}_M = \sum_{i=1}^{M} w_i f(x_i) \quad \cdots (*)$$

$$\left(※_1 \ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{pmatrix} = \begin{pmatrix} X_1(w) \\ X_2(w) \\ \vdots \\ X_3(w) \end{pmatrix} \text{ is random samples.} \quad ※_2 \ \text{The case of } w_i = \frac{1}{M} \text{ shows Uniform weights.} \right)$$

⑦ Generalised Monte Carlo approximation

This approximation to $\mathbb{E}[f(x)]$ use (*), but ...

(i) Use a different (and probably non-independent) joint distribution with PDF $\tilde{\pi}_X(x_1, x_2, \cdots, x_M)$ for $\{X_i\}$

(ii) Use non-uniform weight $\{w_i\}$

$$\left(※ \ w_i > 0 \text{ and } \sum_{i=0}^{M} w_i = 1 \right)$$

Then how to approximate by standard quadrature rules?

→ (a) select PDF as product of Dirac delta functions

$$\tilde{\pi}_M(x_1, x_2, \cdots, x_M) = \prod_{i=1}^{M} \delta(x_i - c_i)$$

(b) properly written as ...

measure : $\tilde{\mu}_M(dx_1, dx_2, \cdots, dx_M) = \prod_{i=1}^{M} \mu_{c_i}(dx_i)$

weights : $w_i = b_i$

However there is a difference between (a) and (b)

→ For given quadrature rule $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_M \end{pmatrix}$ are always same. But in (b), the value of $\mu_M(dx) = \sum_{i=1}^{M} w_i \mu_{x_i}(dx)$ will be different each time we draw from $(X_1, X_2, \cdots, X_M)$ with $\tilde{\pi}_X$

This make the case of $M \to \infty$ complicated.