Noisy Monte Carlo using Order Statistics

Artan Boriçi

University of Tirana
Department of Physics

Why Noisy Monte Carlo?

- Noisy Monte Carlo: A Monte Carlo simulation with extra noise
- Example: det A = det D(U')/det D(U) = exp(Tr Log (A))
- Log(A) difficult to evaluate
- But $Z^\dagger$log (A) Z, Z random easier task

Noisy Determinant

If $Z_j \in \{-1,1\}$ random i.i.d.

$X_j \equiv - Z_j^T \log A \, Z_j$

$\Rightarrow$ (*Central Limit Theorem*)

X_j i.i.d., moreover $X_j \sim N(\mu,\sigma)$

with $\mu \equiv E(X) = - Tr \log A$

and $\sigma^2 \equiv Var(X) = 2 \sum_{j \neq l} [Re(\log A)_{jl}]^2$

Noisy Determinant

- Bias in the noisy estimation :
  $E[\exp(-X)] = \exp(\mathrm{Tr}\log A)\exp(\sigma^2/2)$
- Obvious way to reduce the bias is to take the average of n estimations:
  $\langle X \rangle = (X_1 + X_2 + \ldots + X_n) / n$
- $\Rightarrow \sigma^2 \rightarrow \sigma^2/n$
- Expensive since $\sigma^2 \sim$ Volume

Noisy Monte Carlo Algorithms

- Kennedy-Kuti algorithm: exp(-X) ≈ 1 - X
  - Works well for small changes
  - Negative probabilities for arbitrary changes
- Kentucky algorithm (Joó-Horváth-Liu):
  - Factors determinant into fractional flavours
  - Noise becomes part of the state space
  - Low acceptance and long autocorrelations

General Idea to Eliminate Bias

- Find an estimator $Y = Y(X_1, X_2, \ldots, X_n)$ such that: $E[\exp(-Y)] = \exp(\mathrm{Tr} \log A)$
- Example: Y Gaussian with $EY = \mu + \sigma^2/2n$, $\mathrm{Var}Y = \sigma^2/n$, then $E[\exp(-Y)] = \exp(-\mu - \sigma^2/2n) \exp(\sigma^2/2n)$
- Does such a function $Y = Y(X_1, X_2, \ldots, X_n)$ exist?

Order Statistics

- Order sequence $X_1,\ldots,X_n$ such that:
  $X_{(1)} \leq \ldots X_{(k)} \ldots \leq X_{(n)}$
  $\Rightarrow X_{(k)}$ an order statistics
- Example: Median $X_{((n+1)/2)}$ (n odd)
  $EX_{((n+1)/2)} = \mu$
- Idea: For $k > (n+1)/2$: $EX_{(k)} \geq \mu$
- What is the value of k for no bias?

Order Statistics Distribution

Let $U_1,\ldots,U_m$ be uniformly distributed in $(0,1)$ and $U_{(1)} \leq \ldots\ U_{(k)} \ldots \leq U_{(m)}$

$\Rightarrow$

The marginal pdf of $U_{(k)}$ is:

$f_k(x) = x^{k-1}\ (1-x)^{n-k}\ n!/[(k-1)!/(n-k)!]$

Proof

For $U_{(k)}$ to lie in the interval $(x, x+\Delta)$ it is required that k-1 of the *Us* lie to the left of x, one lies in $(x, x+\Delta)$, and n-k lie to the right. To first order, the probability for this to happen is:

$$f_k(x) = x^{k-1}\,\Delta\,(1-x)^{n-k}\; n!/[(k-1)!\;1!\;(n-k)!]$$

where the factorial terms arise because we don't care which of the *Us* lies in the specified intervals.

Formal Proof

Step 1. Find $F_k(x) = P(U_{(k)} < x)$

Step 2. Differentiate $F_k(x)$ w.r.t. x

Step 1. Construct Binomial Variates:

$$Y_n = \omega_1 + \omega_2 + \ldots + \omega_n$$

$$\omega_i = \mathrm{I}\{U_i < x\}$$

$$\Rightarrow F_k(x) =$$

$$P(Y_n \geq k) = \sum_{j=k,\ldots,n} C(n,j)\, x^j\, (1-x)^{n-j}$$

Formal Proof

Step 2. Differentiate $F_k(x)$ w.r.t. x

$f_k(x) =$

$$= \Sigma_{j=k,\ldots,n} C(n,j)\, [jx^{j-1} (1-x)^{n-j} - (n-j)\, x^j\, (1-x)^{n-1-j}]$$

$$= \Sigma_{j=k,\ldots,n} (T_{j-1} - T_j)$$

where

$$T_j = C(n,j)\, (n-j)\, x^j\, (1-x)^{n-1-j}$$

Since $T_n = 0$ the sum telescopes down to T_{k-1}, which is the required result.

What is the $X_{(k)}$ Distribution ?

Let $X_1,\ldots,X_m$ be Normally distributed with pdf $f(x)$, expectation μ and variance σ^2

Proposition.

$$P(X_k < x) \equiv p(x) = ½ + ½ \,\mathrm{erf}\,[(x - \mu)/\sigma\sqrt{2}]$$

Theorem.

The marginal pdf of $X_{(k)}$ is:

$$f_k(x) = f(x)\, p(x)^{k-1}\, [1-p(x)]^{n-k}\, n!/[(k-1)!/(n-k)!]$$

Examples

Robust Statistics: Median

Extreme Value Theory: Maximum: $X_{(n)}$

Genralization of Central Limit Theorem for Tail Distributions.

eg. Gumbel Distribution

$$f(x) = exp(-e^{-x})$$

Other Distributions, Fréchet, Weibull

Asymptotic Theory

- Let $k(\alpha) = [n\alpha] + 1$, $0 < \alpha < 1$
  $x(\alpha)$ a quantile of the normalized distribution of X:
  - $\alpha = ½ + ½\ \mathrm{erf}[x(\alpha)/\sqrt{2}]$

  ⇒For large n: $E[X_{k(\alpha)}] = \mu + x(\alpha)\ \sigma$

  $Var[X_{k(\alpha)}] = \sigma^2/n\ 2\pi\ \alpha(1-\alpha)\ \exp[x(\alpha)^2]$
- Bias eliminated if:
  - $x(\alpha) = \sigma/n\ \pi\ \alpha(1-\alpha)\ \exp[x(\alpha)^2]$
  - For small $x(\alpha) \approx \pi\sigma/4n$

Noisy Metropolis Algorithm

- Propose a new configuration
- Compute $X_1, X_2, \ldots, X_n$ and find $k(\alpha)$
- Accept/Reject with probability $\min\{1, \exp[-X_{(k(\alpha))}]\}$

Properties of the Algorithm

- Detailed Balance is satisfied in average
- N should be large enough for the asymptotic theory to hold
- σ is unknown: use S from data
- $\Rightarrow$ error $\approx \exp[\sigma\, x(\alpha)\, (\sigma/S - 1)]$
- Use large sample size to control

Testing Algorithm

Schwinger Model on the Lattice

- $N_f = 2$ staggered fermions
- 16x16 lattice
- Bare mass $m = 0.01$
- Coupling $\beta = 5$
- Three simulations:
  - 1. Exact determinant
  - 2. Noisy determinant $n = 30$
  - 3. Noisy determinant $n = 40$

Conclusions

- A Noisy Algorithm without sign violations
- Applications:
  - Global updating when possible
  - Taking the square root for staggered simulations