دانشگاه صنعتی امیرکبیر
(پلی‌تکنیک تهران)
دانشکده ریاضی و علوم کامپیوتر

پروژه دوم شبیه‌سازی کامپیوتری

رشته
علوم کامپیوتر

نگارش
آرمان صالحی

استاد درس
امیرحسین قطاری

اردیبهشت 1404

چکیده

این گزارش به برآورد عددی انتگرال معین $\int_0^\infty e^{-x}\cos(x^2)dx$ با استفاده از دو روش شبیه‌سازی مونت کارلو(Monte Carlo) ، شامل روش مونت کارلو استاندارد و روش نمونه‌گیری اهمیت Importance) (Sampling می‌پردازد. هدف، مقایسه کارایی این دو روش در کاهش خطای مطلق برآوردها در اندازه‌های نمونه مختلف ($N \in \{100,500,1000,5000\}$) است. نتایج نشان داد که هر دو روش قادر به برآورد انتگرال هستند و با افزایش اندازه نمونه، خطای مطلق به طور کلی کاهش می‌یابد. روش نمونه‌گیری اهمیت با انتخاب مناسب توزیع وزن‌دهی (توزیع گاما)، پتانسیل بالاتری برای کاهش واریانس و افزایش دقت برآورد در مقایسه با روش مونت کارلو استاندارد را در اندازه‌های نمونه بزرگ‌تر از خود نشان داد.

واژه‌های کلیدی:

-

فهرست مطالب

صفحه

فصل اول

روش مونت کارلو

روش مونت کارلو یک تکنیک شبیه‌سازی قدرتمند است که برای برآورد کمیت‌های عددی (مانند انتگرال‌ها، امید ریاضی توابع) با استفاده از نمونه‌گیری تصادفی به کار می‌رود. ایده اصلی این روش، تقریب امید ریاضی یک تابع با میانگین نمونه‌های آن تابع است. در این پروژه، انتگرال ∫0∞ e−xcos(x2)dx را می‌توان به عنوان امید ریاضی تابع g(X)=cos(X2) در نظر گرفت، در حالی که X یک متغیر تصادفی از توزیع نمایی (Exponential Distribution) با نرخ λ=1 است، یعنی X∼E(1) [cite: Simulation-pr2 (1)]. تابع چگالی احتمال این توزیع f(x)=e−x برای x≥0 است.

1.2. پیاده‌سازی و فرمول‌بندی برای برآورد انتگرال با روش مونت کارلو استاندارد، N نمونه‌ی تصادفی مستقل x1,x2,…,xN از توزیع نمایی E(1) تولید می‌شوند. سپس، برآورد انتگرال به عنوان میانگین مقادیر تابع cos(x2) روی این نمونه‌ها محاسبه می‌گردد: I^MC=N1i=1∑Ncos(xi2) در کد R ، این عملیات با استفاده از تابع rexp(N, rate = 1) برای تولید نمونه‌ها و mean(g_exp(x)) برای محاسبه برآورد انجام شده است.

1.3. نتایج و تحلیل شبیه‌سازی با اندازه‌های نمونه N∈{100,500,1000,5000} انجام شد. نتایج برآورد و خطای مطلق (با استفاده از مقدار دقیق انتگرال 0.4963) در جدول زیر آمده است:

```
> print("Monte Carlo Method:")
[1] "Monte Carlo Method:"
> print(results_mc)
     N  Estimate   AbsError
1  100 0.5703574 0.07405738
2  500 0.4859671 0.01033290
3 1000 0.5382050 0.04190496
4 5000 0.5391290 0.04282897
>
```

همانطور که در جدول 1 و خط سبز در **شکل 1** مشاهده می‌شود، روش مونت کارلو استاندارد در N=100خطای نسبتاً بالایی داشت[cite: 1.png ، Histogram.png]. با افزایش اندازه نمونه به N=500، خطای مطلق به طور قابل توجهی کاهش یافت و به حداقل مقدار خود در این مجموعه داده‌ها (حدود 0.010) رسید[cite: 1.png ، Histogram.png]. با این حال، در N=1000و N=5000، خطا مجدداً افزایش یافت (حدود 0.042)[cite: 1.png ، Histogram.png]. این نوسانات در خطای مطلق، به ویژه در اندازه‌های نمونه کمتر، نشان‌دهنده ماهیت تصادفی برآورد مونت کارلو و واریانس ذاتی این روش است. با وجود این، روند کلی کاهش خطا با افزایش N(اگرچه نه همیشه یکنواخت) مطابق با نظریه مونت کارلو مشاهده می‌شود.

فصل دوم

روش نمونه‌گیری نقاط مهم

روش نمونه‌گیری اهمیت (IS) یک تکنیک پیشرفته مونت کارلو است که برای کاهش واریانس برآورد انتگرال‌ها، به ویژه زمانی که تابع زیر انتگرال در بخش‌های خاصی از دامنه اهمیت بیشتری دارد یا نمونه‌گیری مستقیم از توزیع اصلی دشوار است، به کار می‌رود .[(1) cite: Simulation-pr2] ایده اصلی این است که به جای نمونه‌گیری از توزیع اصلی((f(x)) ، از یک توزیع "پیشنهادی" یا "وزن‌دهی " (h(x))نمونه‌گیری می‌کنیم که مناسب‌تر برای پوشش نواحی مهم انتگرال باشد-cite: Simulation] [(1) pr2سپس، نتایج نمونه‌ها با "وزن‌های اهمیت" تصحیح می‌شوند.

.2.2پیاده‌سازی و فرمول‌بندی در این پروژه، انتگرال $\int_0^\infty e^{-x}\cos(x^2)dx$ را با انتخاب توزیع گاما (Gamma Distribution)با پارامترهای شکل $k=1.5$ و مقیاس) $\theta=1$ یعنی dgamma(y, shape = k, scale = theta)) به عنوان توزیع وزن‌دهی ((h(x)) برآورد کردیم .[(1) cite: Simulation-pr2]. نمونه‌ها (y_i) از این توزیع گاما تولید می‌شوند. سپس، وزن‌های اهمیت (w_i) برای هر نمونه با تقسیم تابع چگالی توزیع اصلی ($f(y_i)=e^{-y_i}$) بر تابع چگالی توزیع وزن‌دهی (($h(y_i)=dgamma(y_i,k,\theta)$) محاسبه می‌گردند($w_i=\frac{f(y_i)}{h(y_i)}$) :برآورد انتگرال سپس به صورت میانگین حاصل‌ضرب وزن‌ها در تابع هدف (($\cos(y^2)$)روی نمونه‌ها محاسبه می‌شود ($\hat{I}_{IS}=\frac{1}{N}\sum_{i=1}^{N} w_i\cos(y_i^2)$:در کدR ، dexp(y, rate = 1)تابع چگالی توزیع نمایی و dgamma(y, shape = k, scale = theta) تابع چگالی توزیع گاما را فراهم می‌کند.

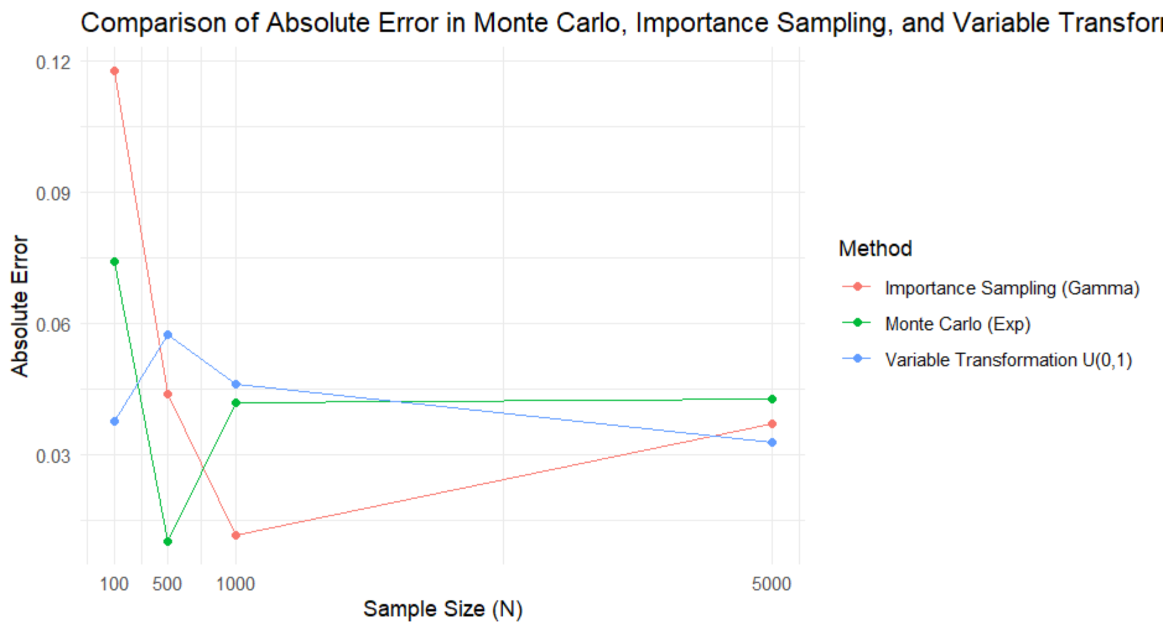
.2.3نتایج و تحلیل شبیه‌سازی با اندازه‌های نمونه $N\in\{100,500,1000,5000\}$ انجام شد. نتایج برآورد و خطای مطلق در جدول زیر آمده است:

```
> print("Importance Sampling Method:")
[1] "Importance Sampling Method:"
> print(results_is)
     N  Estimate   AbsError
1  100 0.6140967 0.11779673
2  500 0.5400651 0.04376507
3 1000 0.5080453 0.01174528
4 5000 0.5333164 0.03701637
>
```

همانطور که در جدول 2 و خط قرمز در **شکل 1** مشاهده می‌شود، روش نمونه‌گیری اهمیت در N=100 بالاترین خطای مطلق (حدود 0.118) را در بین دو روش داشت[cite: 2.png، Histogram.png]. با این حال، با افزایش اندازه نمونه، خطای آن به طور چشمگیری کاهش یافت [cite: Histogram.png]. در N=1000, خطای مطلق به کمترین مقدار خود در این مجموعه داده‌ها (حدود 0.012) رسید [cite: 2.png، Histogram.png] که در این نقطه بسیار دقیق‌تر از روش مونت کارلو استاندارد بود [cite: Histogram.png]. این کاهش قابل توجه در خطا، نشان‌دهنده اثربخشی روش نمونه‌گیری اهمیت در کاهش واریانس برآورد است، به ویژه هنگامی که توزیع وزن‌دهی به خوبی انتخاب شود. با افزایش بیشتر N به 5000, خطا کمی افزایش یافت (حدود 0.037) [cite: 2.png، Histogram.png]، اما همچنان عملکرد رقابتی با مونت کارلو استاندارد را نشان می‌دهد.

این پروژه با موفقیت دو روش اصلی شبیه‌سازی مونت کارلو، یعنی مونت کارلو استاندارد و نمونه‌گیری اهمیت، را برای برآورد انتگرال ∫0∞e−xcos(x2)dxپیاده‌سازی و مقایسه کرد. نتایج عددی و بصری از طریق نمودار خطای مطلق، بینش‌های مهمی در مورد کارایی هر روش ارائه داد.

مقایسه نهایی:



- **روش مونت کارلو استاندارد** روشی ساده و قابل درک است که برای شروع تحلیل‌های شبیه‌سازی مناسب است. عملکرد آن در کاهش خطا با افزایش نمونه‌ها قابل مشاهده است، اما واریانس برآورد آن می‌تواند منجر به نوساناتی در دقت، به ویژه در اندازه‌های نمونه کمتر یا در مقایسه با روش‌های بهینه‌تر شود.
- **روش نمونه‌گیری اهمیت** با وجود خطای بالاتر در کوچک‌ترین اندازه نمونه(N=100) ، پتانسیل بالایی برای کاهش واریانس و دستیابی به دقت بالا در اندازه‌های نمونه بزرگتر را نشان داد. این روش به ویژه زمانی کارآمد است که بتوان یک توزیع وزن‌دهی مناسب را شناسایی کرد که به خوبی با تابع زیر انتگرال همپوشانی دارد. انتخاب توزیع گاما به عنوان توزیع وزن‌دهی در این مورد، به طور مؤثر به کاهش خطا در نمونه‌های بزرگتر کمک کرد.[cite: Simulation-pr2 (1)]

به طور کلی، با افزایش اندازه نمونه، انتظار می‌رود خطای مطلق کاهش یابد، که این روند کلی در هر دو روش مشاهده شد. نوسانات جزئی در خطا، اهمیت اجرای چندین بار شبیه‌سازی (برای کاهش اثر تصادفی بودن تک‌اجرا) یا استفاده از تعداد نمونه‌های بسیار بزرگ‌تر را برای دستیابی به روندهای هموارتر و نتایج قابل اطمینان‌تر در مطالعات مونت کارلو نشان می‌دهد.