



وزارت علوم تحقیقات و فناوری
دانشگاه فردوسی
دانشکده علوم ریاضی

پایان نامه جهت اخذ درجه کارشناسی ارشد آمار ریاضی

هموار سازی مونت کارلو برای سری های زمانی غیر خطی

استاد راهنما
دکتر حسینعلی نیرومند

استاد مشاور
دکتر حسن دوستی

نگارنده
یاسر سنچولی

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

وزارت علوم تحقیقات و فن آوری

دانشگاه فردوسی

دانشکده علوم ریاضی

پایان نامه جهت اخذ درجه کارشناسی ارشد آمار ریاضی

هموار سازی مونت کارلو برای سری های زمانی غیر خطی

استاد راهنما

دکتر حسینعلی نیرومند

استاد مشاور

دکتر حسن دوستی

نگارنده

یاسر سنچولی

تقدیم بہ ہمہ عزیزانم

سپاسگزاری

زندگی صحنه یکتای هنرمندی ماست
هر کسی نغمه خود خواند و از صحنه رود

صحنه پیوسته بجاست
خرم آن نغمه که مردم بسیار ندیده یاد.

چکیده

هموار سازی مونت کارلو برای سری های زمانی غیر خطی

نگارنده

یاسر سنجولی

ما در این پایان نامه، روش هایی برای انجام محاسبات هموار سازی در مدل های عمومی فضای حالت ارائه می کنیم. این روش ها به ارائه مختصری از توابع هموار کننده و سیر تکاملی آن ها در طول زمان با استفاده از نظریه های نمونه گیری و باز نمونه گیری ترتیبی استناد می کند. بخصوص، تکنیک های جدیدی برای تولید یافته های نمونه از دنباله های حالت های گذشته ارائه شده اند. این تکنیک ها در یک فرآیند هموار سازی پیشرو و پسرو که می تواند به عنوان نقطه ی مقابل غیر خطی و غیر نرمال هموارکننده های شبیه سازی استاندارد کالمن^۱ در حالت نرمال خطی مدنظر قرار بگیرد، انجام شده اند. همگرایی میانگین مربعات خطای مسیر های هموار شده اثبات شده است و اعتبار روش های پیشنهادی ما را تایید می کند.

^۱Calman

فهرست

۱	پیشگفتار
۳	۱ مونته کارلوی زنجیر مارکف
۴	۱.۱ انتگرال گیری مونته کارلو
۴	۱.۱.۱ مقدمه ای بر زنجیر های مارکف
۸	۲.۱ الگوریتم هستینگ-متروپلیس
۱۰	۱.۲.۱ نمونه گیری هستینگ-متروپلیس به عنوان یک زنجیر مارکف
۱۱	۲.۲.۱ سوختن نمونه گیر
۱۲	۳.۲.۱ انتخاب یک توزیع پیشنهادی مناسب
۱۴	۴.۲.۱ خودهمبستگی و تورم اندازه ی نمونه
۱۵	۳.۱ نمونه بردار گیبس
۱۶	۲ هموارسازی و پالایش ذره
۱۶	۱.۲ پالایش و درستمایی حاشیه ای
۱۹	۲.۲ هموارسازی
۲۰	۱.۲.۲ روش بازگشتی پسرو-پیشرو
۲۱	۲.۲.۲ فرمول ۲- فیلتر تعمیم یافته
۲۲	۳.۲ روش های مونته کارلوی ترتیبی
۲۳	۱.۳.۲ مفاهیم پایه ای روش های مونته کارلو
۲۴	۲.۳.۲ نمونه گیری نقاط مهم
۲۶	۳.۳.۲ نمونه گیری نقاط مهم ترتیبی
۲۸	۴.۳.۲ باز نمونه گیری
۲۹	۴.۲ پالایش ذره
۲۹	۱.۴.۲ استفاده از روش های مونته کارلوی ترتیبی برای پالایش
۳۱	۵.۲ هموارسازی ذره
۳۲	۱.۵.۲ تقریب با تاخیر ثابت
۳۲	۲.۵.۲ هموارسازی پسرو-پالایش پیشرو
۳۴	۳.۵.۲ فرمول ۲- فیلتر تعمیم یافته

۳۶	۳	پالایش از طریق شبیه سازی
۳۷	۱.۳	صافی ذره
۳۷	۱.۱.۳	تعریف صافی ذره
۳۹	۲.۳	نمونه گیری چگالی پیش بینی تجربی
۳۹	۱.۲.۳	نمونه گیری - باز نمونه گیری نقاط مهم
۴۰	۲.۲.۳	انطباق
۴۰	۳.۲.۳	عدم پذیرش و نمونه گیری مونت کارلوی زنجیر مارکف
۴۱	۴.۲.۳	نقطه ضعف صافی ذره
۴۲	۳.۳	متغیر کمکی
۴۲	۱.۳.۳	مطالب پایه ای
۴۳	۲.۳.۳	نمونه گیری- باز نمونه گیری نقاط مهم بر پایه ی متغیر کمکی

۴۶	۴	هموار سازی مونت کارلو برای سری های زمانی غیر خطی
۴۹	۱.۴	پالایش و هموار سازی با استفاده از نمونه گیری ترتیبی نقاط مهم
۵۰	۲.۴	هموار سازی با استفاده از شبیه سازی پسرو
۵۸	۳.۴	خصوصیات مدل
۶۰	۴.۴	پالایش و هموار سازی
۶۱	۵.۴	نتایج

۶۹ منابع

۷۳ واژه نامه ی انگلیسی به فارسی

لیست تصاویر

۱۰	نتیجه ۵۰۰ مقدار اول	۱.۱
۱۰	یک نتیجه ی نمونه گیری هستینگ-متروپولیس	۲.۱
۱۴	نتایج اجرای نمونه بردار تحت ۲ توزیع پیشنهادی χ^2_{10} و χ^2_2	۳.۱
۱.۴	داده های شبیه سازی شده از مدل سری زمانی غیر خطی. (دنباله (a) حالت های پنهان x_t ؛ دنباله (b) داده های مشاهده شده y_t)	
۵۳	برآورد های چگالی پالایش از خروجی های فیلتر ذره	۲.۴
۵۳	مسیر های هموارسازی انتخاب شده از $p(x_{1:100} y_{1:100})$. (حالت های درست شبیه سازی شده با '* نشان داده شده است.)	۳.۴
۵۴	برآوردهای نمودار ستونی چگالی های هموارسازی $p(x_t y_{1:100})$. (حالت های درست شبیه سازی شده با '* نشان داده شده است.)	۴.۴
۵۴	برآورد چگالی هسته برای $p(x_{3:4} y_{1:100})$	۵.۴
۵۵	نمودار پراکندگی نقاط انتخاب شده برای $p(x_{3:4} y_{1:100})$	۶.۴
۵۶	برآورد چگالی هسته برای $p(x_{27:28} y_{1:100})$	۷.۴
۵۶	نمودار پراکندگی نقاط انتخاب شده برای $p(x_{27:28} y_{1:100})$	۸.۴
۵۷	برآورد چگالی هسته برای $p(x_{27:28} y_{1:100})$ با استفاده از $\sigma_v^2 = 1$ و $\sigma_w^2 = 9$	۹.۴
۵۷	نمودار پراکندگی نقاط انتخاب شده برای $p(x_{50:51} y_{1:100})$	۱۰.۴
۶۲	داده های سخنرانی	۱۱.۴
۶۲	داده های سخنرانی پرسر و صدا	۱۲.۴
۶۳	داده های سخنرانی پرسر و صدا؛ $t=800, \dots, 1000$	۱۳.۴
۶۴	میانگین فیلتر شده و $\frac{5}{45}$ از صدک ها برای برآورد های ضرایب همبستگی جزئی مختلف-زمان	۱۴.۴
۶۴	سخنرانی پالایش شده، میانگین پسین برآورد شده (-)، داده های پرسروصدا (...)	۱۵.۴
۶۵	۱۰ یافته از فرآیند سیگنال هموارسازی شده	۱۶.۴
۶۵	۴ یافته از چگالی هموارسازی برای فرآیند سیگنال با داده های پرسروصدا که با خط نقطه چین نشان داده شده است.	۱۷.۴
۶۵	۱۰ میانگین یافته از چگالی هموارسازی با داده های پرسروصدا که با خط نقطه چین نشان داده شده است.	۱۸.۴
۶۶	۱۰ یافته از چگالی هموارسازی برای برآورد های ضرایب همبستگی جزئی با زمان متغیر	۱۹.۴

- ۲۰.۴ ۱۰ یافته از چگالی هموارسازی برای برآورد های ضرایب همبستگی جزئی با زمان متغیر با استفاده از روش هموارسازی شبیه سازی پیشنهاد شده، (N=۲۰۰۰) ۶۷
- ۲۱.۴ ۱۰ یافته از چگالی هموارسازی برای برآورد های ضرایب همبستگی جزئی با زمان متغیر با استفاده از روش بر پایه ی مسیر استاندارد، (N=۲۰۰۰) ۶۷
- ۲۲.۴ ۱۰ یافته از چگالی هموارسازی برای برآورد های ضرایب همبستگی جزئی با زمان متغیر با استفاده از روش هموارسازی شبیه سازی پیشنهاد شده، (N=۵۰۰) ۶۸
- ۲۳.۴ ۱۰ یافته از چگالی هموارسازی برای برآورد های ضرایب همبستگی جزئی با زمان متغیر با استفاده از روش بر پایه ی مسیر استاندارد، (N=۵۰۰) ۶۸

پیشگفتار

یک محدودیت عمده در استفاده گسترده ی روش های بیزی، بدست آوردن توزیع پسین است که اغلب لازم است از یک تابع چند متغیره انتگرال گرفته شود. از آنجایی که این محاسبه بسیار سخت می باشد، می توان روش های کوتاه انتگرال گیری مستقیم را پیشنهاد نمود. ما در اینجا روی مونت کارلوی زنجیر مارکف تمرکز می کنیم تا بتوانیم نمونه های مستقیمی را از توزیع های پیچیده ی مورد نظر شبیه سازی کنیم.

روش های مونت کارلوی زنجیر مارکف بسیار پرکاربرد می باشند زیرا در این روش ها از مقادیر نمونه ی قبلی برای تولید مقادیر تصادفی نمونه ی بعدی استفاده شده و یک زنجیر مارکف تولید می گردد. وقوع یک روش منحصر به فرد مونت کارلو زنجیر مارکف در اوایل سال ۱۹۹۰ (نمونه گیر گیس) کاربرد وسیعی برای کلاس گسترده ای از مسائل بیزی پیدا کرد.

ما در فصل اول به معرفی انتگرال مونت کارلو پرداخته و مقدمه ای از زنجیر مارکف را ارائه می کنیم. سپس با بیان الگوریتم هستینگ-متروپلیس، به رابطه ی این الگوریتم با تولید دنباله مارکف اشاره می نماییم.

اما در فصل دوم به معرفی پالایش و هموارسازی می پردازیم. از آنجایی که مسائل برآورد بهینه برای مدل های فضای حالت غیر خطی اغلب راه حل های تحلیلی را نمی پذیرند، روش های پالایش ذره بعنوان یک کلاس رایج از الگوریتم های حل عددی مسائل برآورد، مورد استفاده قرار گرفت و هم اکنون نیز بصورت روز مره در زمینه های مختلفی مانند کامپیوتر، اقتصادسنجی، رباتیک و ناوبری مورد استفاده قرار می گیرد.

روش های پالایش ذره، مجموعه ای از روش های قدرتمند و انعطاف پذیر بر پایه شبیه سازی می باشند که نمونه های بصورت تقریبی توزیع شده مطابق توزیع های پسین $p(x_{1:n}|y_{1:n})$ را ارائه می نمایند و کمک به محاسبه تقریبی $p(y_{1:n})$ می کنند.

در فصل سوم یک سری زمانی y_t که $t = 1, 2, \dots, n$ را وقتی که بطور شرطی مستقل از

یک حالت مشاهده نشده α_t می باشد را مدل سازی می کنیم. ما این کار را با برآورد چگالی $f(\alpha_t|y_1, y_2, \dots, y_t) = f(\alpha_t|Y_t); t = 1, 2, \dots, n$ که به سختی محاسبه می شود، انجام می دهیم. همچنین صورت های پارامتری را برای هر دو چگالی اندازه $f(y_t|\alpha_t)$ و چگالی انتقال حالت $f(\alpha_{t+1}|\alpha_t)$ در نظر می گیریم. همچنین روش های انجام پالایش بوسیله شبیه سازی را مد نظر قرار می دهیم و مفاهیم صافی ذره و متغیر کمکی را بیان می کنیم.

در فصل چهارم نیز به ارائه چند مدل برای انجام هموارسازی مورد نظر می پردازیم و چند مثال برای روشن شدن مطلب ارائه می کنیم. سپس هموارسازی را با استفاده از نمونه گیری ترتیبی نقاط مهم و همچنین شبیه سازی پسرو انجام می دهیم و با استفاده از نتایج بدست آمده، به معرفی بهترین مدل ممکن برای چنین روشی اقدام می کنیم. با توجه به نتایج بدست آمده، مدل معرفی شده بهترین جواب را در امر هموارسازی به ما ارائه می کند.

فصل ۱

مونت کارلوی زنجیر مارکف

مقدمه

یک محدودیت عمده در استفاده گسترده ی روش های بیزی، بدست آوردن توزیع پسین است که اغلب لازم است از یک تابع چند متغیره انتگرال گرفته شود. از آنجایی که این محاسبه بسیار سخت می باشد، می توان روش های کوتاه انتگرال گیری مستقیم را پیشنهاد نمود.^۱ (اسمیت^۱ ۱۹۹۱) ما در اینجا روی مونت کارلوی زنجیر مارکف^۲ تمرکز می کنیم تا بتوانیم نمونه های مستقیمی را از توزیع های پیچیده ی مورد نظر شبیه سازی کنیم.

روش های مونت کارلوی زنجیر مارکف بسیار پرکاربرد می باشند زیرا در این روش ها از مقادیر نمونه ی قبلی برای تولید مقادیر تصادفی نمونه ی بعدی استفاده شده و یک زنجیر مارکف تولید می گردد. وقوع یک روش منحصر به فرد مونت کارلو زنجیر مارکف در اوایل سال ۱۹۹۰ (نمونه گیر گیبس^۳) کاربرد وسیعی برای کلاس گسترده ای از مسائل بیزی پیدا کرد.

تلاش فیزیک دانان برای محاسبه ی انتگرال های پیچیده، با معرفی آن ها به عنوان امید ریاضی برای برخی از توزیع ها و سپس تقریب این امید ریاضی ها بوسیله ی انتخاب نمونه از آن توزیع ها صورت می گرفت.

^۱ Smith

^۲ Markov Chain Monte Carlo (MCMC)

^۳ Gibbs sampler

۱.۱ انتگرال گیری مونت کارلو

مبدا روش مونت کارلو روشی بود که توسط فیزیک دانان برای محاسبه ی انتگرال ها با استفاده از تولید اعداد تصادفی ایجاد شده بود. فرض کنیم که می خواهیم انتگرال پیچیده ی زیر را بدست بیاوریم

$$\int_a^b h(x)dx \quad (1.1)$$

اگر بتوانیم $h(x)$ را به حاصلضرب یک تابع مانند $f(x)$ و یک تابع چگالی احتمال مانند $p(x)$ که روی (a, b) تعریف شده است تجزیه کنیم، خواهیم داشت

$$\int_a^b h(x)dx = \int_a^b f(x)p(x)dx = E_{p(x)}[f(x)] \quad (2.1)$$

بنابراین انتگرال اولیه را می توان به صورت امید ریاضی $f(x)$ روی چگالی $p(x)$ بیان کرد. حال اگر ما تعداد زیادی از متغیرهای تصادفی مانند X_1, X_2, \dots, X_n را از چگالی $p(x)$ انتخاب کنیم، آنگاه

$$\int_a^b h(x)dx = E_{p(x)}[f(x)] \simeq \frac{1}{n} \sum_{i=1}^n f(x_i) \quad (3.1)$$

که عبارت فوق به انتگرال گیری مونت کارلو موسوم می باشد. انتگرال مونت کارلو می تواند برای تخمین توزیع های پسین (حاشیه ای پسین) مورد نیاز در تحلیل های بیزی نیز استفاده شود. انتگرال $\int f(y|x)p(x)dx$ را در نظر بگیرید. آن را به صورت زیر تقریب می زنیم

$$\hat{I}(y) = \frac{1}{n} \sum_{i=1}^n f(y|x_i) \quad (4.1)$$

که x_i از چگالی $p(x)$ انتخاب شده اند. خطای استاندارد برآورد شده ی مونت کارلو عبارتست از

$$SE^*[\hat{I}(y)] = \frac{1}{n} \left(\frac{1}{n-1} \sum_{i=1}^n (f(y|x_i) - \hat{I}(y))^2 \right)$$

۱.۱.۱ مقدمه ای بر زنجیرهای مارکف

فرض می کنیم که X_t یک متغیر تصادفی در زمان t باشد و همچنین فرض می کنیم که فضای حالت نیز دامنه ی مقادیر ممکن X باشد.

متغیر تصادفی یک فرآیند مارکف است اگر احتمال های انتقال بین مقدار های مختلف در فضای حالت فقط به حالت فعلی متغیر تصادفی بستگی داشته باشد یعنی

$$Pr(X_{t+1} = s_j | X_0 = s_k, \dots, X_t = s_i) = Pr(X_{t+1} = s_j | X_t = s_i) \quad (5.1)$$

برای یک متغیر تصادفی مارکف تنها اطلاعات حالت فعلی برای پیش بینی حالت آینده متغیر تصادفی نیاز است و دانستن مقدار حالت های گذشته در احتمال انتقال تغییری ایجاد نمی کند. یک زنجیر مارکف به دنباله ای از متغیر های تصادفی (X_0, X_1, \dots, X_n) اشاره دارد که توسط فرآیند مارکف تولید شده اند. یک زنجیر توسط احتمال های انتقالش معرفی می شود. بنابراین $P(i, j) = P(i \rightarrow j)$ احتمال این است که فرآیند از حالت s_i به s_j بوسیله یک گام واحد انتقال می یابد.

$$P(i, j) = P(i \rightarrow j) = P(X_{t+1} = s_j | X_t = s_i) \quad (6.1)$$

از علامت $P(i \rightarrow j)$ برای یک انتقال از i به j استفاده می کنیم. فرض کنید $\pi_j(t) = P(X_t = s_j)$ برابر با احتمال این باشد که زنجیر در حالت j در زمان t است. همچنین فرض می کنیم $\pi(t)$ بردار سطری با احتمال های فضای حالت در گام t باشد. ما زنجیر را با مشخص کردن یک بردار آغاز کننده $\pi(0)$ شروع می کنیم. اغلب همه ی عناصر $\pi(0)$ صفر می باشند، به جز برای یک عنصر منحصر به فرد 1 که مطابق با شروع فرآیند در آن حالت خاص است.

احتمال اینکه زنجیر مقدار حالت s_i در زمان $t + 1$ داشته باشد، توسط معادله چپمن-کولموگروف بدست می آید که حاصل جمع احتمال های حالتی خاص در گام فعلی و احتمال انتقال از این حالت به حالت s_i است.

$$\begin{aligned} \pi_i(t + 1) &= P(X_{t+1} = s_i) \\ &= \sum_k P(X_{t+1} = s_i | X_t = s_k) \cdot P(X_t = s_k) \\ &= \sum_k P(k \rightarrow i) \cdot \pi_k(t) = \sum_k P(k, i) \cdot \pi_k(t) \end{aligned} \quad (7.1)$$

تکرار پی در پی معادله ی چپمن-کولموگروف، سیر تکاملی زنجیر را توضیح می دهد. ما می توانیم معادله چپمن-کولموگروف را بطور خلاصه در یک ماتریس بنویسیم. ماتریس احتمال انتقال P را به عنوان ماتریسی که i و j امین عنصر آن $p(i, j)$ است، تعریف می کنیم و $p(i, j)$ نیز احتمال انتقال از حالت i به حالت j است. بنابراین معادله ی چپمن-کولموگروف می شود

$$\pi(t + 1) = \pi(t) \cdot P \quad (8.1)$$

با استفاده از ماتریس P مشاهده می کنیم که چگونه به سرعت فرمول چپمن-کولموگروف تکرار می شود

$$\pi(t) = \pi(t-1).P = (\pi(t-2)P)P = \pi(t-2)P^2 \quad (9.1)$$

ادامه ی این روش نشان می دهد که

$$\pi(t) = \pi(0).P^t \quad (10.1)$$

تعریف احتمال انتقال گام n -ام (P_{ij}^n) به عنوان احتمال اینکه فرآیند در مرحله معلوم j است که نقطه ی شروع آن در مرحله i یعنی n مرحله قبل است، می باشد بنابراین

$$P_{ij}^n = P(X_{t+n} = s_j | X_t = s_i) \quad (11.1)$$

چنین بر می آید که P_{ij}^n درست عضو i, j ام P^n می باشد.

در نهایت یک زنجیر مارکف را ساده نشدنی گوئیم، هرگاه همه ی حالت ها با هم در ارتباط باشند و یک عضو دلخواه بتواند به هر عضو دیگری برود. (هر چند ممکن است بیشتر از یک گام شود). همچنین یک زنجیر مارکف، غیر دوره ای گفته می شود اگر تعداد گام های مورد نیاز برای حرکت بین دو حالت مضربی از تعدادی عدد نباشد. به عبارت دیگر زنجیر مجبور نباشد در چرخه ای با طول ثابت بین حالت های معین باشد.

مثال ۱. فرض کنید که فضای حالتی به صورت (بارانی، آفتابی، ابری) داشته باشیم و آب و

هوا از یک زنجیر مارکف پیروی کند. بنابراین احتمال هوای فردا به هوای امروز وابسته است.

فرض می کنیم که احتمال انتقال که امروز بارانی داده شده باشد به صورت زیر باشد:

$$Pr(\text{بارانی امروز} | \text{بارانی فردا}) = 0.5$$

$$Pr(\text{بارانی امروز} | \text{آفتابی فردا}) = 0.25$$

$$Pr(\text{بارانی امروز} | \text{ابری فردا}) = 0.25$$

سطر اول ماتریس احتمال انتقال به صورت $(0.5, 0.25, 0.25)$ می باشد. فرض می کنیم که بقیه

ی ماتریس به صورت زیر داده شده باشد:

$$P = \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{pmatrix}$$

۱. مونت کارلوی زنجیر مارکف

توجه کنید که زنجیر مارکف، ساده نشدنی است یعنی همه ی حالت ها با هم در ارتباط می باشند.

حال اگر فرض کنیم امروز هوا آفتابی باشد، هوای ۲ روز آینده چگونه خواهد بود؟

در اینجا $\pi(0) = (0 \ 1 \ 0)$ می باشد و

$$\pi(2) = \pi(0)P^2 = (0/375 \ 0/25 \ 0/375)$$

$$\pi(7) = \pi(0)P^7 = (0/4 \ 0/2 \ 0/4)$$

بر عکس، فرض می کنیم امروز بارانی باشد، آن گاه $\pi(0) = (1 \ 0 \ 0)$ ، بنابراین هوای مورد انتظار

به صورت زیر است:

$$\pi(2) = (0/4375 \ 0/1875 \ 0/375)$$

$$\pi(7) = (0/4 \ 0/2 \ 0/4)$$

توجه داشته باشید که پس از مقدار زمان کافی، هوای مورد انتظار مستقل از مقدار شروع کننده

است. به عبارت دیگر زنجیر به حالت ساکن رسیده است و مقادیر احتمال مستقل از مقدار شروع

کننده می باشد.

همان طور که در بالا اشاره شد، حالت ساکن π^* ، حالتی است که بردارهای احتمال در هر مرحله

معین داده شده، مستقل از شرایط آغازی است.

بخش ساکن ایجاب می کند

$$\pi^* = \pi^* \cdot P \quad (12.1)$$

به عبارت دیگر، π^* مقادیر ویژه ی P با $\lambda = 1$ می باشد.

شرط های یک بخش ساکن این است که زنجیر ساده نشدنی و غیر تناوبی باشد. وقتی یک زنجیر

تناوبی است می تواند در یک مسیر مشخص بین حالت ها تناوب کند و بنابراین هرگز یک بخش

ساکن را نتیجه نمی دهد. (در حقیقت این مسیر گردش، بخش ساکن این زنجیر است.)

کمی استدلال نشان خواهد داد که اگر P مقادیر ویژه ای برابر با ۱ - نداشته باشد، آن گاه غیر

تناوبی است.

شرط کافی برای یک حالت ساکن منحصر به فرد این است که معادلات تعادل تفصیلی برقرار باشند

(به ازای هر i و j).

$$P(j \rightarrow k)\pi_j^* = P(k \rightarrow j)\pi_k^* \quad (13.1)$$

اگر معادلات فوق برقرار باشند به ازای هر i و j زنجیر مارکف را برگشت پذیر و معادله ی فوق را

شرط برگشت پذیری گوئیم.

۲.۱ الگوریتم هستینگ-متروپلیس

یک مشکل در کاربرد انتگرال گیری مونت کارلو، نمونه گیری از توزیع های احتمال پیچیده ی $p(x)$ می باشد. تلاش برای حل این مشکل، ریشه های روش های مونت کارلوی زنجیر مارکف می باشند.

فرض کنید که هدف ما بدست آوردن نمونه هایی از توزیع $p(x)$ باشد و $p(x) = (f(x)/k)$ که در آن k یک ثابت نرمال ساز است که می تواند مجهول باشد و بدست آوردن آن نیز دشوار باشد. الگوریتم متروپلیس^۴ یک دنباله از نمونه های توزیع فوق تولید می کند که در شرایط زیر صدق می کنند:

۱- با هر مقدار آغازی θ_0 که در $f(\theta_0) > 0$ صدق می کند، شروع شود.

۲- با استفاده از مقدار فعلی θ ، یک نقطه ی داوطلب θ^* را از توزیع جهش دار $q(\theta_1, \theta_2)$ نمونه گیری می کنیم، که $q(\theta_1, \theta_2)$ برابر با احتمال بازگشت مقدار معلوم θ_2 به مقدار گذشته ی θ_1 است. تنها محدودیت روی چگالی جهش دار در الگوریتم متروپلیس این است که متقارن باشد یعنی

$$q(\theta_2, \theta_1) = q(\theta_1, \theta_2) \quad (14.1)$$

۳- با معلوم بودن نقطه ی داوطلب θ^* ، نسبت چگالی در نقطه ی داوطلب (θ^*) و نقطه ی فعلی (θ_{t-1}) را محاسبه می کنیم:

$$\alpha = \frac{p(\theta^*)}{p(\theta_{t-1})} = \frac{f(\theta^*)}{f(\theta_{t-1})} \quad (15.1)$$

۴- اگر جهش چگالی را افزایش دهد، $(\alpha > 1)$ ، نقطه ی داوطلب را می پذیریم و قرار می دهیم $(\theta_t = \theta^*)$ و به مرحله ی ۲ می رویم. اما اگر جهش چگالی را کاهش داد، $(\alpha < 1)$ ، آنگاه با احتمال α نقطه ی داوطلب را می پذیریم، در غیر این صورت $(\alpha = 1)$ آن را رد کرده و به مرحله ۲ می رویم.

ما می توانیم نمونه گیری متروپلیس را با محاسبه ی عبارت زیر خلاصه نماییم:

$$\alpha = \min\left(\frac{f(\theta^*)}{f(\theta_{t-1})}, 1\right) \quad (16.1)$$

و بنابراین نقطه ی داوطلب را با احتمال α می پذیریم.

با این روش یک زنجیر مارکف $(\theta_0, \theta_1, \dots, \theta_k, \dots)$ تولید می شود بطوری که احتمال گذار از θ_t

^۴Metropolis

به θ_{t+1} فقط به θ_t بستگی دارد و به $(\theta_0, \dots, \theta_{t-1})$ بستگی ندارد.

هستینگ^۵ در سال ۱۹۷۰ با استفاده از تابع احتمال گذار دلخواه $q(\theta_1, \theta_2) = Pr(\theta_1 \rightarrow \theta_2)$ الگوریتم متروپلیس را تعمیم داد و احتمال قبولی نقطه ی داوطلب را به صورت زیر قرار داد :

$$\alpha = \min\left(\frac{f(\theta^*)q(\theta^*, \theta_{t-1})}{f(\theta_{t-1})q(\theta_{t-1}, \theta^*)}, 1\right) \quad (17.1)$$

این الگوریتم را هستینگ-متروپلیس می نامیم.

فرض اینکه توزیع پیشنهادی متقارن باشد، الگوریتم متروپلیس اصلی را نتیجه می دهد.

مثال ۲. توزیع χ^2 با پارامتر مقیاس وارون زیر را در نظر بگیرید:

$$p(\theta) = C.\theta^{-n/2}.\exp\left(\frac{-a}{2\theta}\right)$$

و فرض کنید که می خواهیم نمونه گیری از این توزیع با $n=5$ درجه آزادی و پارامتر $a=4$ را با استفاده از الگوریتم متروپلیس شبیه سازی کنیم.

همچنین فرض می کنیم که توزیع تولید کننده ی عضو داوطلب خود را توزیع یکنواخت روی $(0, 100)$ در نظر گرفته ایم. به وضوح، احتمال توده ی داده های بالای ۱۰۰ وجود دارد، اما فرض می کنیم که این احتمال به اندازه ی کافی کوچک باشد که بتوان آن را نادیده گرفت.

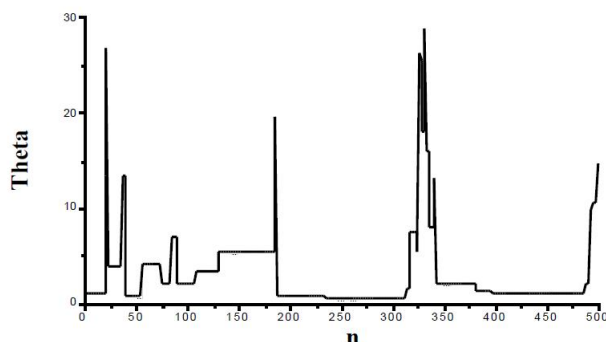
حال الگوریتم را اجرا می کنیم. به عنوان مقدار شروع کننده $\theta_0 = 1$ را در نظر می گیریم و فرض می کنیم که توزیع یکنواخت، مقدار داوطلب $\theta^* = 39/82$ را می دهد. بنابراین

$$\alpha = \min\left(\frac{f(\theta^*)}{f(\theta_{t-1})}, 1\right) = \min\left(\frac{(39/82)^{-2/5}.\exp(-\frac{2}{39/82})}{(1)^{-2/5}}.\exp(-\frac{2}{2/1}), 1\right) = 0.0007$$

چون در این حالت $\alpha < 1$ می باشد، مقدار θ^* با احتمال 0.0007 پذیرفته می شود. بنابراین ما بطور تصادفی، U را از توزیع یکنواخت $(0, 1)$ انتخاب کرده و θ^* را اگر $U \leq \alpha$ باشد، می پذیریم. در این حالت عضو داوطلب رد می شود و مقدار داوطلب دیگری را از توزیع پیشنهادی انتخاب و طبق بالا ادامه می دهیم. نتیجه ۵۰۰ مقدار اول θ^* در شکل زیر نشان داده شده است. توجه داشته باشید که دوره های طولانی مسطحی در شکل وجود دارد (مطابق با همه مقادیر θ^* که رد شده اند). چنین زنجیر هایی را بطور ضعیف آمیخته می گویند.

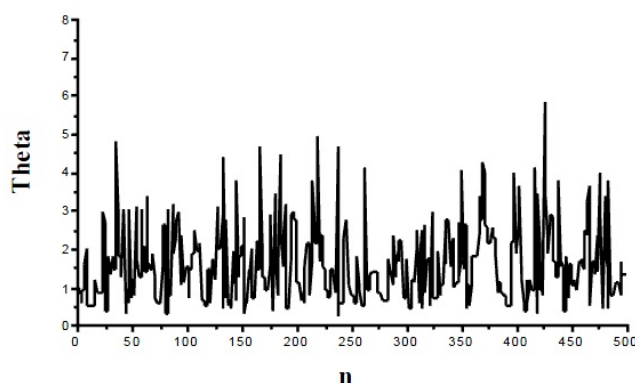
در مقابل فرض می کنیم که توزیع پیشنهادی ما χ^2_1 باشد، آن گاه توزیع داوطلب متقارن نمی باشد

^۵Hasting



شکل ۱.۱: نتیجه ۵۰۰ مقدار اول

و می بایست از الگوریتم هستینگ-متروپلیس استفاده کنیم. در این حالت یک نتیجه ی نمونه گیری هستینگ-متروپلیس در شکل زیر نشان داده شده است. توجه داریم که سری زمانی در شکل زیر شبیه به اغتشاش خالص به نظر می آید و به چنین زنجیر هایی بطور خوب آمیخته می گویند.



شکل ۲.۱: یک نتیجه ی نمونه گیری هستینگ-متروپلیس

۱.۲.۱ نمونه گیری هستینگ-متروپلیس به عنوان یک زنجیر مارکف

برای نشان دادن اینکه نمونه گیری هستینگ-متروپلیس یک زنجیر مارکف تولید می کند، کافی است نشان دهیم که هسته ی گذار هستینگ-متروپلیس در معادلات تفصیلی صدق می کند. تحت الگوریتم هستینگ-متروپلیس از $q(x, y) = Pr(x \rightarrow y|q)$ نمونه گیری می کنیم و انتقال را با احتمال $\alpha(x, y)$ می پذیریم، بنابراین هسته ی گذار به صورت زیر خواهد بود:

$$Pr(x \rightarrow y) = q(x, y)\alpha(x, y) = q(x, y)\min\left(\frac{p(y)q(y, x)}{p(x)q(x, y)}, 1\right) \quad (18.1)$$