

«به نام خدا»

پنجشنبه ۱۱

تیر

July 2015 2

۱۵ رمضان ۱۴۳۶

۹۹۵۲۲۲۲۱

امیر رضا ولیپشته

تقریب ۱۰

① مادر این سوال یک سری جدول احتمال داریم که P_t از جدول با توجه به اینکه از چه روشی استفاده می کنیم بهار ۱۰۰۰۰ نمونه می گیریم احتمال مطلوب خود را مناسب می کنیم

روش likelihood: دلیل sort اول این است که اول Parent ها به child ها تقسیم می شوند

در این روش برخلاف روش Rejection Prior که نمونه را پس داریم و به تریای evidence را چک می کنیم در این جا ما زمانی که می خواهیم نمونه برداری کنیم با توجه به اینکه

آن متغیر در evidence هست یا نه تصمیم می گیریم که در کل دو حالت داریم ① در evidence ها باشد که در این صورت مقدار آن را همان مقدار evidence می گذاریم

② اگر در evidence نباشد در این صورت مانند روش های قبل با توجه به تابع sample-vertex آن را نمونه برداری می کنیم

حال باتوجه به اینکه از نمونه برداریم یا از evidence
ها مقدارش را fix کنیم به هر نمود یک وزن اختصاص می دهیم

به ازای هر متغیر در sample

① اگر نمونه برداری نشده باشد به آن در sample
ضرب می شود

② اگر نمونه برداری نشده باشد به وزنی که برابر با احتمال
آن متغیر باتوجه به value ها تا آن لحظه است از CPT آن
متغیر با استفاده از تابع findrow پیدا شده و در وزن
کلی sample ضرب می شود

دلیل انجام این کار و جمع کردن وزن ها به جای نمونه sample
ها این است که توزیع یکنواخت شود و بتوان از این نمونه
برداری برای به دست آوردن احتمالات شرطی دیگر نیز استفاده

کرده در آخر باتوجه به اینکه queue در sample درک دارد flag

ایا صفر می شود good ها شماره را احتمال می بینیم
sample

۲) Gibbs sampling

ابتدا درباره آرایین روش هم topological sort انجام می دهیم

تا ابتدا از Patent ها ر به از λ ها نمونه برداری کنیم

در این روش نمونه برداری به این شکل است که ابتدا اگر متغیر
موجود در sample در evidence ها با λ fix می شود

آزاد می شود به شکل کاملاً random با احتمال ۵۰٪ (در صورتی که
false یا true set می شود)

سپس ما در یک تعداد iteration این نمونه را به دور می بچرخیم
به این شکل که هر بار یکی از متغیرهای غیر evidence را

سفید می کنیم و از آن با توجه به value های دیگر نمونه برداری
می کنیم حال هر چه تعداد این کار بیشتر باشد بهبود بیشتری
حاصل می شود

در واقع تفاوت روشی gibbs و likelihood اینهاست که ما
گاهی در likelihood نمونه های دارای احتمال خیلی کم (درون کره) داریم

ولی اینها احتمال وقوع نمونه بالا است زیرا با توجه
به Evidence ها داریم نمونه را به طور منصفانه و در واقع
نمونه هایی را حاصل می کنیم که احتمال وقوع Evidence ها در آنها
بالا است نه اینکه صرفاً در آنها برقرار اند.

در آخر ~~good-samples~~ ^{sample} ها را می شماریم

اگر $fuel$ ها را $flag$ یک می نشود و $good-sample$
شماره می شوند. در نهایت احتمال محاسبه می شود.

تابع $sample-vertex$:

ابتدا جدول $cp+$ که جدول احتمال هست را می گیریم

هر $node$ یک شماره دارد $A=0$ و $B=1$ و $C=2$ و $D=3$

و $E=4$

بعد در جدول آن متغیر با توجه به Parent هایش $node$

که در آن احتمال آن متغیر $True$ است را می داریم

شماره جدول

احتمال آن را در x می ریزد و یک key

$random$ تولید می کند که اگر کوچک تر از x باشد $True$ و در غیر صورت $false$ می گذارد

(۳)

در بین این روش ها روش $Rejection Prior$ اندر
با توجه به احتمال مطلوب تغییر نمی کند زیرا تنها یک نمونه
تولید می کند و بعد احتمال را از بین آنها محاسبه می کند
روش $likelihood$ نمونه های بیشتری تولید می کند و دقیق تر است
چون $evidence$ ها حتماً در نمونه ها قرار دارند

اما روش $data$ از کل متفاوت است و بهتر است
برای جاهای لازم استفاده شود که تعداد متغیر همچنین وابسته
آنها زیاد و تعداد پارامترها زیاد است $iteration$ ها که نمونه خود را
به دست می بخشد اگر در هر متغیرهای $evidence$ تاثیر نداشته
باشند (همان $lagged$) آن وقت اصلاً نمونه های فوری تولید
نمی شوند.

همانطور که در $input 1$ دیده می شود در $query$

بین روش های Likelihood و Rejection Prior
اختلاف زیادی وجود ندارد و Likelihood دقیق تر است

اما در روش Gibbs در هر 7 بار 2 بار اختلاف

زیاد با معیار دیگر Sample ها وجود دارد. محورها

در ۶: A و D و E را هم Update می کنند و Relation آنها

که عملاً اثر کمتری روی هم ندارند

در ۲ که اختلاف کم است. A و B و E و Update

می شوند و چون A و B ارتباط مستقیم دارند نمونه

های بیشتری تولید می شوند

و مثلاً در ۱: فقط Evidence است و B و A و E

را هم Update می کنند و چون با هم Parent و Child

یکدیگر را به خوبی می بیند و در نمونه های مناسب تولید

کننده خوبی حاصل می شود