سوال اول: استنباط متغیر



سوال دوم: Diffusion Models

آ)

اگر واریانس x1 را برابر ۱ بگذاریم، می‌توانیم واریانس را بدون تغییر نگه داریم.

که این برای هر time step صدق می‌کند. پس:

ب)

ج)

د) به طور کلی وقتی یک prior گاوسی به شکل و یک احتمال خطی ـ گاوسی به شکل داریم، راه حل کلی برای posterior به شکل است. که و .

با اعمال کردن و همچنین احتمال خطی ـ گاوسی در این معادله کلی به posterior زیر می‌رسیم.

ه)

و)

ر)

سوال سوم: Score Matching

Langevin Dynamics

آ)

اگر ترم مربوط به نویز وجود نداشته باشد:

این معادله فرآیندی را توصیف می‌کند که در آن نمونه به صورت تکراری در جهت گرادیان تابع log p(x) تنظیم می‌شود.

با یک نقطه تصادفی در شروع می‌کنیم. در هر تکرار داریم:

که گرادیان تابع چگالی احتمال در است و به تند ترین صعود اشاره می‌کند. نیز اندازه گام است که مقدار بزرگی update را کنترل می‌کند. قانون به روز رسانی به طور موثر عمل صعود گرادیان را در انجام می‌دهد. با شروع از یک نقطه تصادفی ، نمونه به طور مکرر در جهتی حرکت می‌کند که با بیشترین سرعت افزایش پیدا کند. این روند تا زمانی ادامه دارد که به ماکسیمم محلی برسد که مربوط به پیک است.

از آن جایی که نقطه اولیه به طور تصادفی در انتخاب می‌شود، ممکن است در هر نقطه از فضای تعریف شده توسط باشد. هر به روز رسانی را در جهت افزایش حرکت می‌دهد. در طی چندین تکرار تضمین می‌شود که از گرادیان بالا می‌رود و در نهایت به اوج می‌رسد. به طور خلاصه، معادله بدست آمده معادل شروع از یک نقطه تصادفی در فضای و انجام gradient ascent رو تابع چگالی احتمال است. این فرآیند منجر به همگرایی در یکی از قله های می‌شود. جایی که چگالی احتمال بالاترین است.

ب)

Score Matching Techniques

سوال چهارم: آنالیز پایداری GAN

بخش اول: تحلیل پایداری پیوسته

آ)

چون برای همه ، حتما نقطه تعادل است زیرا ما در نظر می‌گیریم که برای همه ، و داریم اگر و تنها اگر که نشان می‌دهد این نقطه قطعا نقطه تعادل یکتا است.

بعلاوه ژاکوبین بصورت زیر است:

با در نظر گرفتن خواهیم داشت:

ب)

در نظر بگیریم که آنگاه:

که نشان می‌دهد برای همه ثابت است.

ج)

بخش دوم: تحلیل پایداری گسسته

آ) مقادیر ویژه ژاکوبین به روز رسانی عملگر ها برای simultaneous gradient descent با حساب می‌شوند. در نظر بگیریم که فقط دارای مقادیر ویژه با قسمت حقیقی منفی می‌باشد. مقادیر ویژه ژاکوبین به روز رسانی عملگر برای simultaneous gradient descent همه در دایره واحد هستند اگر و تنها اگر:

اثبات:

برای simultaneous gradient descent داریم:

از این رو . بنابراین مقادیر ویژه با حساب می‌شوند.

برای اینکه ببینیم چه زمانی می‌گوییم با و . پس:

*که کوچک تر از ۱ است اگر و تنها اگر:*

*برای اثبات افزایش یکنوای :*

*ب)*

ج)

بخش سوم: instance noise

آ)   
موقع استفاده از instance noise داریم: