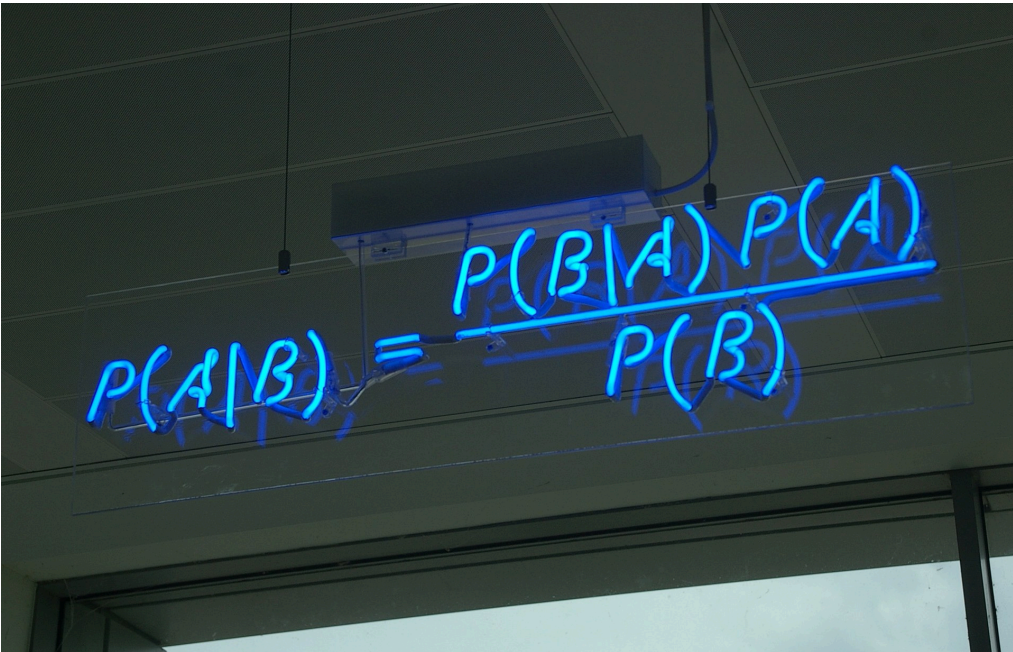


## 19 生成扩散模型漫谈 (三): DDPM = 贝叶斯 + 去噪

Jul By 苏剑林 | 2022-07-19 | 135621位读者 | [引用](#)

到目前为止，笔者给出了生成扩散模型DDPM的两种推导，分别是《生成扩散模型漫谈 (一): DDPM = 拆楼 + 建楼》中的通俗类比方案和《生成扩散模型漫谈 (二): DDPM = 自回归式VAE》中的变分自编码器方案。两种方案可谓各有特点，前者更为直白易懂，但无法做更多的理论延伸和定量理解，后者理论分析上更加完备一些，但稍显形式化，启发性不足。



贝叶斯定理 (来自维基百科)

在这篇文章中，我们再分享DDPM的一种推导，它主要利用到了贝叶斯定理来简化计算，整个过程的“推敲”味道颇浓，很有启发性。不仅如此，它还跟我们后面将要介绍的DDIM模型有着紧密的联系。

[点击阅读全文...](#)

分类: [信息时代](#) 标签: [概率](#), [生成模型](#), [DDPM](#), [扩散](#)

98 评论

[阅读全文](#)

## 6 生成扩散模型漫谈 (二): DDPM = 自回归式VAE

Jul By 苏剑林 | 2022-07-06 | 125233位读者 | 引用

在文章《生成扩散模型漫谈 (一): DDPM = 拆楼 + 建楼》中, 我们为生成扩散模型DDPM构建了“拆楼-建楼”的通俗类比, 并且借助该类比完整地推导了生成扩散模型DDPM的理论形式。在该文章中, 我们还指出DDPM本质上已经不是传统的扩散模型了, 它更多的是一个变分自编码器VAE, 实际上DDPM的原论文中也是将它按照VAE的思路进行推导的。

所以, 本文就从VAE的角度来重新介绍一版DDPM, 同时分享一下自己的Keras实现代码和实践经验。

Github地址: <https://github.com/bojone/Keras-DDPM>

## 多步突破

在传统的VAE中, 编码过程和生成过程都是一步到位的:

$$\text{编码: } x \rightarrow z, \quad \text{生成: } z \rightarrow x \quad (1)$$

[点击阅读全文...](#)

分类: 信息时代 标签: vae, 生成模型, DDPM, 扩散

68 评论

[阅读全文](#)

## 27 生成扩散模型漫谈 (四): DDIM = 高观点DDPM

Jul By 苏剑林 | 2022-07-27 | 203777位读者 | 引用

相信很多读者都听说过甚至读过克莱因的《高观点下的初等数学》这套书, 顾名思义, 这是在学到了更深入、更完备的数学知识后, 从更高的视角重新审视过往学过的初等数学, 以得到更全面的认知, 甚至达到温故而知新的效果。类似的书籍还有很多, 比如《重温微积分》、《复分析: 可视化方法》等。

回到扩散模型，目前我们已经通过三篇文章从不同视角去解读了DDPM，那么它是否也存在一个更高的理解视角，让我们能从中得到新的收获呢？当然有，《Denoising Diffusion Implicit Models》介绍的DDIM模型就是经典的案例，本文一起来欣赏它。

## 思路分析

在《生成扩散模型漫谈（三）：DDPM = 贝叶斯 + 去噪》中，我们提到过该文章所介绍的推导跟DDIM紧密相关。具体来说，文章的推导路线可以简单归纳如下：

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}) \xrightarrow{\text{推导}} p(\mathbf{x}_t | \mathbf{x}_0) \xrightarrow{\text{推导}} p(\mathbf{x}_{t-1} | \mathbf{x}_t, \mathbf{x}_0) \xrightarrow{\text{近似}} p(\mathbf{x}_{t-1} | \mathbf{x}_t) \quad (2)$$

[点击阅读全文...](#)

分类：信息时代 标签：微分方程, 生成模型, DDPM, 扩散

136 评论

[阅读全文](#)

### 12 生成扩散模型漫谈（七）：最优扩散方差估计（上）

Aug By 苏剑林 | 2022-08-12 | 74572位读者 | [引用](#)

对于生成扩散模型来说，一个很关键的问题是生成过程的方差应该怎么选择，因为不同的方差会明显影响生成效果。

在《生成扩散模型漫谈（二）：DDPM = 自回归式VAE》我们提到，DDPM分别假设数据服从两种特殊分布推出了两个可用的结果；《生成扩散模型漫谈（四）：DDIM = 高观点DDPM》中的DDIM则调整了生成过程，将方差变为超参数，甚至允许零方差生成，但方差为0的DDIM的生成效果普遍差于方差非0的DDPM；而《生成扩散模型漫谈（五）：一般框架之SDE篇》显示前、反向SDE的方差应该是一致的，但这原则上在 $\Delta t \rightarrow 0$ 时才成立；《Improved Denoising Diffusion Probabilistic Models》则提出将它视为可训练参数来学习，但会增加训练难度。

所以，生成过程的方差究竟该怎么设置呢？今年的两篇论文《Analytic-DPM: an Analytic Estimate of the Optimal Reverse Variance in Diffusion Probabilistic Models》和

《Estimating the Optimal Covariance with Imperfect Mean in Diffusion Probabilistic Models》算是给这个问题提供了比较完美的答案。接下来我们一起欣赏一下它们的结果。

[点击阅读全文...](#)

分类：信息时代 标签：优化, 生成模型, DDPM, 扩散

41 评论

[阅读全文](#)

### 3 生成扩散模型漫谈（五）：一般框架之SDE篇

Aug By 苏剑林 | 2022-08-03 | 185264位读者 | [引用](#)

在写生成扩散模型的第一篇文章时，就有读者在评论区推荐了宋飏博士的论文《Score-Based Generative Modeling through Stochastic Differential Equations》，可以说该论文构建了一个相当一般化的生成扩散模型理论框架，将DDPM、SDE、ODE等诸多结果联系了起来。诚然，这是一篇好论文，但并不是一篇适合初学者的论文，里边直接用到了随机微分方程（SDE）、Fokker-Planck方程、得分匹配等大量结果，上手难度还是颇大的。

不过，在经过了前四篇文章的积累后，现在我们可以尝试去学习一下这篇论文了。在接下来的文章中，笔者将尝试从尽可能少的理论基础出发，尽量复现原论文中的推导结果。

## 随机微分

在DDPM中，扩散过程被划分为了固定的 $T$ 步，还是用《生成扩散模型漫谈（一）：DDPM = 拆楼 + 建楼》的类比来说，就是“拆楼”和“建楼”都被事先划分为了 $T$ 步，这个划分有着相当大的人为性。事实上，真实的“拆”、“建”过程应该是没有刻意划分的步骤的，我们可以将它们理解为一个在时间上连续的变换过程，可以用随机微分方程（Stochastic Differential Equation, SDE）来描述。

[点击阅读全文...](#)

分类: 信息时代 标签: 微分方程, 生成模型, DDPM, 扩散

157 评论

[阅读全文](#)

## 8 生成扩散模型漫谈 (六): 一般框架之ODE篇

Aug By 苏剑林 | 2022-08-08 | 105023位读者 | [引用](#)

上一篇文章《生成扩散模型漫谈 (五): 一般框架之SDE篇》中, 我们对宋飏博士的论文《Score-Based Generative Modeling through Stochastic Differential Equations》做了基本的介绍和推导。然而, 顾名思义, 上一篇文章主要涉及的是原论文中SDE相关的部分, 而遗留了被称为“概率流ODE (Probability flow ODE)”的部分内容, 所以本文对此做个补充分享。

事实上, 遗留的这部分内容在原论文的正文中只占了一小节的篇幅, 但我们需要新开一篇文章来介绍它, 因为笔者想了很久后发现, 该结果的推导还是没办法绕开Fokker-Planck方程, 所以我们需要一定的篇幅来介绍Fokker-Planck方程, 然后才能请主角ODE登场。

## 再次反思

我们来大致总结一下上一篇文章的内容: 首先, 我们通过SDE来定义了一个前向过程 (“拆楼”):

$$d\boldsymbol{x} = \boldsymbol{f}_t(\boldsymbol{x})dt + g_t d\boldsymbol{w} \tag{3}$$

[点击阅读全文...](#)

分类: 信息时代 标签: 微分方程, 生成模型, DDPM, 扩散

67 评论

[阅读全文](#)

## 18 生成扩散模型漫谈 (八): 最优扩散方差估计 (下)

Aug By 苏剑林 | 2022-08-18 | 42288位读者 | [引用](#)

在上一篇文章《生成扩散模型漫谈（七）：最优扩散方差估计（上）》中，我们介绍并推导了Analytic-DPM中的扩散模型最优方差估计结果，它是直接给出了已经训练好的生成扩散模型的最优方差的一个解析估计，实验显示该估计结果确实能有效提高扩散模型的生成质量。

这篇文章我们继续介绍Analytic-DPM的升级版，出自同一作者团队的论文《Estimating the Optimal Covariance with Imperfect Mean in Diffusion Probabilistic Models》，在官方Github中被称为“Extended-Analytic-DPM”，下面我们也用这个称呼。

## 结果回顾

上一篇文章是在DDIM的基础上，推出DDIM的生成过程最优方差应该是

$$\sigma_t^2 + \gamma_t^2 \bar{\sigma}_t^2 \quad (4)$$

其中 $\bar{\sigma}_t^2$ 是分布 $p(\mathbf{x}_0|\mathbf{x}_t)$ 的方差，它有如下的估计结果（这里取“方差估计2”的结果）：

$$\bar{\sigma}_t^2 = \frac{\bar{\beta}_t^2}{\bar{\alpha}_t^2} \left( 1 - \frac{1}{d} \mathbb{E}_{\mathbf{x}_t \sim p(\mathbf{x}_t)} [\|\epsilon_\theta(\mathbf{x}_t, t)\|^2] \right) \quad (5)$$

[点击阅读全文...](#)

分类：信息时代 标签：优化, 生成模型, DDPM, 扩散

20 评论

[阅读全文](#)

### 30 生成扩散模型漫谈（九）：条件控制生成结果

Aug By 苏剑林 | 2022-08-30 | 135856位读者 | [引用](#)

前面的几篇文章都是比较偏理论的结果，这篇文章我们来讨论一个比较有实用价值的主题——条件控制生成。

作为生成模型，扩散模型跟VAE、GAN、flow等模型的发展史很相似，都是先出来了无条件生成，然后有条件生成就紧接而来。无条件生成往往是为了探索效果上限，而

有条件生成则更多是应用层面的内容，因为它可以实现根据我们的意愿来控制输出结果。从DDPM至今，已经出来了很多条件扩散模型的工作，甚至可以说真正带火了扩散模型的就是条件扩散模型，比如脍炙人口的文生图模型DALL·E 2、Imagen。

在这篇文章中，我们对条件扩散模型的理论基础做个简单的学习和总结。

## 技术分析

从方法上来看，条件控制生成的方式分两种：事后修改（Classifier-Guidance）和事前训练（Classifier-Free）。

[点击阅读全文...](#)

分类：信息时代 标签：概率, 生成模型, DDPM, 扩散

87 评论 [阅读全文](#)