$x_0 \sim q_{data}(x)$

define markov chain

$x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4 \rightarrow x_5 \rightarrow x_T$

x_0 = Real data

$x_T \approx \mathcal{N}(0, I)$ → (pure Gaussian Noise)

THE FORWARD PROCESS

$$q(x_t \mid x_{t-1}) = \mathcal{N}\left(\underbrace{\sqrt{\alpha_t}\, x_{t-1}}, \underbrace{(1-\alpha_t) I}\right)$$

means

$$x_t = \sqrt{\alpha_t}\, x_{t-1} + \sqrt{1-\alpha_t}\, \epsilon_t$$

$\epsilon_t \sim \mathcal{N}(0, I)$

$\boxed{\alpha_t = 1 - \beta_t}$

$\boxed{\beta_t \in (0,1)}$

Small variance schedule

T = Total diffusion steps

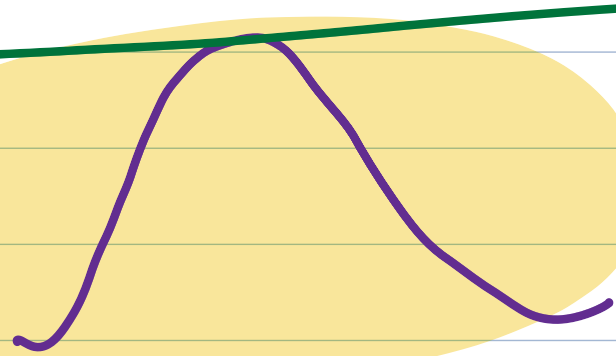
Compute variance

$$Var(x_t) = \alpha_t \, Var(x_{t-1}) + (1-\alpha_t) \, Var(\epsilon_t)$$

Both variances are I,

$$Var(x_t) = \alpha_t I + (1-\alpha_t) I = I$$

Variance



how much spread out random values are

$$x \sim \mathcal{N}(0,1) \quad \text{--- (i)}$$

Mean = 0
Variance = 1 — values may be -3 to 3

$$x \sim \mathcal{N}(0,100) \quad \text{— (ii)}$$

+ much more spread out values around -30 to 30

Variance → Scale of randomness

Scaling a random variable

$y = a x$

$Var(y) = a^2 Var(x)$

if $a=2$

$Var(x)=1$

$Var(2x)=4$

Second

$\boxed{Var(y) = Var(x) + Var(z)}$

$y = x+z$

$$x_t = \sqrt{\alpha}\, x_{t-1} + \sqrt{1-\alpha}\, \epsilon$$

$\boxed{\text{Variance} = \alpha}$

variance $(1-\alpha)$

$\boxed{\text{Add} \rightarrow Var(x_t) = \alpha + 1 - \alpha = 1}$

Because we want

$\boxed{x_t \sim \mathcal{N}(0,1)}$

If variance keep on increasing

- Distribution would explode
- Reverse process becomes unstable

<u>If variance shrank:</u>

- everything collapses to zero

If we donot use square roots then

$$x_t = \alpha x_{t-1} + (1-\alpha)\epsilon$$

$$\boxed{Var = \alpha^2 + (1-\alpha)^2}$$

Not equal to 1.

$\alpha = 0.9$

$\cancel{0}.9^2 + 0.1^2 = 0.81 + 0.01$

$= 0.82$

Variance shrinks

Think of variance like energy

Total energy = signal energy + Noise energy

$$\boxed{x_t = \sqrt{\alpha_t}\, x_{t-1} + \sqrt{1-\alpha_t}\, \epsilon_t}$$

Forward process of a DDPM

This is Not a linear Schedule

All will look similar

one can define series of conditional distributions

$$q(x_t \mid x_{t-1}) = \mathcal{N}\Big(\underbrace{x_t}_{\text{output}} ; \underbrace{\sqrt{\alpha_t}\, x_{t-1}}_{\text{mean}}, \underbrace{(1-\alpha_t)}_{\text{var}} I\Big)$$

Forward, encoding distribution

$\beta_t \in (0,1)$ small

= small variance

"

variance small variance

$$\alpha_t + \beta_t = 1$$

$$\beta_t = \underbrace{1-\alpha_t}$$

$$\sqrt{\alpha_t}$$

$$\alpha_t$$