Problem 1. Пусть $X_0 \sim p_{\text{data}}$ — объект из распределения данных, $t = 0, 1, \dots, T$ – последовательность моментов времени, а $(X_t, t = 0, \dots, T)$ — процесс, сохраняющий дисперсию:

$$X_t = \sqrt{1-\beta_t}\, X_{t-1} + \sqrt{\beta_t}\, \varepsilon_t,$$

где $\varepsilon_t \sim \mathcal{N}(0, I)$ — независимая с $X_1, \dots, X_{t-1}$ величина. Положим $\alpha_t = \prod_{s=1}^{t}(1-\beta_s)$.

Покажите, что условное распределение перехода от момента времени t к моменту времени $t-1$ при известном изначальном объекте $X_0 = x_0$ равно

$$p_{t-1|t,0}(x_{t-1}|x_t, x_0) := p_{X_{t-1}|X_t,X_0}(x_{t-1}|x_t, x_0) = \mathcal{N}\left(x_{t-1} \mid \mu_t(x_t, x_0),\, \sigma_t^2\, \mathbf{I}\right),$$

$$\mu_t(x_t, x_0) = \frac{\sqrt{\alpha_{t-1}}\beta_t}{1-\alpha_t}x_0 + \frac{\sqrt{1-\beta_t}(1-\alpha_{t-1})}{1-\alpha_t}x_t, \;\; \sigma_t^2 = \frac{1-\alpha_{t-1}}{1-\alpha_t}\beta_t.$$

Применим формулу Байеса:

$$p_{X_{t-1}|X_t,X_0}(x_{t-1}|x_t, x_0) = \frac{p_{X_t|X_{t-1},X_0}(x_t|x_{t-1}, x_0) \cdot p_{X_{t-1}|X_0}(x_{t-1}|x_0)}{p_{X_t|X_0}(x_t|x_0)}$$

Мы явно можем посчитать каждое из распределений (при $X_0 = x_0$):

$$p_{X_t|X_{t-1}}(x_t|x_{t-1}) = \mathcal{N}(x_t|\sqrt{1-\beta_t}x_{t-1}, \beta_t\mathbf{I})$$

$$p_{X_{t-1}|X_0}(x_{t-1}|x_0) = \mathcal{N}(x_{t-1}|x_0\sqrt{\alpha_{t-1}}, (1-\alpha_{t-1})\mathbf{I})$$

$$p_{X_t|X_0}(x_t|x_0) = \mathcal{N}(x_t|x_0\sqrt{\alpha_t}, (1-\alpha_t)\mathbf{I})$$

Пусть мы понимаем, что нужное нам распределение имеет вид нормального на x_{t-1}. Тогда посмотрим на его функциональный вид:

$$p_{X_{t-1}|X_t,X_0}(x_{t-1}|x_t, x_0) \propto \exp\left[-\frac{1}{2}\frac{(x_{t-1}-\mu)^\top(x_{t-1}-\mu)}{\sigma^2}\right] = \exp\left[-\frac{1}{2}\left(\frac{x_{t-1}^\top x_{t-1}}{\sigma^2} + \frac{\mu^\top\mu}{\sigma^2} - 2\frac{x_{t-1}^\top\mu}{\sigma^2}\right)\right]$$

Теперь подставим плотности в дробь выше:

$$p_{X_{t-1}|X_t,X_0}(x_{t-1}|x_t, x_0) \propto$$

$$\exp\left[-\frac{1}{2}\left(\frac{(x_t-\sqrt{1-\beta_t}x_{t-1})^\top(x_t-\sqrt{1-\beta_t}x_{t-1})}{\beta_t} + \frac{(x_{t-1}-\sqrt{\alpha_{t-1}}x_0)^\top(x_{t-1}-\sqrt{\alpha_{t-1}}x_0)}{1-\alpha_{t-1}} - \frac{(x_t-\sqrt{\alpha_t}x_0)^\top(x_t-\sqrt{\alpha_t}x_0)}{1-\alpha_t}\right)\right]$$

$$= \exp\left[-\frac{1}{2}\left(\frac{x_t^\top x_t - 2\sqrt{1-\beta_t}x_t^\top x_{t-1} + (1-\beta_t)x_{t-1}^\top x_{t-1}}{\beta_t} + \frac{x_{t-1}^\top x_{t-1} - 2\sqrt{\alpha_{t-1}}x_{t-1}^\top x_0 + \alpha_{t-1}x_0^\top x_0}{1-\alpha_{t-1}} - C\right)\right] =$$

Последнюю скобку расскрывать не было смысла, ведь она константа по x_{t-1}. Теперь осталось сгруппировать слогаемые под функциональный вид выше:

$$= \exp\left[-\frac{1}{2}\left(x_{t-1}^\top x_{t-1}\left(\frac{1-\beta_t}{\beta_t} + \frac{1}{1-\alpha_{t-1}}\right) - 2x_{t-1}^\top\left(\frac{\sqrt{1-\beta_t}x_t}{\beta_t} + \frac{x_0\sqrt{\alpha_{t-1}}}{1-\alpha_{t-1}}\right) - C\right)\right]$$

У первого слагаемого стоит обратная дисперсия:

$$\frac{1}{\sigma_t^2} = \frac{1-\beta_t}{\beta_t} + \frac{1}{1-\alpha_{t-1}} = \frac{1-\beta_t-\alpha_{t-1}(1-\beta_t)+\beta_t}{\beta_t(1-\alpha_{t-1})} = \frac{1-\alpha_t}{\beta_t(1-\alpha_{t-1})} \implies \sigma_t^2 = \frac{\beta_t(1-\alpha_{t-1})}{1-\alpha_t}$$

У второго слагаемого стоит часто матожидания и дисперсии:

$$\mu_t(x_t, x_0) = \left(\frac{\sqrt{1-\beta_t}x_t}{\beta_t} + \frac{x_0\sqrt{\alpha_{t-1}}}{1-\alpha_{t-1}}\right)\frac{\beta_t(1-\alpha_{t-1})}{1-\alpha_t} = \frac{\sqrt{1-\beta_t}(1-\alpha_{t-1})}{1-\alpha_t}x_t + \frac{\beta_t\sqrt{a_{t-1}}}{1-\alpha_t}x_0$$

Problem 2. Пусть X_t — процесс, заданный стохастическим дифференциальным уравнением

$$dX_t = \frac{x_1 - X_t}{1-t}dt + \sqrt{\gamma}dW_t,$$

где x_1 — константа и $X_0 = x_0$ — константа. Найдите $\mathbb{E}X_t$ и $\mathbb{V}arX_t$.

Сначала запишем дискретизацию:

$$X_{t+h} = X_t + \frac{x_1 - X_t}{1-t}h + \sqrt{\gamma h}\varepsilon_t$$

Теперь считаем матожидание. Просто навешиваем его с двух сторон и переходим к производной:

$$m_{t+h} = m_t + \frac{x_1}{1-t}h - \frac{m_t}{1-t}h$$

$$\frac{m_{t+h} - m_t}{h} = \frac{x_1}{1-t} - \frac{m_t}{1-t}$$

$$h \to 0: \ \frac{d}{dt}(m_t) = \frac{x_1}{1-t} - \frac{m_t}{1-t}$$

Теперь пользуемся домножением на экспоненту и производной произведения:

$$\frac{d}{dt}\left(m_t \cdot \exp\left(\int_0^t \frac{1}{1-s}ds\right)\right) = \frac{x_1}{1-t} \cdot \exp\left(\int_0^t \frac{1}{1-s}ds\right)$$

Теперь берём интеграл с двух сторон от 0 до t:

$$m_t \cdot \exp\left(\int_0^t \frac{1}{1-s}ds\right) - m_0 \cdot \exp\left(\int_0^0 \frac{1}{1-s}ds\right) = \int_0^t \frac{x_1}{1-s} \cdot \exp\left(\int_0^s \frac{1}{1-h}dh\right)ds$$

Заменяем m_0 на x_0. Ну и заметим, что справа под интегралом стоит производная:

$$m_t \cdot \exp\left(\int_0^t \frac{1}{1-s}ds\right) - x_0 = x_1\left(\exp\left(\int_0^t \frac{1}{1-s}ds\right) - 1\right)$$

$$m_t = \frac{x_0}{\alpha_t} + x_1\left(1 - \frac{1}{\alpha_t}\right)$$

Ну и оказывается, тут α_t считается, так что считаем:

$$\exp\left(\int_0^t \frac{1}{1-s}ds\right) = \exp\left(\log(1-t)^{-1}\right) = \frac{1}{1-t}$$

$$\implies m_t = x_0(1-t) + tx_1$$

Теперь разбираемся с дисперсией. План точно такой же:

$$v_{t+h} = v_t\left(1 - \frac{h}{1-t}\right)^2 + h\gamma$$

Только тут уберём квадрта с помощью Тейлора:

$$v_{t+h} = v_t\left(1 - \frac{2h}{1-t}\right) + h\gamma + o(h)$$

$$\frac{v_{t+h} - v_t}{h} = -\frac{2}{1-t}v_t + \gamma + o(1)$$

$$h \to 0: \ \frac{d}{dt}(v_t) = -\frac{2}{1-t}v_t + \gamma$$

$$\frac{d}{dt}\left(v_t \cdot \exp\left(\int_0^t \frac{2}{1-s}ds\right)\right) = \gamma \exp\left(\int_0^t \frac{2}{1-s}ds\right)$$

$$v_t \cdot \exp\left(\int_0^t \frac{2}{1-s}ds\right) - v_0 \cdot \exp\left(\int_0^0 \frac{2}{1-s}ds\right) = \int_0^t \gamma \exp\left(\int_0^s \frac{2}{1-h}dh\right)ds$$

Так как x_0 зафиксировано, то дисперсия у неё нулевая:

$$v_t \cdot \alpha_t = \gamma \int_0^t \alpha_s ds$$

Считаем α_t:

$$\alpha_t = \exp(\log(1-t)^{-2}) = \frac{1}{(1-t)^2}$$

$$\int_0^t \frac{1}{(1-s)^2} ds = -\frac{t}{t-1}$$

$$v_t = \gamma(1-t)^2 \frac{t}{1-t} = \gamma(1-t)t$$

Solution: $\mathbb{E}X_t = x_0(1-t) + tx_1, \;\; \mathbb{V}arX_t = \gamma(1-t)t$