Диффузионные модели, восстановление изображений через обратные линейные задачи

Тингир Бадмаев

18 марта 2022

> Вернуть утраченное — да, возможно,
> сохранить то, что есть, — любой ценой
>
> Тони Старк, *Мстители: Финал*

Linear Inverse Problems

- Linear Inverse Problems

$$y = Hx + z,$$
$$x \in \mathbb{R}^n, \quad y \in \mathbb{R}^m, \quad H \in \mathbb{R}^{m \times n}$$
$$z \sim \mathcal{N}(0, \sigma_y^2 I)$$

- **Supervised решения** - обучить модель на парах $(x_i, y_i)_{i=1}^N$

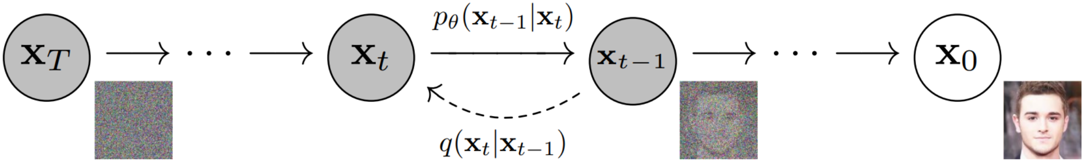
$$x_\theta = f_\theta(y)$$

- **Unsupervised решения** - моделировать апостериорное распределение $p_\theta(x|y) \approx p(x|y) \propto p_{data}(x)p(y|x)$.
  Один из подходов использовать нейросеть для предсказания score фукнции и динамику Ланжевена

$$x_{t+1} = x_t + \alpha \nabla_{x_t} \log p_\theta(x_t|y) + \sqrt{2\alpha} z_t, \quad z_t \sim \mathcal{N}(0, I)$$

DDPM, зашумление

- Общая схема



- Зашумление данных

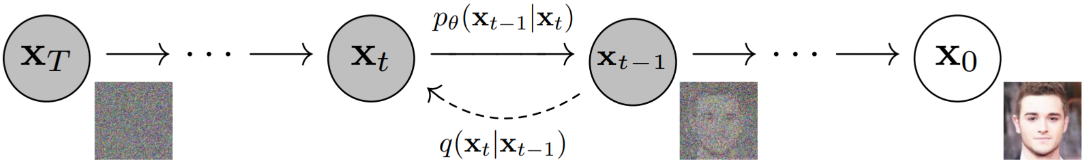
$$q(x_{1:T}|x_0) = \prod_{t=1}^{T} q(x_t|x_{t-1}), \quad q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I)$$

- Зашумление на несколько шагов

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)I), \quad \alpha_t = 1-\beta_t, \quad \bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$$

DDPM, удаление шума

- Общая схема



- Удаление шума

$$p_\theta(x_{0:T}) = p(x_T)\prod_{t=1}^{T} p_\theta(x_{t-1}|x_t), \quad p_\theta(x_T) = \mathcal{N}(x_T; 0, I)$$

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t))$$

DDPM

- Зашумление данных

$$q(x_{1:T}|x_0) = \prod_{t=1}^{T} q(x_t|x_{t-1}), \quad q(x_t|x_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I)$$

- По формуле Байеса:

$$q(x_{t+1}|x_t, x_0) = \frac{q(x_t|x_{t+1}, x_0)q(x_{t+1}|x_0)}{q(x_t|x_0)}$$

$$q(x_{1:T}|x_0) = q(x_1|x_0)\prod_{t=1}^{T-1} q(x_{t+1}|x_t, x_0) = q(x_1|x_0)\prod_{t=1}^{T-1} \frac{q(x_t|x_{t+1}, x_0)q(x_{t+1}|x_0)}{q(x_t|x_0)}$$

$$= q(x_T|x_0)\prod_{t=1}^{T-1} q(x_t|x_{t+1}, x_0)$$

- Удаление шума

$$p_\theta(x_{0:T}) = p(x_T)\prod_{t=0}^{T-1} p_\theta(x_t|x_{t+1})$$

- ELBO

$$\mathbb{E}_{q(x_{0:T})}\left[D_{KL}(q(x_T|x_0)||p_\theta(x_T)) + \sum_{t=1}^{T-1} D_{KL}(q(x_t|x_{t+1}, x_0)||p_\theta(x_t|x_{t+1})) - \log p_\theta(x_0|x_1)\right]$$

Схема семплирования в DDPM с помощью репараметризации

- Процесс зашумления и денойзинга
  - $q(x_{1:T}|x_0) = q(x_T|x_0) \prod_{t=1}^{T-1} q(x_t|x_{t+1}, x_0)$
  - $p_\theta(x_{0:T}) = p(x_T) \prod_{t=0}^{T-1} p_\theta(x_t|x_{t+1})$
- Репараметризация для вар. вывода DDPM
  - Вспомним, что:

$$x_t(x_0, \epsilon) = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t}\epsilon, \quad \epsilon \sim \mathcal{N}(0, I)$$
$$x_0 = \frac{1}{\sqrt{\bar{\alpha}_t}} \left(x_t(x_0, \epsilon) - \sqrt{1 - \bar{\alpha}_t}\epsilon\right)$$

  - Тогда ELBO и предсказанное x_0 из момента времени $t + 1$ выглядят:

$$\mathcal{L}_t = \mathbb{E}_q \left[D_{KL}(q(x_t|x_{t+1}, x_0, y)||p_\theta(x_t|x_{t+1}, y))\right]$$
$$x_{\theta,t} = \frac{1}{\sqrt{\alpha_{t+1}}} \left(x_{t+1} - \sqrt{1 - \bar{\alpha}_{t+1}}\epsilon_\theta(x_{t+1}, t + 1)\right)$$
$$\mathcal{L}_t = \mathbb{E}_{q,\epsilon} \left[\frac{\beta_{t+1}^2}{2\sigma_{t+1}^2 \alpha_{t+1}(1-\bar{\alpha}_{t+1})} \|\epsilon - \epsilon_\theta(x_{t+1}, t + 1)\|^2\right]$$

  - Цепочка - $x_T \xrightarrow{\theta} x_{\theta,T-1} \xrightarrow{q_{T-1}} x_{T-1} \xrightarrow{\theta} \ldots \xrightarrow{\theta} x_{\theta,0} = x_0$

DDRM

- Зашумление данных

$$q(x_{1:T}|x_0, y) = q(x_T|x_0, y)\prod_{t=1}^{T-1} q(x_t|x_{t+1}, x_0, y)$$

- Удаление шума

$$p_\theta(x_{0:T}|y) = p(x_T|y)\prod_{t=0}^{T-1} p_\theta(x_t|x_{t+1}, y)$$

- ELBO

$$\mathbb{E}_{q(x_{0:T})}\left[D_{KL}(q(x_T|x_0, y)||p_\theta(x_T|y))\right.$$
$$\left.+\sum_{t=1}^{T-1} D_{KL}(q(x_t|x_{t+1}, x_0, y)||p_\theta(x_t|x_{t+1}, y)) - \log p_\theta(x_0|x_1, y)\right.$$

SVD для перехода в другую систему координат

- **SVD**. Для матриц $H \in \mathbb{R}^{m \times n}$ выполнено, что $H = U \Sigma V^T$, где $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ ортогональные матрицы, $\Sigma \in \mathbb{R}^{m \times n}$ прямоугольная диагональная матрица с числами $(s_1, s_2, \ldots, s_k, 0, 0, \ldots, 0)$, $s_1 \geqslant s_2 \geqslant s_3 \geqslant \ldots \geqslant s_k > 0$ на главной диагонали, $k \leqslant min(n, m)$.
- **Псевдообратной матрицей** для Σ назовем матрицу $\Sigma^+ \in \mathbb{R}^{n \times m}$, у которой на главной диагонали стоят числа $(1/s_1, 1/s_2, \ldots, 1/s_k, 0, 0, \ldots, 0)$.
- SVD преобразование для линейной обраной задачи

$$y = Hx_0 + z, z \sim \mathcal{N}(0, \sigma_y^2 I),$$
$$H = U \Sigma V^T,$$
$$\bar{x}_t^i = (V^T x_t)^i,\ \bar{y}^i = (\Sigma^+ U^T y)^i,$$
$$\bar{y} = \Sigma^+ U^T y = V^T x_0 + \Sigma^+ U^T z = \bar{x}_0 + \Sigma^+ U^T z,$$
$$\bar{y}^i \sim \mathcal{N}(\bar{x}_0^i, \frac{\sigma_y^2}{s_i^2}), \text{ если } s_i > 0.$$

Диффузия в новой системе координат

- Зашумление данных

$$q(\bar{x}_{1:T}|x_0, y) = q(\bar{x}_T|x_0, y) \prod_{t=1}^{T-1} q(\bar{x}_t|x_{t+1}, x_0, y)$$

- Удаление шума

$$p_\theta(\bar{x}_{0:T}|y) = p(\bar{x}_T|y) \prod_{t=0}^{T-1} p_\theta(\bar{x}_t|x_{t+1}, y)$$

- ELBO

$$\mathbb{E}_{q(\bar{x}_{0:T})} [D_{KL}(q(\bar{x}_T|x_0, y)||p_\theta(\bar{x}_T|y))$$
$$+ \sum_{t=1}^{T-1} D_{KL}(q(\bar{x}_t|x_{t+1}, x_0, y)||p_\theta(\bar{x}_t|x_{t+1}, y)) - \log p_\theta(\bar{x}_0|x_1, y)$$

Построение вариационного вывода для семплирования DDRM I

- Процесс зашумления

$$q(\bar{x}_{1:T}|x_0, y) = q(\bar{x}_T|x_0, y)\prod_{t=1}^{T-1} q(\bar{x}_t|x_{t+1}, x_0, y)$$

- Зафиксируем переходы для апостериорных распределений:

$$q(\bar{x}_T^i|x_0, y) = \begin{cases} \mathcal{N}(\bar{y}^i, \sigma_T^2 - \frac{\sigma_y^2}{s_i^2}), \text{ если } s_i > 0, \\ \mathcal{N}(\bar{x}_0^i, \sigma_T^2), \text{ если } s_i = 0. \end{cases}$$

$$q(\bar{x}_t^i|x_{t+1}, x_0, y) = \begin{cases} \mathcal{N}(\bar{x}_0^i(1-\eta_b) + \eta_b\bar{y}^i, \sigma_t^2 - \frac{\sigma_y^2}{s_i^2}\eta_b^2), \text{ если } s_i > 0 \text{ и } \sigma_t >= \frac{\sigma_y}{s_i}, \\ \mathcal{N}(\bar{x}_0^i + \sqrt{1-\eta^2}\sigma_t\frac{\bar{y}^i - \bar{x}_0^i}{\sigma_y/s_i}, \eta^2\sigma_t^2), \text{ если } s_i > 0 \text{ и } \sigma_t < \frac{\sigma_y}{s_i}, \\ \mathcal{N}(\bar{x}_0^i + \sqrt{1-\eta^2}\sigma_t\frac{\bar{x}_{t+1}^i - \bar{x}_0^i}{\sigma_{t+1}}, \eta^2\sigma_t^2), \text{ если } s_i = 0. \end{cases},$$

где $\eta, \eta_b \in (0; 1]$.

Вар. вывод DDRM II, прямой процесс

- Процесс зашумления

$$q(\bar{x}_{1:T}|x_0, y) = q(\bar{x}_T|x_0, y)\prod_{t=1}^{T-1} q(\bar{x}_t|x_{t+1}, x_0, y)$$

- Какая тогда динамика для зушемления из 0 момента в момент t? Предположение:

$$q(x_t|x_0) = \mathcal{N}(x_0, \sigma_t^2 I)$$

$$\Updownarrow$$

$$q(\bar{x}_t|x_0) = \mathcal{N}(\bar{x}_0, \sigma_t^2 I), \text{ потому что}$$

$$\bar{x}_t = V^T x_t,$$

$$x_t = V\bar{x}_t, V \text{ - ортогональная матрица и}$$

$$\text{у всех координат вектора дисперсия } \sigma_t^2.$$

Вар. вывод DDRM III, прямой процесс

- Hints:
  - $p(z_1|z_0) = \mathcal{N}(z_0, V_1), p(z_2|z_1) = \mathcal{N}(\alpha z_1, V_2) \to p(z_2|z_0) = \mathcal{N}(\alpha z_0, \alpha^2 V_1 + V_2)$
  - $p(z_1) = \mathcal{N}(\mu_1, V_1), p(z_2) = \mathcal{N}(\mu_2, V_2) \to p(z_1 + z_2) = \mathcal{N}(\mu_1 + \mu_2, V_1 + V_2)$
  - $\bar{y}^i \sim q(\bar{y}^i|x_0) = \mathcal{N}(\bar{x}_0^i, \frac{\sigma_y^2}{s_i^2})$, если $s_i > 0$, и $\bar{y}^i = \bar{x}_0^i$, если $s_i = 0$
- Докажем для момента времени T предположение о $q(\bar{x}_T^i|x_0)$
  - $s_i = 0$ и $q(\bar{x}_T^i|x_0, y) = \mathcal{N}(\bar{x}_0^i, \sigma_T^2)$ - выполнено по построению

    $\bar{y}^i = \bar{x}_0^i$ и $q(\bar{x}_T^i|x_0) = q(\bar{x}_T^i|x_0, y)$
  - $s_i > 0$ и $q(\bar{x}_T^i|x_0, y) = \mathcal{N}(\bar{y}^i, \sigma_T^2 - \frac{\sigma_y^2}{s_i^2})$

    По Hint-1 $q(\bar{x}_T^i|x_0) = \mathcal{N}(1 * \bar{x}_0^i, 1 * 1 * \frac{\sigma_y^2}{s_i^2} + \sigma_T^2 - \frac{\sigma_y^2}{s_i^2}) = \mathcal{N}(\bar{x}_0^i, \sigma_T^2)$

Вар. вывод DDRM IV, прямой процесс

- Hints:
  - $p(z_1|z_0) = \mathcal{N}(z_0, V_1), p(z_2|z_1) = \mathcal{N}(\alpha z_1, V_2) \rightarrow p(z_2|z_0) = \mathcal{N}(\alpha z_0, \alpha^2 V_1 + V_2)$
  - $p(z_1) = \mathcal{N}(\mu_1, V_1), p(z_2) = \mathcal{N}(\mu_2, V_2) \rightarrow p(z_1 + z_2) = \mathcal{N}(\mu_1 + \mu_2, V_1 + V_2)$
  - $\bar{y}^i \sim q(\bar{y}^i|x_0) = \mathcal{N}(\bar{x}_0^i, \frac{\sigma_y^2}{s_i^2})$, если $s_i > 0$, и $\bar{y}^i = \bar{x}_0^i$, если $s_i = 0$
- Докажем для момента времени $t < T$ предположение о $q(\bar{x}_t^i|x_0)$
  - $s_i > 0$, $\sigma_t >= \frac{\sigma_y}{s_i}$, и $q(\bar{x}_t^i|x_{t+1}, x_0, y) = \mathcal{N}(\bar{x}_0^i(1 - \eta_b) + \eta_b \bar{y}^i, \sigma_t^2 - \frac{\sigma_y^2}{s_i^2}\eta_b^2)$.
    В подсчете плотности мы не используем x_{t+1}, тогда можем убрать из условия x_{t+1}. С помощью Hint-1 посчитаем распределение $\eta_b \bar{y}|x_0$ и с помощью Hint-2 посчитаем $q(\bar{x}_t|x_0)$.

$$\text{Среднее} = \bar{x}_0^i(1 - \eta_b) + \eta_b \bar{x}_0^i = \bar{x}_0^i,$$

$$\text{Дисперсия} = \sigma_t^2 - \frac{\sigma_y^2}{s_i^2}\eta_b^2 + \frac{\sigma_y^2}{s_i^2}\eta_b^2 = \sigma_t^2.$$

Вар. вывод DDRM V, прямой процесс

- Hints:
  - $p(z_1|z_0) = \mathcal{N}(z_0, V_1), p(z_2|z_1) = \mathcal{N}(\alpha z_1, V_2) \to p(z_2|z_0) = \mathcal{N}(\alpha z_0, \alpha^2 V_1 + V_2)$
  - $p(z_1) = \mathcal{N}(\mu_1, V_1), p(z_2) = \mathcal{N}(\mu_2, V_2) \to p(z_1 + z_2) = \mathcal{N}(\mu_1 + \mu_2, V_1 + V_2)$
  - $\bar{y}^i \sim q(\bar{y}^i|x_0) = \mathcal{N}(\bar{x}_0^i, \frac{\sigma_y^2}{s_i^2})$, если $s_i > 0$, и $\bar{y}^i = \bar{x}_0^i$, если $s_i = 0$
- Докажем для момента времени $t < T$ предположение о $q(\bar{x}_t^i|x_0)$
  - $s_i > 0$, $\sigma_t < \frac{\sigma_y}{s_i}$, и $q(\bar{x}_t^i|x_{t+1}, x_0, y) = \mathcal{N}(\bar{x}_0^i + \sqrt{1-\eta^2}\sigma_t \frac{\bar{y}^i - \bar{x}_0^i}{\sigma_y/s_i}, \eta^2\sigma_t^2)$.
    В подсчете плотности мы не используем x_{t+1}, тогда можем убрать из условия x_{t+1}. С помощью Hint-1 посчитаем распределение для $\sqrt{1-\eta^2}\sigma_t \frac{\bar{y}^i - \bar{x}_0^i}{\sigma_y/s_i}|x_0$ и с помощью Hint-2 посчитаем $q(\bar{x}_t|x_0)$.

$$\text{Среднее } = \bar{x}_0^i + \sqrt{1-\eta^2}\sigma_t \frac{\bar{x}_0^i - \bar{x}_0^i}{\sigma_y/s_i} = \bar{x}_0^i,$$
$$\text{Дисперсия } = \eta^2\sigma_t^2 + (1-\eta^2)\sigma_t^2 = \sigma_t^2.$$

Вар. вывод DDRM VI, прямой процесс

- Hints:
  - $p(z_1|z_0) = \mathcal{N}(z_0, V_1), p(z_2|z_1) = \mathcal{N}(\alpha z_1, V_2) \rightarrow p(z_2|z_0) = \mathcal{N}(\alpha z_0, \alpha^2 V_1 + V_2)$
  - $p(z_1) = \mathcal{N}(\mu_1, V_1), p(z_2) = \mathcal{N}(\mu_2, V_2) \rightarrow p(z_1 + z_2) = \mathcal{N}(\mu_1 + \mu_2, V_1 + V_2)$
  - $\bar{y}^i \sim q(\bar{y}^i|x_0) = \mathcal{N}(\bar{x}_0^i, \frac{\sigma_y^2}{s_i^2})$, если $s_i > 0$, и $\bar{y}^i = \bar{x}_0^i$, если $s_i = 0$
- Докажем для момента времени $t < T$ предположение о $q(\bar{x}_t^i|x_0)$
  - $s_i = 0$, и $q(\bar{x}_t^i|x_{t+1}, x_0, y) = \mathcal{N}(\bar{x}_0^i + \sqrt{1-\eta^2}\sigma_t \frac{\bar{x}_{t+1}^i - \bar{x}_0^i}{\sigma_{t+1}}, \eta^2\sigma_t^2)$.
    Докажем по индукции, что $q(\bar{x}_t^i|x_0) = \mathcal{N}(\bar{x}_0^i, \sigma_t^2)$, если $s_i = 0$.
    База - для момента времени $t = T$ предположение выполнено. Допустим, что $q(\bar{x}_{t+1}^i|x_0) = \mathcal{N}(\bar{x}_0^i, \sigma_{t+1}^2)$, докажем для $q(\bar{x}_t^i|x_0)$.
    Доказательство аналогично двум предыдущим случаям. Теперь условная вероятность не зависит от y, поскольку $s_i = 0$, в спектральном пространсве, в котором мы делаем диффузию в векторе $\bar{y}$ нет информации о $i-$ой компоненте, поэтому опускаем условие.
    Повторяем рассуждения о среднем и дисперсии с помощью Hint-1, Hint-2.

Вар. вывод DDRM VII, обратный процесс

- Удаление шума

$$p_\theta(\bar{x}_{0:T}|y) = p(\bar{x}_T|y)\prod_{t=0}^{T-1} p_\theta(\bar{x}_t|x_{t+1}, y)$$

- Для DDRM процесс денойзинга будет выглядеть:

$$p(\bar{x}_T^i|y) = \begin{cases} \mathcal{N}(\bar{y}^i, \sigma_T^2 - \frac{\sigma_y^2}{s_i^2}), \text{ если } s_i > 0, \\ \mathcal{N}(0, \sigma_T^2), \text{ если } s_i = 0. \end{cases}$$

$$p(\bar{x}_t^i|x_{t+1}, y) = \begin{cases} \mathcal{N}(\bar{x}_{\theta,t}^i(1-\eta_b) + \eta_b\bar{y}^i, \sigma_t^2 - \frac{\sigma_y^2}{s_i^2}\eta_b^2), \text{ если } s_i > 0 \text{ и } \sigma_t >= \frac{\sigma_y}{s_i}, \\ \mathcal{N}(\bar{x}_{\theta,t}^i + \sqrt{1-\eta^2}\sigma_t\frac{\bar{y}^i - \bar{x}_{\theta,t}^i}{\sigma_y/s_i}, \eta^2\sigma_t^2), \text{ если } s_i > 0 \text{ и } \sigma_t < \frac{\sigma_y}{s_i}, \\ \mathcal{N}(\bar{x}_{\theta,t}^i + \sqrt{1-\eta^2}\sigma_t\frac{\bar{x}_{t+1}^i - \bar{x}_{\theta,t}^i}{\sigma_{t+1}}, \eta^2\sigma_t^2), \text{ если } s_i = 0. \end{cases}$$

Вар. вывод DDRM VIII, алгоритм семплирования

- Получить $\bar{x}_T \sim p(\bar{x}_T|y)$ - семплировать с помощью y, если можем($s_i > 0$)
- Определить x_T - по правилу $x_t = V\bar{x}_t$
- Получить $x_{\theta,T-1}$ с помощью DDPM - предсказать x_0 из момента времени T
- Перевести $x_{\theta,T-1}$ в $\bar{x}_{\theta,T-1}$ - перейти в спектральное пространство, чтобы зашумлять используя y
- Зашумить $\bar{x}_{\theta,T-1}$ с помощью $q(\bar{x}_t|x_{t+1}, x_t, y)$ - получим $\bar{x}_{T-1}$
- Перевести обратно в пространство пикселей и получить x_{T-1}
- Повторить процесс с 3 пункта, если не дошли до времени $t = 0$

Таким образом, мы получили итерационный процесс на основе диффузионных моделей(которые можно ускорять!), который на каждом шаге обновляет наш объект с учетом информации из y, если это возможно. Поскольку изначально перед нами стояла задача восстановления картинок, мы хотим как можно лучше сохнить имеющуюся информацию.

Есть две прямые динамики зашумления

Возникает вопрос - а как это мы используем DDPM, хотя наша новая построенная динамика для семплирования в пространстве пикселей использует правило перехода $q(x_t|x_0) \sim \mathcal{N}(x_t|x_0, \sigma_t^2 I)$?

- NCSN подход с $q(x_t|x_0) \sim \mathcal{N}(x_t|x_0, \sigma_t^2 I)$
- DDPM подход с $q(x_t|x_0) \sim \mathcal{N}(x_t|\sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)I)$

На самом деле после этапа обучения две динамики эквивалентны.

$$x_t = x_0 + \sigma_t \epsilon$$
$$\color{blue}{x_t} = \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon$$
$$\frac{x_t}{\sqrt{1+\sigma_t^2}} = \frac{x_0}{\sqrt{1+\sigma_t^2}} + \frac{\sigma_t}{\sqrt{1+\sigma_t^2}}\epsilon$$
$$\alpha_t = 1/(1+\sigma_t^2),\ \color{blue}{x_t} = x_t/\sqrt{1+\sigma_t^2}$$

Примеры линейных операторов с $O(n)$ паматью

- **Денойзинг**. Тогда линейный оператор $H = I$, и его SVD разложение U, Σ, V^T - это единичные матрицы, хранить которые мы можем за $O(n)$ памяти.
- **Закраска**. В данном случае $H = I\Sigma P^T$, где P - матрица перестановки, Σ матрица размера $k \times n$, с единицами на главной диагонали, где n - общее число пикселей, k - число незамаскированных пикселей. Матрицу перестановки мы можем хранить за $O(n)$ памяти.
- **Повышение разрешения картинки**. (на след. слайде)
- **Колоризация**. (аналогично увеличению разрешения)

Примеры линейных операторов с $O(n)$ памятью

- **Повышение разрешения картинки.**
  Посмотрим на модельный пример для матрицы размера 4×4, мы хотим уменьшить её в размере в 2 раза по обеим сторонам. Вытянем матрицу в вектор x длины 16. Наше ядро H будет выглядеть следующим образом:
  $$H = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} P,$$
  где P это матрица перестановки, которая первый патч(левый верхний угол) переводит в первые 4 элемента, второй патч(правый верхний угол) во вторые 4 элемента вектора x и т.д.. Тогда y имеет размерность 4, что соответствует low-res. картинке.

- **Повышение разрешения картинки.**
  Нам понадобится произведение Кронекера:

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}.$$

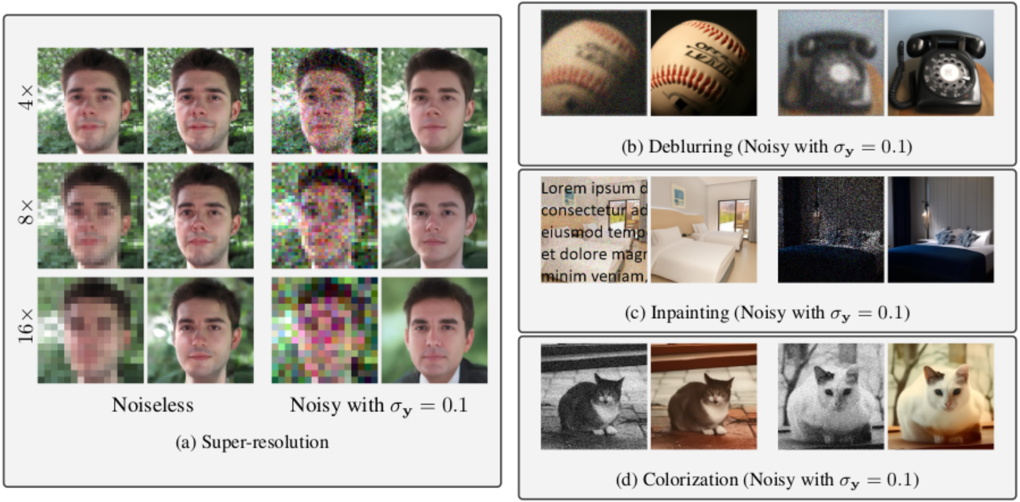
$y^i = (Hx)^i = k^T p^i$, где k - вектор размера r^2 со значениями $\dfrac{1}{r^2}$, p^i векторизованный $i-$ый патч размера $r \times r$. Тогда

$$H = (I \otimes k^T)P,$$
$$H \in \mathbb{R}^{\frac{n}{r^2} \times n},\ I \in \mathbb{R}^{\frac{n}{r^2} \times \frac{n}{r^2}},\ k^T \in \mathbb{R}^{1 \times r^2},\ P \in \mathbb{R}^{n \times n}.$$

Примеры линейных операторов с $O(n)$ памятью

- **Повышение разрешения картинки**.
  Разложим вектор $k^T = U_k \Sigma_k V_k^T$, по свойству кронекерового произведения:
  $$H = (I \otimes k^T)P = ((III) \otimes (U_k \Sigma_k V_k^T))P = (I \otimes U_k)(I \otimes \Sigma_k)(I \otimes V_k^T)P,$$
  $$U_k \in \mathbb{R}^{1 \times 1}, \Sigma_k \in \mathbb{R}^{1 \times r^2}, V_k^T \in \mathbb{R}^{r^2 \times r^2}.$$
  $I \otimes U_k$ диагональная матрица с элементали либо всеми 1, либо всеми -1. В Σ_k ровно один нелувой элемент и это вектор-строка. Тогда в матрице $I \otimes \Sigma_k$ в кажой строке(строк $\frac{n}{r^2}$) будет ровно один ненулевой элемент равный сингулярному числу. Самый большой интерес представляет матрица $I \otimes V_k^T$, она будет иметь блочно-диагональный вид. Блоки на главной диагонали будут матрицами V_k, все остальные блоки равны $\mathbb{O}$. Т. о., мы храним матрицу H за $O(n)$ памяти.

Пример семплов



(a) Super-resolution

(b) Deblurring (Noisy with $\sigma_{\mathbf{y}} = 0.1$)

(c) Inpainting (Noisy with $\sigma_{\mathbf{y}} = 0.1$)

(d) Colorization (Noisy with $\sigma_{\mathbf{y}} = 0.1$)