Введение в искусственный интеллект. Современное компьютерное зрение

Тема семинара: Генеративные состязательные сети

Бабин Д.Н., Иванов И.Е., Петюшко А.А.

кафедра Математической Теории Интеллектуальных Систем

План семинара

1. Вывод формул!

План семинара

1. Вывод формул!
2. Условные состязательные генеративные сети

План семинара

1. Вывод формул!
2. Условные состязательные генеративные сети
3. Вариационный автоэнкодер

Энтропия для унитарного кодирования

- Пусть $p = (p_1, \ldots, p_n)$, $\quad \sum_{i=1}^{n} p_i = 1, 0 \leq p_i \leq 1.$

Энтропия для унитарного кодирования

- Пусть $p = (p_1, \ldots, p_n), \quad \sum_{i=1}^{n} p_i = 1, 0 \leq p_i \leq 1.$
- **Энтропия** $H(p) = -\sum_{i=1}^{n} p_i \log p_i$

Энтропия для унитарного кодирования

- Пусть $p = (p_1, \ldots, p_n), \quad \sum_{i=1}^{n} p_i = 1, 0 \leq p_i \leq 1.$
- **Энтропия** $H(p) = -\sum_{i=1}^{n} p_i \log p_i$
- Унитарное кодирование для класса y_c:
  $y_{one-hot} = (0, \ldots, 1, \ldots, 0), \quad y_i = 0, i \neq y_c, \quad y_i = 1, i = y_c$

Энтропия для унитарного кодирования

- Пусть $p = (p_1, \ldots, p_n), \quad \sum_{i=1}^{n} p_i = 1, 0 \leq p_i \leq 1.$
- **Энтропия** $H(p) = -\sum_{i=1}^{n} p_i \log p_i$
- Унитарное кодирование для класса y_c:
  $y_{one-hot} = (0, \ldots, 1, \ldots, 0), \quad y_i = 0, i \neq y_c, \quad y_i = 1, i = y_c$
- В этом случае для $i = y_c$: $p_i \log p_i = 1 \cdot \log 1 = 1 \cdot 0 = 0$

Энтропия для унитарного кодирования

- Пусть $p = (p_1, \dots, p_n), \quad \sum_{i=1}^n p_i = 1, 0 \le p_i \le 1.$
- **Энтропия** $H(p) = -\sum_{i=1}^n p_i \log p_i$
- Унитарное кодирование для класса y_c:
  $y_{one-hot} = (0, \dots, 1, \dots, 0), \quad y_i = 0, i \ne y_c, \quad y_i = 1, i = y_c$
- В этом случае для $i = y_c$: $p_i \log p_i = 1 \cdot \log 1 = 1 \cdot 0 = 0$
- А для $i \ne y_c$ воспользуемся правилом Лопиталя: $p_i \log p_i = \lim_{x \to 0} x \log x = \lim_{x \to 0} \frac{\log x}{1/x} = \lim_{x \to 0} \frac{(\log x)'}{(1/x)'} = \lim_{x \to 0} \frac{1/x}{-1/x^2} = -\lim_{x \to 0} x = 0$

Энтропия для унитарного кодирования

- Пусть $p = (p_1, \ldots, p_n), \quad \sum_{i=1}^{n} p_i = 1, 0 \le p_i \le 1.$
- **Энтропия** $H(p) = -\sum_{i=1}^{n} p_i \log p_i$
- Унитарное кодирование для класса y_c:
  $y_{one-hot} = (0, \ldots, 1, \ldots, 0), \quad y_i = 0, i \neq y_c, \quad y_i = 1, i = y_c$
- В этом случае для $i = y_c$: $p_i \log p_i = 1 \cdot \log 1 = 1 \cdot 0 = 0$
- А для $i \neq y_c$ воспользуемся правилом Лопиталя: $p_i \log p_i = \lim_{x \to 0} x \log x = \lim_{x \to 0} \frac{\log x}{1/x} = \lim_{x \to 0} \frac{(\log x)'}{(1/x)'} = \lim_{x \to 0} \frac{1/x}{-1/x^2} = -\lim_{x \to 0} x = 0$
- Таким образом, разобрали все случаи, и $H(p) = -\sum_{i=1}^{n} p_i \log p_i = 0.$

Неравенство Гиббса (1)

- Пусть $p = (p_1, \ldots, p_n)$, $\sum_{i=1}^{n} p_i = 1, 0 \leq p_i \leq 1$,

Неравенство Гиббса (1)

- Пусть $p = (p_1, \ldots, p_n)$, $\sum_{i=1}^{n} p_i = 1, 0 \leq p_i \leq 1$,
- $q = (q_1, \ldots, q_n)$, $\sum_{i=1}^{n} q_i = 1, 0 \leq q_i \leq 1$.

Неравенство Гиббса (1)

- Пусть $p = (p_1, \ldots, p_n)$, $\sum_{i=1}^{n} p_i = 1, 0 \le p_i \le 1$,
- $q = (q_1, \ldots, q_n)$, $\sum_{i=1}^{n} q_i = 1, 0 \le q_i \le 1$.

Неравенство Гиббса

$\sum_{i=1}^{n} p_i \log p_i \ge \sum_{i=1}^{n} p_i \log q_i$ для любых распределений p, q, при этом равенство достигается только при $p = q$.

Доказательство.

- При $x > 0$ имеем $\log x \le x - 1$, причем равенство достигается только при $x = 1$.

Неравенство Гиббса (1)

- Пусть $p = (p_1, \ldots, p_n)$, $\sum_{i=1}^{n} p_i = 1, 0 \leq p_i \leq 1$,
- $q = (q_1, \ldots, q_n)$, $\sum_{i=1}^{n} q_i = 1, 0 \leq q_i \leq 1$.

Неравенство Гиббса

$\sum_{i=1}^{n} p_i \log p_i \geq \sum_{i=1}^{n} p_i \log q_i$ для любых распределений p, q, при этом равенство достигается только при $p = q$.

Доказательство.

- При $x > 0$ имеем $\log x \leq x - 1$, причем равенство достигается только при $x = 1$.
- Пусть I — множество индексов, в которых $p_i > 0$.

Неравенство Гиббса (1)

- Пусть $p = (p_1, \ldots, p_n), \quad \sum_{i=1}^{n} p_i = 1, 0 \leq p_i \leq 1,$
- $q = (q_1, \ldots, q_n), \quad \sum_{i=1}^{n} q_i = 1, 0 \leq q_i \leq 1.$

Неравенство Гиббса

$\sum_{i=1}^{n} p_i \log p_i \geq \sum_{i=1}^{n} p_i \log q_i$ для любых распределений p, q, при этом равенство достигается только при $p = q$.

Доказательство.

- При $x > 0$ имеем $\log x \leq x - 1$, причем равенство достигается только при $x = 1$.
- Пусть I — множество индексов, в которых $p_i > 0$.
- Тогда $\sum_{i \in I} p_i \log q_i - \sum_{i \in I} p_i \log p_i = \sum_{i \in I} p_i \log \frac{q_i}{p_i} \leq \sum_{i \in I} p_i (\frac{q_i}{p_i} - 1) = \sum_{i \in I} q_i - \sum_{i \in I} p_i \leq 1 - \sum_{i \in I} p_i = 1 - 1 = 0$

Неравенство Гиббса (2)

Доказательство.

- В случае $p_i = 0$:
  - $p_i \log p_i = 0$ (см. ранее),

Неравенство Гиббса (2)

Доказательство.

- В случае $p_i = 0$:
  - $p_i \log p_i = 0$ (см. ранее),
  - Тогда $\sum_{i:p_i=0} p_i \log q_i - \sum_{i:p_i=0} p_i \log p_i = \sum_{i:p_i=0} p_i \log q_i \leq \sum_{i:p_i=0} p_i \log 1 = 0.$

Неравенство Гиббса (2)

Доказательство.

- В случае $p_i = 0$:
  - $p_i \log p_i = 0$ (см. ранее),
  - Тогда $\sum_{i:p_i=0} p_i \log q_i - \sum_{i:p_i=0} p_i \log p_i = \sum_{i:p_i=0} p_i \log q_i \le \sum_{i:p_i=0} p_i \log 1 = 0$.
- Таким образом, $\sum_{i=1}^{n} p_i \log p_i \ge \sum_{i=1}^{n} p_i \log q_i$, при этом равенство возможно только при $\frac{q_i}{p_i} = 1$ в случае $i \in I$,

Доказательство.

- В случае $p_i = 0$:
  - $p_i \log p_i = 0$ (см. ранее),
  - Тогда $\sum_{i:p_i=0} p_i \log q_i - \sum_{i:p_i=0} p_i \log p_i = \sum_{i:p_i=0} p_i \log q_i \leq \sum_{i:p_i=0} p_i \log 1 = 0$.
- Таким образом, $\sum_{i=1}^{n} p_i \log p_i \geq \sum_{i=1}^{n} p_i \log q_i$, при этом равенство возможно только при $\frac{q_i}{p_i} = 1$ в случае $i \in I$,
- Но тогда и $p_i = q_i = 0, \quad i \notin I$. ■

Функция близости Кульбака-Лейблера

Определение

Функция близости (divergence) двух распределений, заданных на одном пространстве функций распределения S — это функция $D(\cdot||\cdot) : S \times S \to \mathbb{R}$, такая что:

1. $D(P||Q) \geq 0$ для всех $P, Q \in S$
2. $D(P||Q) = 0 \Leftrightarrow P = Q$

Функция близости Кульбака-Лейблера

Определение

Функция близости (divergence) двух распределений, заданных на одном пространстве функций распределения S — это функция $D(\cdot||\cdot) : S \times S \to \mathbb{R}$, такая что:

1. $D(P||Q) \geq 0$ для всех $P, Q \in S$
2. $D(P||Q) = 0 \Leftrightarrow P = Q$

Функция близости Кульбака-Лейблера

$D_{KL}(P||Q) = \sum_x p(x) \log \frac{p(x)}{q(x)}$

Функция близости Кульбака-Лейблера

Определение

Функция близости (divergence) двух распределений, заданных на одном пространстве функций распределения S — это функция $D(\cdot||\cdot) : S \times S \to \mathbb{R}$, такая что:

1. $D(P||Q) \geq 0$ для всех $P, Q \in S$
2. $D(P||Q) = 0 \Leftrightarrow P = Q$

Функция близости Кульбака-Лейблера

$D_{KL}(P||Q) = \sum_x p(x) \log \frac{p(x)}{q(x)}$

Доказательство. $D_{KL}(P||Q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = \sum_x p(x) \log p(x) - \sum_x p(x) \log q(x) \geq 0$, при этом равенство достигается только при $P = Q$ (следствие из неравенства Гиббса). ∎

Метрика на основе функции близости Йенсена-Шеннона (1)

Функция близости Йенсена-Шеннона

$D_{JS}(P||Q) = \frac{1}{2}D_{KL}(P||M) + \frac{1}{2}D_{KL}(Q||M)$, где $M = \frac{1}{2}(P + Q)$.

Метрика на основе функции близости Йенсена-Шеннона (1)

Функция близости Йенсена-Шеннона

$D_{JS}(P||Q) = \frac{1}{2}D_{KL}(P||M) + \frac{1}{2}D_{KL}(Q||M)$, где $M = \frac{1}{2}(P+Q)$.

Определение метрики

Функция $d : X \times X \to \mathbb{R}$ называется метрикой на пространстве X, если:

1. $d(x,y) = 0 \Leftrightarrow x = y$,
2. $d(x,y) = d(y,x)$,
3. $d(x,z) \leq d(x,y) + d(y,z)$.

Метрика на основе функции близости Йенсена-Шеннона (1)

Функция близости Йенсена-Шеннона

$D_{JS}(P||Q) = \frac{1}{2}D_{KL}(P||M) + \frac{1}{2}D_{KL}(Q||M)$, где $M = \frac{1}{2}(P+Q)$.

Определение метрики

Функция $d : X \times X \to \mathbb{R}$ называется метрикой на пространстве X, если:

1. $d(x,y) = 0 \Leftrightarrow x = y$,
2. $d(x,y) = d(y,x)$,
3. $d(x,z) \leq d(x,y) + d(y,z)$.

Теорема

$\sqrt{D_{JS}(P||Q)}$ — метрика.

Метрика на основе функции близости Йенсена-Шеннона (2)

Доказательство.

1. $\sqrt{D_{JS}(P||Q)} = 0 \Leftrightarrow D_{JS}(P||Q) = 0 \Leftrightarrow D_{KL}(P||M) = D_{KL}(Q||M) = 0, P = Q = \frac{1}{2}(P+Q)$ (из свойств функции близости D_{KL});

[1] Endres D. M., Schindelin J. E. "A new metric for probability distributions". 2003.

Метрика на основе функции близости Йенсена-Шеннона (2)

Доказательство.

1. $\sqrt{D_{JS}(P||Q)} = 0 \Leftrightarrow D_{JS}(P||Q) = 0 \Leftrightarrow D_{KL}(P||M) = D_{KL}(Q||M) = 0, P = Q = \frac{1}{2}(P+Q)$ (из свойств функции близости D_{KL});
2. $\sqrt{D_{JS}(P||Q)} = \sqrt{\frac{1}{2}D_{KL}(P||M) + \frac{1}{2}D_{KL}(Q||M)} = \sqrt{\frac{1}{2}D_{KL}(Q||M) + \frac{1}{2}D_{KL}(P||M)} = \sqrt{D_{JS}(Q||P)}$;

[1] Endres D. M., Schindelin J. E. "A new metric for probability distributions". 2003.

Метрика на основе функции близости Йенсена-Шеннона (2)

Доказательство.

1. $\sqrt{D_{JS}(P||Q)} = 0 \Leftrightarrow D_{JS}(P||Q) = 0 \Leftrightarrow D_{KL}(P||M) = D_{KL}(Q||M) = 0, P = Q = \frac{1}{2}(P+Q)$ (из свойств функции близости D_{KL});
2. $\sqrt{D_{JS}(P||Q)} = \sqrt{\frac{1}{2}D_{KL}(P||M) + \frac{1}{2}D_{KL}(Q||M)} = \sqrt{\frac{1}{2}D_{KL}(Q||M) + \frac{1}{2}D_{KL}(P||M)} = \sqrt{D_{JS}(Q||P)}$;
3. А с неравенством треугольника здесь все действительно сложнее. За деталями — см. дополнительную литературу[1]. ■

[1] Endres D. M., Schindelin J. E. "A new metric for probability distributions". 2003.

GAN и оптимальное распределение D/G (1)

Если обозначить за p_g — распределение $G(z)$, то оригинальную функцию потерь

$$\min_G \max_D V(D, G) = \min_G \max_D [\mathbb{E}_{x \sim p_{data}} \log D(x) + \mathbb{E}_{z \sim p_z} \log(1 - D(G(z)))]$$

можно переписать как

GAN и оптимальное распределение D/G (1)

Если обозначить за p_g — распределение $G(z)$, то оригинальную функцию потерь

$$\min_G \max_D V(D,G) = \min_G \max_D [\mathbb{E}_{x \sim p_{data}} \log D(x) + \mathbb{E}_{z \sim p_z} \log(1 - D(G(z)))]$$

можно переписать как

$$\min_G \max_D V(D,G) = \min_G \max_D [\mathbb{E}_{x \sim p_{data}} \log D(x) + \mathbb{E}_{x \sim p_g} \log(1 - D(x))]$$

GAN и оптимальное распределение D/G (1)

Если обозначить за p_g — распределение $G(z)$, то оригинальную функцию потерь

$$\min_G \max_D V(D,G) = \min_G \max_D [\mathbb{E}_{x\sim p_{data}} \log D(x) + \mathbb{E}_{z\sim p_z} \log(1 - D(G(z)))]$$

можно переписать как

$$\min_G \max_D V(D,G) = \min_G \max_D [\mathbb{E}_{x\sim p_{data}} \log D(x) + \mathbb{E}_{x\sim p_g} \log(1 - D(x))]$$

Теорема

$\arg\max_D V(D,G) = D_G = \frac{p_{data}}{p_{data}+p_g}$.

GAN и оптимальное распределение D/G (1)

Если обозначить за p_g — распределение $G(z)$, то оригинальную функцию потерь

$$\min_G \max_D V(D, G) = \min_G \max_D [\mathbb{E}_{x \sim p_{data}} \log D(x) + \mathbb{E}_{z \sim p_z} \log(1 - D(G(z)))]$$

можно переписать как

$$\min_G \max_D V(D, G) = \min_G \max_D [\mathbb{E}_{x \sim p_{data}} \log D(x) + \mathbb{E}_{x \sim p_g} \log(1 - D(x))]$$

Теорема

$\arg\max_D V(D, G) = D_G = \frac{p_{data}}{p_{data} + p_g}$.

Доказательство. $V(D, G) = \mathbb{E}_{x \sim p_{data}} \log D(x) + \mathbb{E}_{x \sim p_g} \log(1 - D(x)) = \int_x (p_{data}(x) \log D(x) + p_g(x) \log(1 - D(x)))dx$

GAN и оптимальное распределение D/G (2)

Теорема

$\arg\max_D V(D, G) = D_G = \frac{p_{data}}{p_{data}+p_g}$.

Доказательство. Найдем точку экстремума функции $f(s) = p_{data} \log s + p_g \log(1 - s)$:

GAN и оптимальное распределение D/G (2)

Теорема

$\arg\max_D V(D, G) = D_G = \frac{p_{data}}{p_{data}+p_g}$.

Доказательство. Найдем точку экстремума функции $f(s) = p_{data} \log s + p_g \log(1-s)$:

- $f'(s) = \frac{p_{data}}{s} - \frac{p_g}{1-s}$,
- $f'(s) = 0 \Leftrightarrow p_{data}(1-s) - p_g s = 0 \Leftrightarrow s = \frac{p_{data}}{p_{data}+p_g}$.

GAN и оптимальное распределение D/G (2)

Теорема

$\arg\max_D V(D,G) = D_G = \frac{p_{data}}{p_{data}+p_g}$.

Доказательство. Найдем точку экстремума функции $f(s) = p_{data}\log s + p_g \log(1-s)$:

- $f'(s) = \frac{p_{data}}{s} - \frac{p_g}{1-s}$,
- $f'(s) = 0 \Leftrightarrow p_{data}(1-s) - p_g s = 0 \Leftrightarrow s = \frac{p_{data}}{p_{data}+p_g}$.

При этом $s = \frac{p_{data}}{p_{data}+p_g}$ — точка максимума функции $f(s)$ (можно проверить, например, что $\lim_{s\to 0+} f'(s) > 0, \lim_{s\to 1-} f'(s) < 0$).

Теорема

$\arg\max_D V(D,G) = D_G = \frac{p_{data}}{p_{data}+p_g}$.

Доказательство. Найдем точку экстремума функции $f(s) = p_{data}\log s + p_g \log(1-s)$:

- $f'(s) = \frac{p_{data}}{s} - \frac{p_g}{1-s}$,
- $f'(s) = 0 \Leftrightarrow p_{data}(1-s) - p_g s = 0 \Leftrightarrow s = \frac{p_{data}}{p_{data}+p_g}$.

При этом $s = \frac{p_{data}}{p_{data}+p_g}$ — точка максимума функции $f(s)$ (можно проверить, например, что $\lim_{s\to 0+} f'(s) > 0, \lim_{s\to 1-} f'(s) < 0$).

Таким образом, $\arg\max_D V(D,G) = D_G = \frac{p_{data}}{p_{data}+p_g}$. ■

GAN и оптимальное распределение D/G (3)

Пусть $C(G) = \max_D V(D, G)$

Теорема

$\arg\min_G C(G) = p_{data}$, $C(p_{data}) = -\log 4$.

GAN и оптимальное распределение D/G (3)

Пусть $C(G) = \max_D V(D, G)$

Теорема

$\arg\min_G C(G) = p_{data}$, $C(p_{data}) = -\log 4$.

Доказательство.

- $C(G) = \mathbb{E}_{x \sim p_{data}} \log D_G(x) + \mathbb{E}_{x \sim p_g} \log(1 - D_G(x)) = \mathbb{E}_{x \sim p_{data}} \log \frac{p_{data}(x)}{p_{data}(x)+p_g(x)} + \mathbb{E}_{x \sim p_g} \log \frac{p_g(x)}{p_{data}(x)+p_g(x)}$,

GAN и оптимальное распределение D/G (3)

Пусть $C(G) = \max_D V(D, G)$

Теорема

$\arg\min_G C(G) = p_{data}, C(p_{data}) = -\log 4.$

Доказательство.

- $C(G) = \mathbb{E}_{x \sim p_{data}} \log D_G(x) + \mathbb{E}_{x \sim p_g} \log(1 - D_G(x)) = \mathbb{E}_{x \sim p_{data}} \log \frac{p_{data}(x)}{p_{data}(x)+p_g(x)} + \mathbb{E}_{x \sim p_g} \log \frac{p_g(x)}{p_{data}(x)+p_g(x)},$
- $\mathbb{E}_{x \sim p_{data}} \log \frac{p_{data}(x)}{p_{data}(x)+p_g(x)} = \mathbb{E}_{x \sim p_{data}} \log \frac{p_{data}(x)}{\frac{p_{data}(x)+p_g(x)}{2}} - \mathbb{E}_{x \sim p_{data}} \log 2 = D_{KL}(p_{data} || \frac{p_{data}(x)+p_g(x)}{2}) - \log 2,$

GAN и оптимальное распределение D/G (3)

Пусть $C(G) = \max_D V(D, G)$

Теорема

$\arg\min_G C(G) = p_{data}$, $C(p_{data}) = -\log 4$.

Доказательство.

- $C(G) = \mathbb{E}_{x\sim p_{data}} \log D_G(x) + \mathbb{E}_{x\sim p_g} \log(1 - D_G(x)) =$
$\mathbb{E}_{x\sim p_{data}} \log \frac{p_{data}(x)}{p_{data}(x)+p_g(x)} + \mathbb{E}_{x\sim p_g} \log \frac{p_g(x)}{p_{data}(x)+p_g(x)}$,
- $\mathbb{E}_{x\sim p_{data}} \log \frac{p_{data}(x)}{p_{data}(x)+p_g(x)} = \mathbb{E}_{x\sim p_{data}} \log \frac{p_{data}(x)}{\frac{p_{data}(x)+p_g(x)}{2}} - \mathbb{E}_{x\sim p_{data}} \log 2 =$
$D_{KL}(p_{data}||\frac{p_{data}(x)+p_g(x)}{2}) - \log 2$,
- $\mathbb{E}_{x\sim p_g} \log \frac{p_g(x)}{p_{data}(x)+p_g(x)} = D_{KL}(p_g||\frac{p_{data}(x)+p_g(x)}{2}) - \log 2$,

GAN и оптимальное распределение D/G (3)

Пусть $C(G) = \max_D V(D, G)$

Теорема

$\arg\min_G C(G) = p_{data}, C(p_{data}) = -\log 4.$

Доказательство.

- $C(G) = \mathbb{E}_{x\sim p_{data}} \log D_G(x) + \mathbb{E}_{x\sim p_g} \log(1 - D_G(x)) =$
  $\mathbb{E}_{x\sim p_{data}} \log \frac{p_{data}(x)}{p_{data}(x)+p_g(x)} + \mathbb{E}_{x\sim p_g} \log \frac{p_g(x)}{p_{data}(x)+p_g(x)},$
- $\mathbb{E}_{x\sim p_{data}} \log \frac{p_{data}(x)}{p_{data}(x)+p_g(x)} = \mathbb{E}_{x\sim p_{data}} \log \frac{p_{data}(x)}{\frac{p_{data}(x)+p_g(x)}{2}} - \mathbb{E}_{x\sim p_{data}} \log 2 =$
  $D_{KL}(p_{data}||\frac{p_{data}(x)+p_g(x)}{2}) - \log 2,$
- $\mathbb{E}_{x\sim p_g} \log \frac{p_g(x)}{p_{data}(x)+p_g(x)} = D_{KL}(p_g||\frac{p_{data}(x)+p_g(x)}{2}) - \log 2,$
- $\Rightarrow C(G) = D_{KL}(p_{data}||\frac{p_{data}(x)+p_g(x)}{2}) + D_{KL}(p_g||\frac{p_{data}(x)+p_g(x)}{2}) - 2\log 2,$

Теорема

$\arg\min_G C(G) = p_{data}, C(p_{data}) = -\log 4.$

Доказательство.

- $\Rightarrow C(G) = 2D_{JS}(p_{data}||p_g) - \log 4 \geq -\log 4,$

GAN и оптимальное распределение D/G (4)

Теорема

$\arg\min_G C(G) = p_{data}, C(p_{data}) = -\log 4.$

Доказательство.

- $\Rightarrow C(G) = 2D_{JS}(p_{data}||p_g) - \log 4 \geq -\log 4,$
- При этом $C(G) = -\log 4 \Leftrightarrow D_{JS}(p_{data}||p_g) = 0 \Leftrightarrow p_{data} = p_g.$ ■

GAN и оптимальное распределение D/G (4)

Теорема

$\arg\min_G C(G) = p_{data}, C(p_{data}) = -\log 4.$

Доказательство.

- $\Rightarrow C(G) = 2D_{JS}(p_{data}||p_g) - \log 4 \geq -\log 4,$
- При этом $C(G) = -\log 4 \Leftrightarrow D_{JS}(p_{data}||p_g) = 0 \Leftrightarrow p_{data} = p_g.$ ■

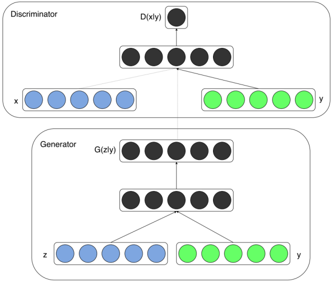
Замечание. Таким образом, генератор G будет пытаться полностью повторять распределение исходных данных: $p_g = p_{data}$.

Условные состязательные генеративные сети

$$\min_G \max_D [\mathbb{E}_{x \sim p_{data}, y \sim p_y} \log D(x, y) + \mathbb{E}_{z \sim p_z, y \sim p_y} \log(1 - D(G(z, y), y))]$$

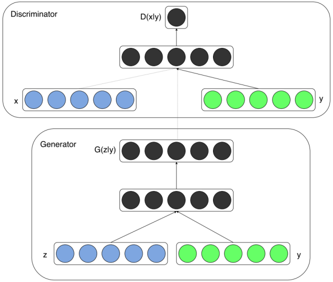
Условные состязательные генеративные сети

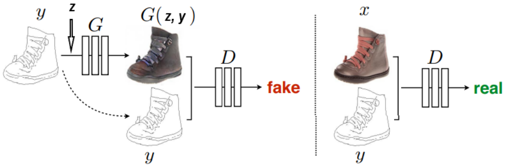
$$\min_G \max_D [\mathbb{E}_{x\sim p_{data}, y\sim p_y} \log D(x,y) + \mathbb{E}_{z\sim p_z, y\sim p_y} \log(1 - D(G(z,y),y))]$$



Условные состязательные генеративные сети

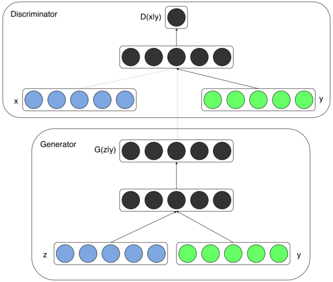
$$\min_G \max_D [\mathbb{E}_{x\sim p_{data}, y\sim p_y} \log D(x,y) + \mathbb{E}_{z\sim p_z, y\sim p_y} \log(1 - D(G(z,y),y))]$$

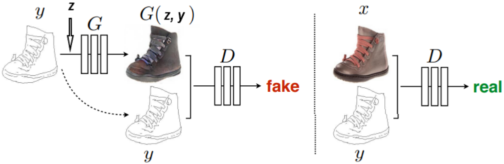




Условные состязательные генеративные сети

$$\min_G \max_D [\mathbb{E}_{x\sim p_{data}, y\sim p_y} \log D(x,y) + \mathbb{E}_{z\sim p_z, y\sim p_y} \log(1 - D(G(z,y),y))]$$



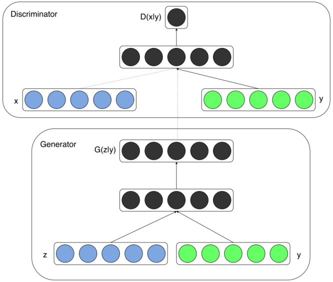


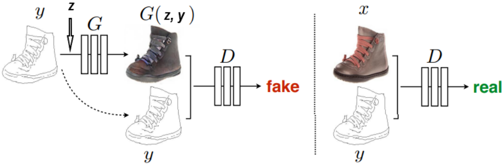
Ссылки для любознательных:

1. Image2Image Translation от Д. Михайлова

Условные состязательные генеративные сети

$$\min_G \max_D [\mathbb{E}_{x\sim p_{data}, y\sim p_y} \log D(x,y) + \mathbb{E}_{z\sim p_z, y\sim p_y} \log(1 - D(G(z,y),y))]$$





Ссылки для любознательных:

1. Image2Image Translation от Д. Михайлова
2. Image Generation от А. Иванюты

Вариационный автоэнкодер

Сначала реальные данные отображаются в простое распределение (например, $z \sim N(\mu, \sigma^2)$) с помощью энкодера Enc, после чего отображаются обратно с помощью декодера Dec. Dec и Enc — нейросети, обучающиеся совместно. При генерации порождается $z \sim N(\mu, \sigma^2)$ и используется только $Dec(z)$.

[2] Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." 2013.

Вариационный автоэнкодер

Сначала реальные данные отображаются в простое распределение (например, $z \sim N(\mu, \sigma^2)$) с помощью энкодера Enc, после чего отображаются обратно с помощью декодера Dec. Dec и Enc — нейросети, обучающиеся совместно. При генерации порождается $z \sim N(\mu, \sigma^2)$ и используется только $Dec(z)$.

Замечания[2]:

- Если просто генерировать случайное $z \sim N(\mu, \sigma^2)$, то не сможем обучать сеть (нет возможности провести обратное распространение ошибки!)

[2] Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." 2013.

Вариационный автоэнкодер

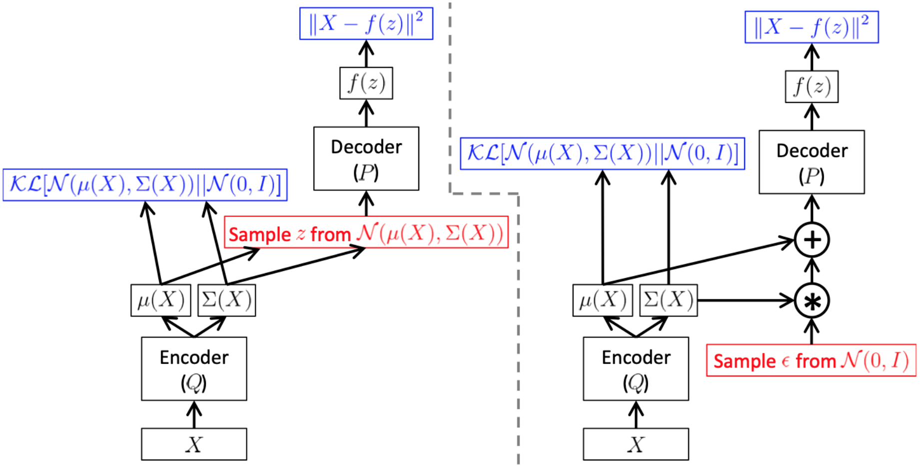
Сначала реальные данные отображаются в простое распределение (например, $z \sim N(\mu, \sigma^2)$) с помощью энкодера Enc, после чего отображаются обратно с помощью декодера Dec. Dec и Enc — нейросети, обучающиеся совместно. При генерации порождается $z \sim N(\mu, \sigma^2)$ и используется только $Dec(z)$.

Замечания[2]:

- Если просто генерировать случайное $z \sim N(\mu, \sigma^2)$, то не сможем обучать сеть (нет возможности провести обратное распространение ошибки!)
- Применяется так называемый репараметризационный трюк (reparametrization trick): генерируется случайное значение $z \sim N(0, 1)$, после чего линейно домножается на σ и прибавляется $\mu \Rightarrow$ можно пропускать градиенты по σ и μ!

[2] Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." 2013.

Вариационный автоэнкодер — схема



Замечание: красным обозначены места без backprop, синим — функции потерь.

Время для вопросов

Спасибо за внимание!