Deep Variational Canonical Correlation Analysis

Ольга Гребенькова

Байесовское мультимоделирование

27 апреля 2022 г.

О чем пойдет речь?

Родственная задача transfer learning

Multiview learning representation — обучение с несколькими представлениями(типами измерений) одного и того же базового сигнала, и цель состоит в том, чтобы изучить полезные свойства каждого представления. Изученные свойства должны раскрывать общие и различные черты в представлениях, что может быть полезно для исследовательского анализа или последующих задач.

Классические подходы

Канонический корреляционный анализ (1936) и его нелинейные варианты, включая ядерные (kernel) варианты (2000-2003), глубокие нейронные сети (2013)

Идея

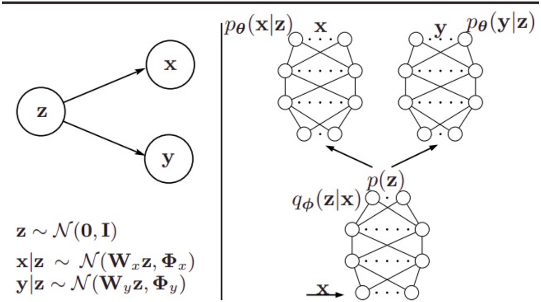
Давайте попробуем использовать подход вариационного автокодировщика.

- Пусть два представления $\mathbf{x}$ и $\mathbf{y}$ независимы. $\mathbf{z}$ – скрытая переменная.
- Совместное распределение:

$$p(\mathbf{x}, \mathbf{y}, \mathbf{z}) = p(\mathbf{z})p(\mathbf{x}|\mathbf{z})p(\mathbf{y}|\mathbf{z})$$

$$p(\mathbf{x}, \mathbf{y}) = \int p(\mathbf{x}, \mathbf{y}, \mathbf{z}) d\mathbf{z}.$$

- Если $p(\mathbf{x}|\mathbf{z})$ и $p(\mathbf{y}|\mathbf{z})$ линейные, получаем схему слева и слабые репрезентативные возможности модели.
- Параметризуем модели $p_\theta(\mathbf{x}|\mathbf{z}; \theta_x)$ и $p_\theta(\mathbf{y}|\mathbf{z}; \theta_y)$. Здесь θ_x, θ_y — веса некоторых DNN.



- Маргинальное правдоподобие $p_\theta(\mathbf{x}, \mathbf{y})$ и $p_\theta(\mathbf{z}|\mathbf{x})$ для нахождения скрытого представления по одному из представлений объекта невычислимы в общем случае.
- Введем
$$q_\phi(\mathbf{z}|\mathbf{x}; \phi_z) \sim p_\theta(\mathbf{z}|\mathbf{x})$$
- ELBO для логарифма правдоподобия:
$$\log p_\theta(\mathbf{x}, \mathbf{y}) \geq L(\mathbf{x}, \mathbf{y}; \theta, \phi) :=$$
$$-D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p(\mathbf{z})) + \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{y}|\mathbf{z})]$$
- VCCA максимизирует вариационную нижнюю оценку логарифма правдоподобия:
$$\max_{\theta, \phi} \frac{1}{N} \sum_{i=1}^{N} L(\mathbf{x}_i, \mathbf{y}_i; \theta, \phi).$$

$$L(\mathbf{x}, \mathbf{y}; \theta, \phi) = D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})||p(\mathbf{z})) + \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{y}|\mathbf{z})]$$

- Для того, чтобы слагаемое с D_{KL} было вычислимо, ограничим $q_\phi(\mathbf{z}|\mathbf{x}; \phi_z)$.
- Для пары семплов $(\mathbf{x}_i, \mathbf{y}_i)$:

$$\log q_\phi(\mathbf{z}_i|\mathbf{x}_i) = \log \mathcal{N}(\mathbf{z}_i; \mu_i, \Sigma_i), \quad \Sigma_i = diag(\sigma_{i1}^2, \ldots, \sigma_{i,d_z}^2)$$

  Здесь $[\mu_i, \Sigma_i] = f(\mathbf{x}_i; \phi_z)$, то есть выход DNN энкодера.
- В этом случае

$$D_{KL}(q_\phi(\mathbf{z}_i|\mathbf{x}_i)||p(\mathbf{z}_i)) = -\frac{1}{2}\sum_{j=1}^{d_z}(1 + \log \sigma_{ij}^2 - \sigma_{ij}^2 - \mu_{ij}^2)$$

- Второе слагаемое будем находим с помощью метода Монте-Карло:

$$\mathbb{E}_{q_\phi(\mathbf{z}_i|\mathbf{x}_i)}[\log p_\theta(\mathbf{x}_i|\mathbf{z}_i) + \log p_\theta(\mathbf{y}_i|\mathbf{z}_i)] \approx \frac{1}{L}\sum_{l=1}^{L} \log p_\theta\left(\mathbf{x}_i|\mathbf{z}_i^l\right) + \log p_\theta\left(\mathbf{y}_i|\mathbf{z}_i^l\right)$$

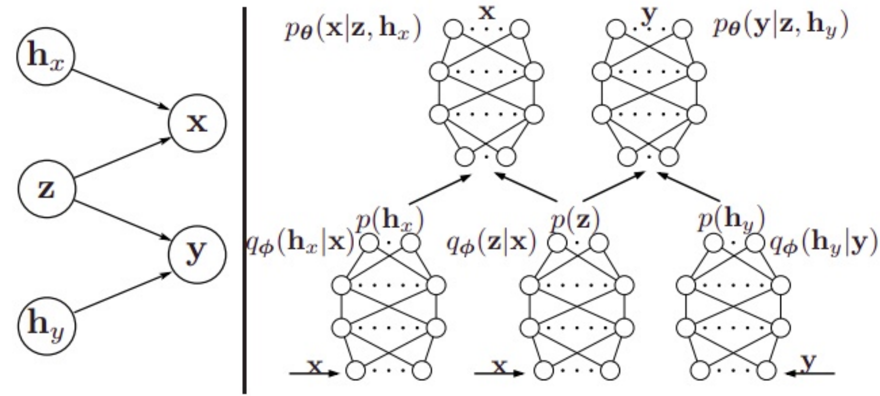


Figure 2. VCCA-private: variational CCA with view-specific private variables.

Эксперимент

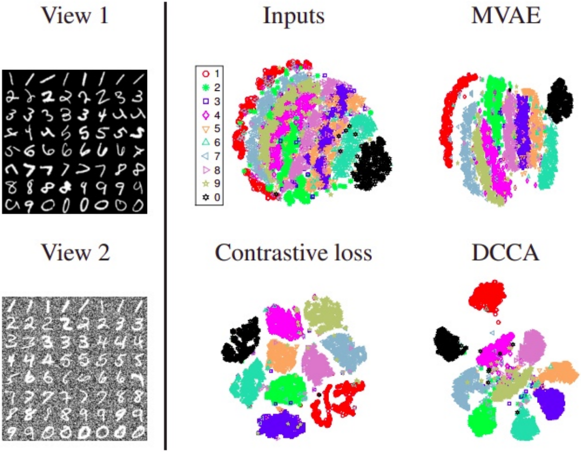


Figure 3. Left: Selection of view 1 images (top) and their corresponding view 2 images (bottom) from noisy MNIST. Right: 2D t-SNE visualization of features learned by previous methods.

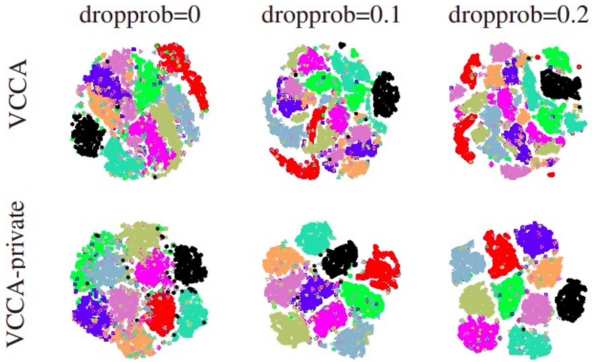


Figure 4. 2D t-SNE visualizations of the extracted shared variables **z** on noisy MNIST test data by VCCA (top row) and VCCA-private (bottom row) for different dropout rates. Here $d_z = 40$.

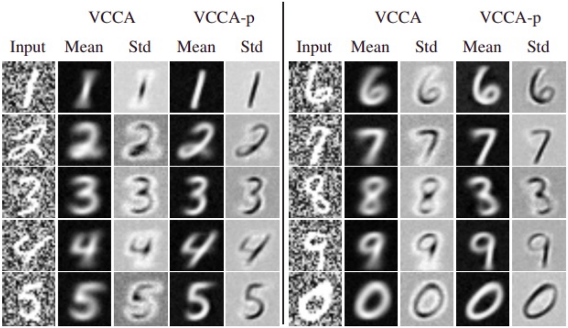


Figure 5. Sample reconstruction of view 2 images from the noisy MNIST test set by VCCA and VCCA-private.

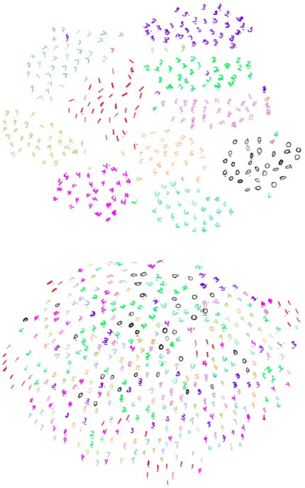


Figure 6. 2D t-SNE embedding of the shared variables $\mathbf{z} \in \mathbb{R}^{40}$ (top) and private variables $\mathbf{h}_x \in \mathbb{R}^{30}$ (bottom).