

# Функциональный анализ данных

Кафедра Интеллектуальных Систем

# Оглавление

<b>I</b>	<b>Multimodal data</b>	<b>5</b>
<b>1</b>	<b>Боева</b>	<b>7</b>
1.1	Введение . . . . .	7
1.2	Определение связности VCCA с нормализующим потоком . . . . .	7
1.3	Выводы . . . . .	9
<b>II</b>	<b>Continuous time and space models</b>	<b>11</b>
<b>III</b>	<b>Physics-informed models</b>	<b>13</b>
<b>IV</b>	<b>Riemannian spaces</b>	<b>15</b>
<b>V</b>	<b>Multilinear models</b>	<b>19</b>



Часть I

Multimodal data



# Глава 1

## Воева

### 1.1 Введение

Канонический корреляционный анализ (ССА) — это классический статистический метод обнаружения скрытой ковариации, лежащей в основе двух или более наблюдаемых случайных векторов. Мы предлагаем новый мультимодальный вариационный автоэнкодер, который позволяет генерировать данные на основе совместного распределения и условно для любого числа сложных модальностей. Векторные представления канонического корреляционного анализа, которые сохраняют общую информацию в разных модальностях, что приводит к более согласованным кросс-модальным генерациям. Объединяя идею VCCA и нормализационных потоков, мы создаем мощный инструмент для моделирования сложных мультимодальных данных. Нормализационные потоки позволяют нам эффективно моделировать сложные распределения данных, обеспечивая точное и гибкое представление совместного распределения. В сочетании с VCCA, который использует вариационные методы для аппроксимации сложных распределений, мы можем достичь высокой точности в генерации данных и улучшенной интерпретации скрытых структур.

### 1.2 Определение связности VCCA с нормализующим потоком

**Формулировка DCCA** Эта порождающая интерпретация ССА часто теряется в нелинейных интерпретациях ССА. Например, в Deep CCA (DCCA, (Andrew et al., 2013)), чтобы расширить ССА до нелинейных отображений с большей мощностью представления, из исходных данных каждого вида извлекаются нелинейные объекты,  $f$  для  $x$  и  $g$  для  $y$ , так что каноническая корреляция выходных данных DNN (измеренная линейным ССА с проекционными матрицами  $U$  и  $V$ ) максимизируется. Формально, учитывая набор данных из  $N$  пар наблюдений  $(x_1, y_1), \dots, (x_N, y_N)$  случайных векторов  $(x, y)$ , DCCA оптимизирует

$$\max_{W_f, W_g, U, V} \text{tr}(U^T f(X)g(Y)^T V)$$

при условии соблюдения

$$U^T f(X)f(X)^T U = V^T g(Y)g(Y)^T V = NI,$$

где  $f(X) = [f(x_1), \dots, f(x_N)]$  и  $g(Y) = [g(y_1), \dots, g(y_N)]$  и  $W_f$  обозначают все весовые параметры DNN  $f$  (и аналогично для  $W_g$ ).

**Формулировка VCCA** Вариационная модель скрытых переменных ССА (Bach и Jordan, 2005) определяет следующее совместное распределение по случайным переменным  $(x, y)$ :

$$p(x, y, z) = p(z)p(x|z)p(y|z), \quad p(x, y) = \int p(x, y, z)dz. \quad (2)$$

Основное предположение этой модели заключается в том, что при условии скрытых переменных  $z \in \mathbb{R}^d$ ,  $x$  и  $y$  независимы. Однако линейные модели наблюдений  $(p(x|z)$  и  $p(y|z)$ , как показано на рисунке 1 (слева)) имеют ограниченную репрезентативную силу. В этой статье мы рассматриваем нелинейные модели наблюдений  $p_\theta(x|z; \theta_x)$  и  $p_\theta(y|z; \theta_y)$ , параметризованные  $\theta_x$  и  $\theta_y$  соответственно, которые могут быть коллекциями весов DNN. В этом

случае маргинальная правдоподобность  $p_\theta(x, y)$  не имеет замкнутой формы. Кроме того, задача вывода  $p_\theta(z|x)$  — задача вывода скрытых переменных при одной из точек зрения — также неразрешима.

Вдохновленные работой Kingma и Welling (2014) по вариационным автоэнкодерам (VAE), мы аппроксимируем  $p_\theta(z|x)$  с помощью условной плотности  $q_\varphi(z|x; \varphi_z)$ , где  $\varphi_z$  — это коллекция параметров другой DNN. Мы можем вывести нижнюю границу для маргинальной правдоподобности данных, используя  $q_\varphi(z|x)$ :

$$\begin{aligned} \log p_\theta(x, y) &= \log \int p_\theta(x, y) q_\varphi(z|x) dz = \int \log p_\theta(x, y) q_\varphi(z|x) dz \\ &= \int q_\varphi(z|x) (\log q_\varphi(z|x) - \log p_\theta(z|x, y) + \log p_\theta(x, y, z) - \log q_\varphi(z|x)) dz \\ &= D_{KL}(q_\varphi(z|x) || p_\theta(z|x, y)) + \mathbb{E}_{q_\varphi(z|x)} [\log p_\theta(x, y, z) - \log q_\varphi(z|x)] \\ &\geq \mathbb{E}_{q_\varphi(z|x)} [\log p_\theta(x, y, z) - \log q_\varphi(z|x)] =: L(x, y; \theta, \varphi) \quad (3) \end{aligned}$$

где мы использовали факт, что KL-дивергенция неотрицательна на последнем шаге. В результате  $L(x, y; \theta, \varphi)$  является нижней границей для логарифма правдоподобности данных  $\log p_\theta(x, y)$ . Подставляя (2) в (3), мы получаем

$$\begin{aligned} L(x, y; \theta, \varphi) &= \int q_\varphi(z|x) (\log p(z) - \log q_\varphi(z|x) + \log p_\theta(x|z) + \log p_\theta(y|z)) dz \\ &= -D_{KL}(q_\varphi(z|x) || p(z)) + \mathbb{E}_{q_\varphi(z|x)} [\log p_\theta(x|z) + \log p_\theta(y|z)]. \quad (4) \end{aligned}$$

VCCA максимизирует эту вариационную нижнюю границу для правдоподобности данных на обучающем наборе:

$$\max_{\theta, \varphi} \frac{1}{N} \sum_{i=1}^N L(x_i, y_i; \theta, \varphi).$$

**Определение нормализующего потока в общем виде** Нормализующие потоки преобразуют простое базовое распределение  $p_Z(\mathbf{z})$  в сложное целевое распределение  $p_X(\mathbf{x})$ , используя серию обратимых и дифференцируемых преобразований  $f_i$ :

$$\mathbf{x} = f_n \circ f_{n-1} \circ \dots \circ f_1(\mathbf{z})$$

Плотность вероятности  $\mathbf{x}$  вычисляется с использованием изменения переменных:

$$p_X(\mathbf{x}) = p_Z(\mathbf{z}) \left| \det \left( \frac{\partial f^{-1}(\mathbf{x})}{\partial \mathbf{x}} \right) \right|$$

или эквивалентно:

$$\log p_X(\mathbf{x}) = \log p_Z(\mathbf{z}) - \sum_{i=1}^n \log \left| \det \left( \frac{\partial f_i}{\partial \mathbf{h}_{i-1}} \right) \right|$$

где  $\mathbf{h}_i = f_i(\mathbf{h}_{i-1})$  и  $\mathbf{h}_0 = \mathbf{z}$

**Связь с нормализующими потоками** Чтобы правильно оценить истинные унимодальные апостериорные распределения  $(p_i(z|x_i))_{i \in [1, m]}$ , унимодальные апостериоры  $(q_{\varphi_i}(z|x_i))_{i \in [1, m]}$  требуют большой гибкости. Однако эти распределения часто выбираются как многомерные гауссовы, что может быть слишком строгим классом распределений и может привести к несогласованности в условных распределениях. В нашей модели мы предлагаем устранить это ограничение, расширив эти распределения с использованием нормализующих потоков. Более точно выражение для унимодального распределения записывается следующим образом:

$$\ln q_{\varphi_i}(z_K|x_i) = \ln q_{\varphi_i}^{(0)}(z_0|x_i) - \sum_{k=1}^K \ln \left| \det \frac{\partial f_{\varphi_i}^{(k)}}{\partial z_{k-1}} \right|,$$

где  $q_{\varphi_i}^{(0)}(z_0|x_i)$  — простое параметризованное распределение, параметры которого задаются нейронными сетями, и  $(f_{\varphi_i}^{(k)})$  являются нормализующими потоками. На практике мы используем многомерные нормальные распределения с

диагональной ковариацией для  $(q_{\varphi_i}^{(0)})_{i \in [1, m]}$ . Поскольку  $q_{\varphi}(z|X)$  является фиксированным на данном этапе обучения, эту задачу можно переписать как:

$$L_{JM}(X) = - \sum_{i=1}^m \mathbb{E}_{q_{\varphi}(z|X)} (\ln q_{\varphi_i}(z|x_i)). \quad (1.1)$$

Для  $i \in [1, m]$  математическое ожидание внутри суммы может быть оценено с использованием выборок из совместного  $q_{\varphi}(z|X)$  и вычислением плотности  $\ln q_{\varphi_i}(z|x_i)$  для этих выборок. Поскольку нам нужно выполнить оценку плотности только для  $q_{\varphi_i}(z|x_i)$  во время обучения, мы решили использовать замаскированные потоки авторегрессии (MAF) (Paragkar et al., 2017), которые позволяют эффективно вычислять его. Уравнение (9) показывает, что во время обучения унимодальные кодировщики получают информацию от объединенного кодировщика: для каждого образца  $X = (x_1, \dots, x_m)$ , скрытая переменная  $z$  выбирается из  $q_{\varphi}(z|X)$ , а затем для каждого  $i \in [1, m]$  вероятность  $q_{\varphi_i}(z|x_i)$  максимизируется. Интересно, что, интегрируя уравнение (1.1) для всего обучающего набора, мы можем показать, что  $q_{\varphi_i}(z|x_i)$  поощряется, чтобы быть близким к среднему распределению  $q_{\text{avg}}(z|x_i) = \mathbb{E}_{\hat{p}((x_j)_{j \neq i}|x_i)}(q_{\varphi}(z|X))$ , где  $\hat{p}$  - наблюдаемое эмпирическое распределение данных.

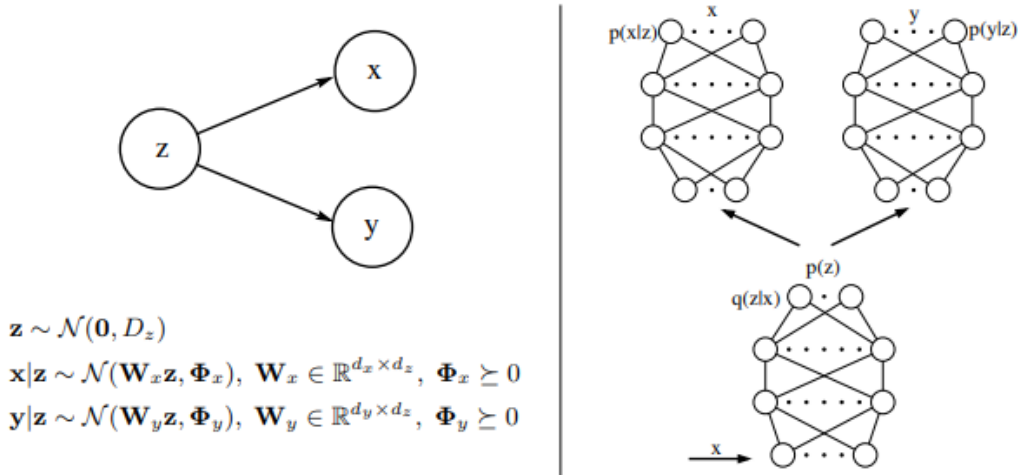


Figure 1: Left: Probabilistic interpretation of CCA (Bach and Jordan, 2005). Right: The deep variational CCA (VCCA) model.

Рис. 1.1. Слева: Вероятностная интерпретация ССА. Справа: Модель глубокого вариационного СОС (VCCA).

### 1.3 Выводы

В данной работе мы представили новый мультимодальный вариационный автоэнкодер, который объединяет идеи канонического корреляционного анализа (ССА) и нормализационных потоков. Этот подход позволяет эффективно моделировать сложные мультимодальные данные, обеспечивая высокую точность генерации данных и улучшенную интерпретацию скрытых структур.

В будущем планируется продолжить исследования в области мультимодальных данных, включая разработку новых методов для улучшения интерпретации скрытых структур и повышения эффективности моделирования сложных распределений.





## Часть II

# Continuous time and space models



## Часть III

# Physics-informed models



Часть IV

Riemannian spaces



blablabla





Часть V

## Multilinear models

