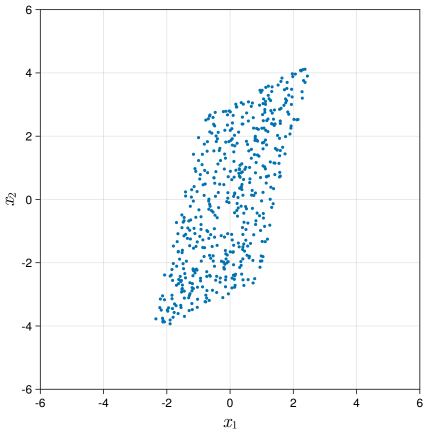
These are exercises for self-study and exam preparation. All material is examinable unless otherwise mentioned.

### Exercise 1. *Generative model of variational autoencoders*

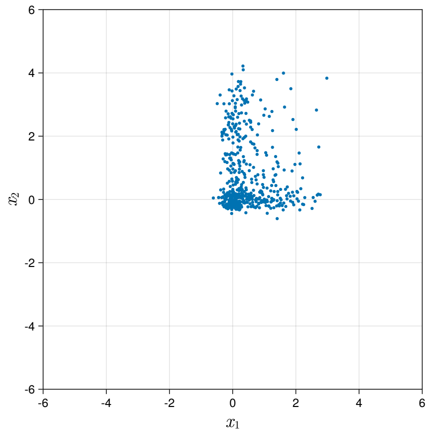
Factor analysis, (noisy) independent component analysis, and variational auto-encoders have the same directed graphical model. However, the three models make different distributional assumptions which leads to models with markedly different properties which we here investigate.

(a) Consider the four scatter plots below that show observed two-dimensional data $(x_1, x_2)$. For each plot indicate the simplest model, from the four listed below, that may have generated the data. Each of the four models should be used once:

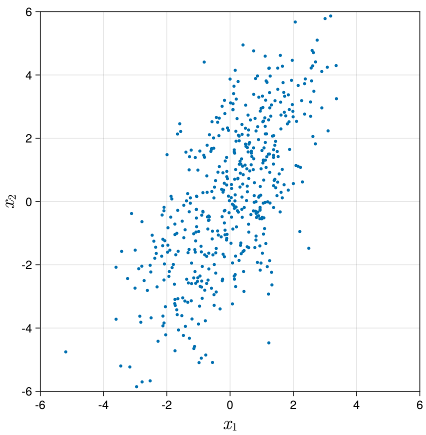
- a factor analysis model
- a Gaussian VAE
- a noise-free square ICA model with super-Gaussian sources
- a noise-free square ICA model with sub-Gaussian sources



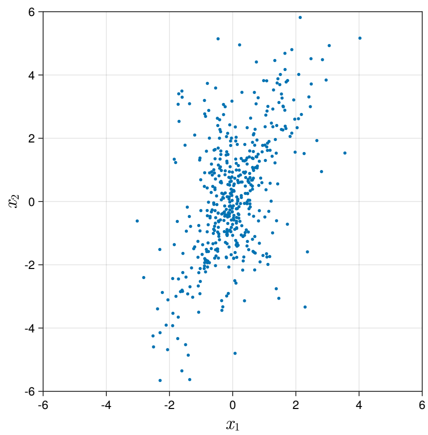

(a)




(b)




(c)




(d)

(b) Briefly explain under which assumptions a Gaussian VAE model as discussed in the course becomes a factor analysis model.

**Exercise 2.** ***ELBO of variational autoencoders***

The general expression for the ELBO for parameter estimation in presence of unobserved variables is

$$\mathcal{L}_{\mathcal{D}}(\boldsymbol{\theta}, q) = \mathbb{E}_{q(\mathbf{h}|\mathcal{D})}\left[\log \frac{p(\mathcal{D}, \mathbf{h}; \boldsymbol{\theta})}{q(\mathbf{h}|\mathcal{D})}\right] \tag{1}$$

where $\mathcal{D}$ are the observed data, $\mathbf{h}$ the unobserved variables, $\boldsymbol{\theta}$ the parameters of the model for the observed and unobserved variables, and $q(\mathbf{h}|\mathcal{D})$ is the variational distribution.

In the case of some VAEs, the above ELBO becomes

$$\begin{aligned} J(\boldsymbol{\theta}, \boldsymbol{\phi}) = &\sum_{i=1}^{n}\sum_{k=1}^{d} \mathbb{E}_{q_{\boldsymbol{\phi}}(\mathbf{h}|\mathbf{v}_i)}\left[v_{ik} \log p_k(\mathbf{h}) + (1 - v_{ik}) \log(1 - p_k(\mathbf{h}))\right] + \\ &+ \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{H}\left(1 + \log(\sigma_j^2(\mathbf{v}_i)) - \sigma_j^2(\mathbf{v}_i) - \mu_j^2(\mathbf{v}_i)\right) \end{aligned} \tag{2}$$

where the $p_k(\mathbf{h})$ are some functions parameterised by $\boldsymbol{\theta}$ that map $\mathbf{h}$ to $[0, 1]$, $\sigma_j(\mathbf{v})$ and $\mu_j(\mathbf{v})$ are some functions parameterised by $\boldsymbol{\phi}$, and the $\mathbf{v}_i = (v_{i1}, \ldots, v_{id})$ are the observed data points. Variable $\mathbf{h}$ is $H$ dimensional.

(a) State two independence assumptions that we made for $p(\mathcal{D}, \mathbf{h}; \boldsymbol{\theta})$ when deriving Equation (2) from Equation (1).

(b) Would you use a VAE trained with the ELBO in Equation (2) for real-valued or binary data? Justify your answer.

(c) Equation (2) uses amortised variational inference. Briefly explain what amortisation refers to in the context of VAEs and state a reason why we may use it.