Exercises Week 5

Advanced Machine Learning (02460)
Technical University of Denmark
Søren Hauberg

February 2024
(Version 1.0)

1 Theoretical exercises

Exercise 5.1 Probabilistic PCA can be equivalently expressed as

$$p(\mathbf{y}|\mathbf{x}) = \mathcal{N}(\mathbf{y}|\mathbf{A}\mathbf{x} + \mathbf{b}, \sigma^2\mathbf{I}) \tag{1}$$

(with latent variables $\mathbf{x}$) or

$$p(\mathbf{y}|\hat{\mathbf{x}}) = \mathcal{N}(\mathbf{y}|\hat{\mathbf{A}}\hat{\mathbf{x}} + \mathbf{b}, \sigma^2\mathbf{I}) \tag{2}$$

(with latent variables $\hat{\mathbf{x}} = \mathbf{R}\mathbf{x}$ and $\hat{\mathbf{A}} = \mathbf{A}\mathbf{R}^\intercal$, where $\mathbf{R}$ is a rotation matrix). The two latent representations $\mathbf{x}$ and $\hat{\mathbf{x}}$ are clearly different (for $\mathbf{R} \neq \mathbf{I}$). *Show that*

1. *Euclidean distances between points in these representations are identical, i.e.* $\|\mathbf{x}_i - \mathbf{x}_j\| = \|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j\|$*, and*
2. *angles between points in these representations are identical, i.e.* $\angle(\mathbf{x}_i - \mathbf{x}_j, \mathbf{x}_k - \mathbf{x}_j) = \angle(\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j, \hat{\mathbf{x}}_k - \hat{\mathbf{x}}_j)$.

Exercise 5.2 Consider a generative model $\mathbf{y} = f(\mathbf{x})$, where $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ and $\mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$. Let

$$f(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ f_3(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} 2x_1^2 + x_2^2 \\ x_1 \\ x_2 \end{pmatrix}. \tag{13}$$

1. *Derive the Jacobian matrix of* f.
2. *Show that the generative model spans an* immersed *manifold.*
3. *Show that the generative model spans an* embedded *manifold.*

Exercise 5.3 Consider the curve $c : [0, 1] \to \mathbb{R}^2$ defined as

$$c(t) = \begin{pmatrix} 2t + 1 \\ -t^2 \end{pmatrix}. \tag{17}$$

1. *Derive an expression of the speed function* $t \mapsto \|\dot{c}_t\|$.
2. *Compute the Euclidean length of the curve. Hint:*

$$\int \sqrt{1 + t^2} \mathrm{d}t = \frac{1}{2} \left(\sqrt{1 + t^2}\; t \sinh^{-1}(t) \right) + C, \tag{18}$$

for some constant C.

2 Programming exercises

The main programming task of this week is to get a sense of how to measure curve lengths. Having an intuition of this is key to understanding the geometry of latent variable models.

Exercise 5.4 Consider the curve $c : [0, 1] \to \mathbb{R}^2$ defined in Eq. 17.

1. *Write a computer program that evaluates the length of the curve c using Eq. 4.2 in the DGGM book.*
2. *If you have completed exercise 5.3:*
   (a) *Did the numerical and the analytical results agree?*
   (b) *Use the analytic expression for* $\|\dot{c}_t\|$ *to write a computer program that evaluates the length of the curve using Eq. 4.5 in the DGGM book. Note that you need to approximate the integral with a sum.*

Exercise 5.5 Consider the Bernoulli VAE that you worked on in Week 2. Train this with a two-dimensional latent space (for ease of plotting).

1. *Write a computer program that evaluates the length of any latent second-order polynomial curve c using Eq. 4.2 in the DGGM book. It is recommended that you write the code to support any callable curve c.*