INFR10086 Machine Learning (MLG) Semester 2, 2025/6

Exercises 1

Lecturer: Hao Tang

Exercise 1. Check that the mean of a 1D Gaussian distribution

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) \tag{1}$$

is μ by showing that $\mathbb{E}[x] = \mu$.

Exercise 2. Show that the log of a multivariate Gaussian distribution

$$\frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) \tag{2}$$

is given as

$$-\frac{1}{2}\mathbf{x}^\top \Sigma^{-1}\mathbf{x} + \boldsymbol{\mu}^\top \Sigma^{-1}\mathbf{x} - \frac{1}{2}\boldsymbol{\mu}^\top \Sigma^{-1}\boldsymbol{\mu} - \frac{1}{2}\log|\Sigma| - \frac{d}{2}\log 2\pi. \tag{3}$$

Exercise 3. Show that a covariance matrix $\Sigma = \mathbb{E}[(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^\top]$ is symmetric and positive semidefinite. A matrix A is positive semidefinite if $\mathbf{x}^\top A\mathbf{x} \geq 0$ for all $\mathbf{x}$.

Exercise 4. In this question, we will look at why the contour of Gaussian distributions consists of ellipses. A general definition of an ellipse (or an ellipsoid in high dimensions) can be written as

$$\left\{\mathbf{x} \in \mathbb{R}^d \,\middle|\, (\mathbf{x}-\mathbf{v})^\top A(\mathbf{x}-\mathbf{v}) = 1\right\}, \tag{4}$$

where $\mathbf{v}$ is where the ellipse is centered and A is a symmetric and positive definite matrix. A matrix A is positive definite if $\mathbf{x}^\top A\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$.

- An axis-aligned ellipse in 2D can be written as

$$\left\{(x, y) \in \mathbb{R}^2 \,\middle|\, \frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{b^2} = 1\right\}, \tag{5}$$

where a and b are the lengths of the two axes and (x_0, y_0) is where the ellipse is centered. Show that equation (5) can be written as equation (4).

- A contour of Gaussian distribution consists of lines where the distribution has the same value. Show that

$$\left\{ \mathbf{x} \in \mathbb{R}^d \,\middle|\, \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu}) \right) = c \right\} \tag{6}$$

is an ellipse for some constant $c > 0$ by rewriting it as equation (4), assuming that the covariance matrix is positive definite.

Exercise 5. In a classification setting, where $x \in \mathbb{R}^d$, $y \in \{1, \ldots, K\}$, and K is the number of classes, show that

$$p(y|x) = \frac{p(x|y)p(y)}{\sum_{y'=1}^{K} p(x|y')p(y')}. \tag{7}$$

Exercise 6. Consider binary classification with the linear classifier

$$y(\mathbf{x}) = \begin{cases} +1 & \text{if } \mathbf{w}^\top \mathbf{x} + b \geq 0 \\ -1 & \text{otherwise} \end{cases} \tag{8}$$

where $\mathbf{x} \in \mathbb{R}^d$, $\mathbf{w}$ is the weight vector, and b is the bias.

- Show that the decision boundary is a straight line when $d = 2$. A line in 2D can be expressed as $y = ax + b$ for some constant $a \in \mathbb{R}$ and $b \in \mathbb{R}$.
- Show that the weight vector $\mathbf{w}$ is a normal vector of the decision boundary.

Exercise 7. Derive a formula for the Euclidean distance between the origin $(0, 0)$ and a line $y = ax + b$, where a and b are arbitrary constants.

Exercise 8. Consider the 2D case for the linear classifier in equation (8). Suppose that the points $(-2, -3)$ and $(4, 1)$ are on the decision bounary and that the point $(2, -3)$ lies in the -1 class region. Find the parameters $(\mathbf{w}, b)$ of the classifier.