Intelligent Learning and Analysis Systems: Machine Learning — Exercise 9

Yauhen Selivonchyk
Łukasz Segiet
Duc Duy Pham
Dominik Schindler

January 29, 2015

1 Linear Separability

Problem (i)

Let

$$\phi : \mathbb{R} \to \mathbb{R},\ x \mapsto \phi(x) = 1 - \left(2\left(x - \frac{1}{2}\right)\right)^2 = 4x - 4x^2.$$

Then

$$\phi(x) \geq 0 \quad \forall x \in S^+$$
$$\phi(x) < 0 \quad \forall x \in S^-.$$

Problem (ii)

The set B is given by a single constraint $\left(x^2 + y^2 + z^2\right) \leq 1$. The very same constraint can be used to find a transformation ϕ with $N = 1$. Thus, let

$$\phi : \mathbb{R}^3 \to \mathbb{R},\ (x, y, z) \mapsto \phi(x, y, z) = 1 - \left(x^2 + y^2 + z^2\right).$$

Then

$$\phi(x, y, z) \geq 0 \quad \forall x \in S^+$$
$$\phi(x, y, z) < 0 \quad \forall x \in S^-.$$

2 Ridge Regression

The values $\lambda \in \left\{0, 10^{-8}, 10^{-6}, 10^{-4}, 10^{-2}, 1\right\}$ were chosen for the ridge regression task. Figure 1 shows the corresponding plots. The lowest $\mathcal{L}^2$ error norm is achieved for $\lambda = 10^{-6}$. Therefore, if the examples are assumed to be uniformly distributed on $[0, 1]$, the best generalization (with respect to squared-error loss function) is achieved for $\lambda = 10^{-6}$.

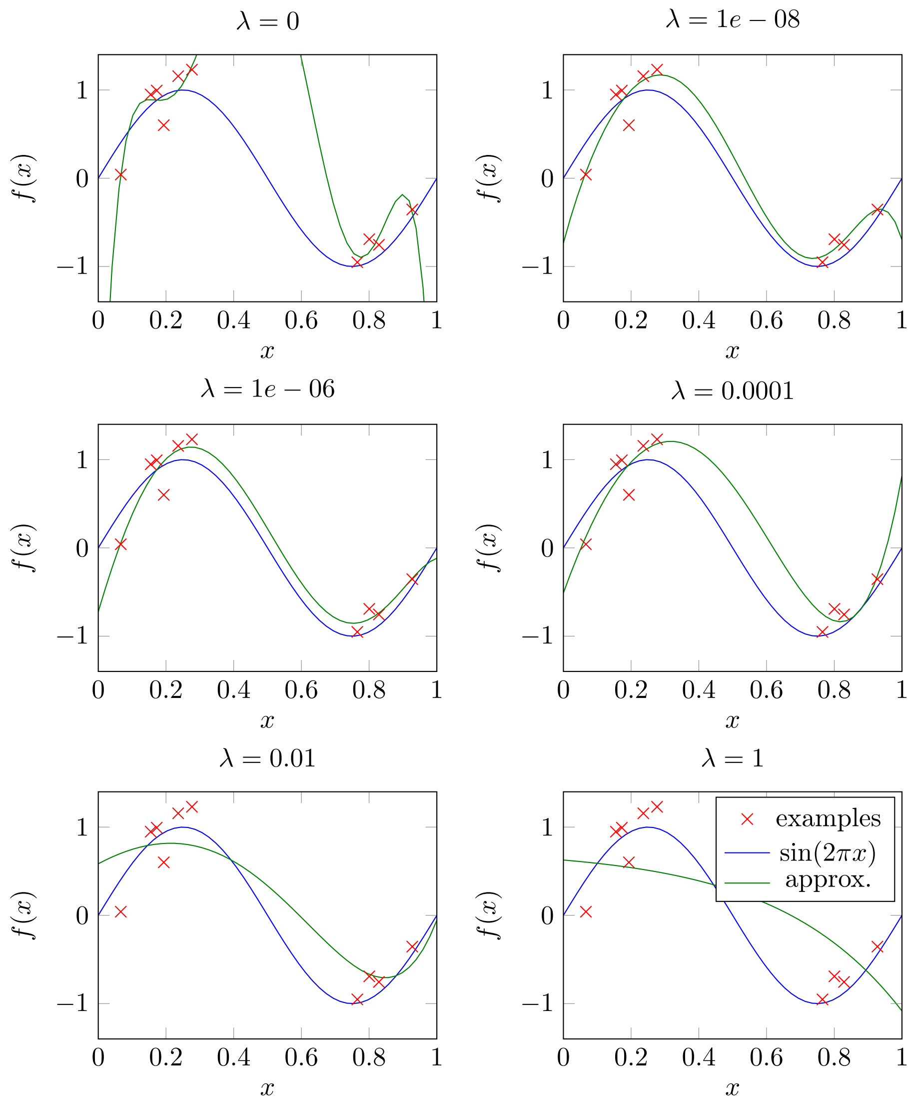

Figure 1: Results of ridge regression for $\lambda \in \{0, 10^{-8}, 10^{-6}, 10^{-4}, 10^{-2}, 1\}$

3 Kernelized Method

First, let

$$\mathrm{coin}: \mathbb{R}^d \to \{+,-\},\ x \mapsto \mathrm{coin}(x) = \begin{cases} + & \text{with probability } 1/2 \\ - & \text{with probability } 1/2 \end{cases}$$

and the euclidean distance

$$d : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R},\ d(x,y) = \|x-y\|_2 = \sqrt{\langle x-y, x-y\rangle}.$$

Then the desired decision function $f : \mathbb{R}^d \to \{+,-\}$ is

$$f(x) = \begin{cases} + & \text{if } d(x,c^+) < d(x,c^-) \\ - & \text{if } d(x,c^+) > d(x,c^-) \\ \mathrm{coin}(x) & \text{if } d(x,c^+) = d(x,c^-). \end{cases}$$

$$\Leftrightarrow f(x) = \begin{cases} + & \text{if } d^2(x,c^+) - d^2(x,c^-) < 0 \\ - & \text{if } d^2(x,c^+) - d^2(x,c^-) > 0 \\ \mathrm{coin}(x) & \text{if } d^2(x,c^+) - d^2(x,c^-) = 0 \end{cases}$$

$$\Leftrightarrow f(x) = \begin{cases} + & \text{if } \langle x-c^+, x-c^+\rangle - \langle x-c^-, x-c^-\rangle < 0 \\ - & \text{if } \langle x-c^+, x-c^+\rangle - \langle x-c^-, x-c^-\rangle > 0 \\ \mathrm{coin}(x) & \text{if } \langle x-c^+, x-c^+\rangle - \langle x-c^-, x-c^-\rangle = 0. \end{cases}$$

The constraint can also be written by a separating hyperplane.

$$f(x) = \begin{cases} + & \text{if } \langle x, (c^- - c^+)\rangle + \frac{1}{2}\left(\langle c^+, c^+\rangle - \langle c^-, c^-\rangle\right) < 0 \\ - & \text{if } \langle x, (c^- - c^+)\rangle + \frac{1}{2}\left(\langle c^+, c^+\rangle - \langle c^-, c^-\rangle\right) > 0 \\ \mathrm{coin}(x) & \text{if } \langle x, (c^- - c^+)\rangle + \frac{1}{2}\left(\langle c^+, c^+\rangle - \langle c^-, c^-\rangle\right) = 0 \end{cases}$$