Mathematics for New Technologies in Finance

Exercise sheet 2

Exercise 2.1 (Stone-Weierstrass theorem [1])

(a) Construct a sequence of polynomials converges pointwisely but not uniformly on $[0, 1]$.

(b) Construct a sequence of polynomials converges uniformly to $x \mapsto |x|$ on $[-1, 1]$. (Hint: Corollary 2.3. in [1])

(c) Prove that ReLU can be approximated uniformly by polynomials on $[-1, 1]$.

(d) Use the universal approximation theory of shallow neural networks on $[0, 1]$ to prove the Stone-Weierstrass theorem.

Exercise 2.2 (Networks on discrete path spaces)

(a) Describe the space of paths $\omega : \{1, \ldots, T\} \to \mathbb{R}^d$ as $\mathbb{R}^{dT}$.

(b) Describe a shallow neural network, which depends on value at time t and on path information up to time t. Formulate a universal approximation theorem in this setting.

Exercise 2.3 (Backpropogation of neural network) Let $\theta = (w, b, a) \in \mathbb{R}^3$ and let σ be the activation function. We consider the shallow neural network $f_\theta \colon \mathbb{R} \to \mathbb{R}$ s.t.

$$f_\theta(x) = a\sigma(wx + b). \tag{1}$$

Then we solve the regression problem with 3 data point $(x_i, y_i) \in \mathbb{R}^2$, $i = 1, 2, 3$ by minimizing the L^2 loss

$$\mathcal{L}_f = \sum_{i=1,2,3} \big(f_\theta(x_i) - y_i\big)^2. \tag{2}$$

(a) When solving the regression, do we compute $\nabla_{x_0}\mathcal{L}_f$ or $\nabla_\theta \mathcal{L}_f$?

(b) Compute $\partial_w f$ and $\partial_b f$ by chain rule. Do you find any intermediate value computed twice in both $\partial_w f$ and $\partial_b f$?

(c) Consider regression problem as a constrained optimization problem

$$\begin{aligned} &\min \sum_{i=1,2,3} l_i \\ &l_i = (\tilde{y}_i - y_i)^2 \\ &\tilde{y}_i = a\sigma(z_i), \qquad\qquad i = 1, 2, 3. \\ &z_i = wx_i + b \end{aligned} \tag{3}$$

Solve it by Lagrange multiplier and relate this with backpropagation.

(d) Generalize this idea to deep neural networks.

Exercise 2.4 (Functional analysis) Let K be a compact subset of $\mathbb{R}^d$.

(a) Let μ be a finite Borel measure on K. Prove that

$$\mathcal{L}_\mu(f) := \int_K f(x)\mu(dx) \tag{4}$$

for $f \in C(K, \mathbb{R})$ is a bounded linear functional.

(b) Let $\mathcal{L}, C(K, \mathbb{R})$ be a positive linear functional, i.e. $\mathcal{L}(f) \geq 0$ for $f \geq 0$. Then $\mathcal{L}$ is continuous.

(c) Prove that

$$\mathcal{F} := \{f \mapsto \sum_{i=1}^{n} \lambda_i f(x_i) \mid \lambda_i \in \mathbb{R}, n \in \mathbb{N}, x_i \in K, i = 1, 2, ..., n\} \tag{5}$$

is point separating and additive.

References

[1] SAMEER CHAVAN. Problems and notes: Uniform convergence and polynomial approximation.

[2] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

[3] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. Deep learning for time series classification: a review. *Data mining and knowledge discovery*, 33(4):917–963, 2019.

[4] Yann LeCun, D Touresky, G Hinton, and T Sejnowski. A theoretical framework for back-propagation. 1:21–28, 1988.