Technische Universität Braunschweig

Institut für Robotik und Prozessinformatik
Prof. J. Steil

Übungen für Grundlagen Maschinelles Lernen

SS 2019 Blatt 4

Übungstermin: 17. 05. 2017

Task 4.1 Cramer-Rao Schranke am Boccia-Beispiel:
Betrachten Sie das Boccia-Datenmodell aus dem Vorlesungsskript mit $y_i = wx^2 + \nu_i$ mit konstanter Anfangsgeschwindigkeit $x = 0.5$ und dem Schätzer

$$\hat{w} = \frac{1}{N}\sum_{i=1}^{N}\frac{y_i}{x^2}$$

Dabei sei $\nu_i \sim \mathcal{N}(0, \sigma^2)$ normalverteiltes Rauschen mit Varianz $\sigma^2 = 0.1$. Zeige experimentell über je 100 Wiederholungen, dass für $N \in \{1, 2, 3, 5, 10, 100, 1000\}$ die Varianz der Schätzung, d.h. $Var(\hat{w})$, gegen die Cramer-Rao Schranke $\frac{\sigma^2}{Nx^4}$ konvergiert. Plotte dazu $Var(\hat{w})$ und $\frac{\sigma^2}{Nx^4}$ in Abhängigkeit von N. Wie veränderen sich die Schranke und die Schätzungen bei Variation von σ^2.

Task 4.2 Regularisierte Polynom-Regression:
Berechnen Sie die Regressionsformel für die quadratische Fehlerfunktion

$$E = \frac{1}{2}\sum_{i=1}^{N}(y(x_i) - \hat{y}(x_i))^2 + \lambda\frac{1}{2}\sum_{m=0}^{M}\hat{w}_m^2(e^m - 1)$$

mit speziellem Regularisierungsterm gewichtet mit Parameter λ für einen Polynomapproximator

$$\hat{y}(x) = \sum_{m=0}^{M}\hat{w}_m x^m.$$

Welche Idee steckt hinter diesem speziellem Regularisierungsansatz und wie wirkt sich die Regularisierung auf die Lösungen der Regression aus?

Task 4.3 Bias-Varianz Dilemma am Beispiel der Polynom-Regression:
Untersuchen Sie den Einfluss der oben eingeführten Regularisierung auf die Schätzgenauigkeit. Als Datenmodell sei das Polynom $y(x, \mathbf{w}) + \nu$ gegeben mit wahren Parametern $w_0 = 0.8$, $w_1 = 2.4$, $w_2 = -1.55$, $w_3 = -0.15$, $w_4 = 0.1$ und $\nu \sim \mathcal{N}(0, \sigma^2)$ mit $\sigma^2 = 2$.

Schätzen Sie wiederholt Modellparameter $\hat{\mathbf{w}}$ für Polynome 8-ten Grades mittels der regularisierten Regression aus Aufgabe 3.2. Erstellen Sie dazu 100 Datensaetze bestehend aus je 20 Datenpunkten $(x_i,\ y(x_i) + \nu_i)$ mit uniform verteilten $x_i \in [-5, 5]$ und wiederholen Sie die Schätzung der Parameter für jeden Datensatz und Regularisierungen $\lambda \in \{0, 10^{-5}, 10^{-4}, \ldots, 10^5\}$. Welchen Einfluss hat λ auf die Aufteilung des Schätzfehlers $\langle||\mathbf{w} - \hat{\mathbf{w}}||^2\rangle$ in Bias- und Varianz-Anteile? Plotten Sie dazu den Bias des Schätzers, d.h. $||\langle\hat{\mathbf{w}}\rangle - \mathbf{w}||^2$, sowie dessen Varianz $\langle||\hat{\mathbf{w}} - \langle\hat{\mathbf{w}}\rangle||^2\rangle$ in Abhängigkeit von λ.