

Maschinelles Lernen und Data Mining
 Sommersemester 2013
Übungsblatt 7

Besprechung des Übungsblattes am 04.07.2013

Aufgabe 7-1 Modellvergleich

Vergleichen Sie Modelle der Regression mit Basisfunktionen. Die Vorhersage für einen Datenpunkt $\mathbf{x}_i \in \mathbb{R}$ sei gegeben durch:

$$f(\mathbf{x}_i, \mathbf{w}) = \sum_{j=1}^{M_\Phi} w_j \phi_j(\mathbf{x}_i)$$

Verwenden Sie die PLS-Lösung $\hat{\mathbf{w}} = (\Phi^T \Phi + \lambda I)^{-1} \Phi^T \mathbf{y}$ mit $\Phi_{i,j} = \phi_j(\mathbf{x}_i) = \mathbf{x}_i^{j-1}$. Gegeben sei der Datensatz $\mathbf{X}, \mathbf{y}$ der Größe $N = 10$, basierend auf einer Rauschvarianz von $\sigma^2 = 0.25$:

$\mathbf{X}$	0.3	0.4	0.8	1.5	1.8	3.6	4	4.3	4.6	5
$\mathbf{y}$	7	4.7	0.6	-1.1	-0.3	4.6	5.5	5.7	3.1	-0.3

Es soll das beste Modell für Basisfunktionen mit $M_\Phi \in \{1, \dots, 6\}$ bestimmt werden. Als Loss-Funktion nehmen Sie im folgenden den mittleren quadratischen Fehler (MSE).

- Bestimmen Sie das beste Modell durch Kreuzvalidierung (5-fach und 10-fach). Unterstützen die paarweisen Tests aus der Vorlesung die Entscheidung des MSE? Welchen Einfluss hat der λ -Parameter?
- Kommen Sie mit den frequentistischen Verfahren (C_p Statistik und AIC) und dem Bayes'schen Verfahren (BIC) zu den gleichen Schlüssen?
- Welchen Einfluss hat die Datengröße N , wenn Sie einen vergleichbaren Datensatz für $N = \{100, 1000\}$ simulieren?

Aufgabe 7-2 Lagrange Multiplikatoren

Wir betrachten ein Optimierungsproblem auf $x \in \mathbb{R}^n$ der Form $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ und $g_j : \mathbb{R}^n \rightarrow \mathbb{R}$:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} & f_0(x) \\ \text{so dass} & f_i(x) \leq 0 \quad \text{für } i = 1, \dots, m \\ & g_j(x) = 0 \quad \text{für } j = 1, \dots, p \end{aligned}$$

f_0 ist die Zielfunktion, f_i und g_j sind Nebenbedingungen (für Ungleichungsbedingungen und exakte Bedingungen). Eine Lagrangefunktion nimmt diese Nebenbedingungen mit in die Zielfunktion auf und optimiert dadurch innerhalb der angegebenen Grenzen:

$$\mathcal{L}(x, \gamma, \lambda) = f_0(x) + \sum_{i=1}^m \gamma_i f_i(x) + \sum_{j=1}^p \lambda_j g_j(x),$$

b.w.

wobei γ und λ reelle Vektoren, die *Lagrange Multiplikatoren* sind. Sie werden auch als duale Variablen bezeichnet. Alle γ_i müssen ≥ 0 sein, die λ_j sind frei wählbar. In einfacheren Problemen kann man sie jeweils durch Minimierung nach den Zielparametern (also x) eindeutig bestimmen.

Optimieren Sie die folgenden Probleme für $n = 2$ mithilfe eines Lagrange-Terms für Zahlen $x_1, x_2 \in \mathbb{R}$, deren Summe 20 ist,

- a) wenn $x_1 \cdot x_2$ maximal sein soll.
- b) wenn $x_1^2 + x_2^2$ minimal sein soll.
- c) wenn $e^{-(5x_1 - x_2)^2}$ maximal sein soll.

Stellen Sie hierzu zunächst jeweils den passenden Lagrange-Term auf, und minimieren Sie diesen nach den Zielparametern x_1 und x_2 . Danach darf auch wieder die Eingangsbedingung ausgenutzt werden.

Aufgabe 7-3 Minimale Oberfläche

Ein geschlossener Karton soll ein Fassungsvermögen von 36 cm^3 haben. Zusätzlich soll die Breite seiner Grundfläche genau die dreifache Länge der Grundfläche betragen.

Berechnen Sie Länge, Breite und Höhe des Kartons mit der kleinsten Oberfläche.