

Übungen zur Vorlesung

Praktische Optimierung, SoSe 2022

Prof. Dr. Günter Rudolph, Dr. Roman Kalkreuth

<https://ls11-www.cs.tu-dortmund.de/people/rudolph/teaching/lectures/POKS/SS2022/lecture.jsp>**Blatt 8, Block B**

01.06.2022

Abgabe: 08.06.2022**Aufgabe 8.1: Metamodellierung (10 Punkte)**

Auf der Übungswebseite stehen Ihnen die Dateien `po22_blat08_data.R` und `po22_blat08_data.RData` zur Verfügung. In der Datei `po22_blat08_data.RData` finden Sie für drei verschiedene Funktionen jeweils die Dataframes `data`, `grid` und `true` mit Nummern 1, 2 und 3 (also z.B. `data1`). Laden Sie diese Daten in R. Die zugehörigen Funktionen sind in der Datei `po22_blat08_data.R` enthalten. `data` enthält jeweils einen Datensatz mit den Variablen (`x` oder `x1`, `x2` oder `x1`, `x2`, `x3`, `x4`) und dem Zielfunktionswert `value`. In `grid` sind Werte für die Variablen angegeben, in denen später die Modelle ausgewertet werden sollen und in `true` sind die zugehörigen Zielfunktionswerte zu finden. Bei `grid` handelt es sich nicht immer um ein Gitter, wie es beispielsweise mit `expand.grid` erstellt werden kann. Die Werte aus `true` dürfen nur verwendet werden, um die Qualität der Modelle abschließend zu evaluieren, nicht zum Anpassen der Modelle oder zum Parametertuning.

Verwenden Sie folgende Methoden, um je ein Metamodell auf Basis der Werte in `data` zu erstellen:

- polynomielles Modell vom Grad $k = 2$ mithilfe von `lm`,
- RBF-Netz mithilfe der Funktion `rbf` aus dem Paket `RSNNS`,
- Kriging mithilfe der Funktion `km` aus dem Paket `DiceKriging`,

Sie sollen bei dem RBF-Netz und beim Kriging die Parameter der Modelle so einstellen, dass Sie ein möglichst gutes Ergebnis erzielen. Wie Sie die Optimierung der Parameter vornehmen bleibt Ihnen überlassen. Beschreiben Sie Ihr Vorgehen mit Text und Code.

Mithilfe aller Modelle sollen Vorhersagen in den Werten, die jeweils in `grid` enthalten sind, gemacht werden. Dabei müssen Sie für die drei Funktionen jeweils 3 Vorhersagen erhalten. Visualisieren Sie für die Funktionen 1 und 2 Ihre Modelle sowie die wahre Zielfunktion. Dazu sollen Sie die Werte aus `grid` sowie die Modellvorhersagen bzw. die wahren Werte aus `true` verwenden. Die jeweils 3 Modelle und die wahre Funktion sollen jeweils in einem `plotly`-Plot dargestellt werden, wobei verschiedene Farben verwendet werden sollen. Die beiden `plotly`-Plots sollen separat als `html`-Datei abgegeben werden, die nur diese beiden Grafiken enthält. Eine sinnvolle Beschriftung sowie eine Legende werden bei den Plots erwartet.

Wie gut Ihre Modelle sind, lässt sich beispielsweise anhand der mittleren absoluten Abweichung und der mittleren quadratischen Abweichung zwischen der Modellvorhersage und dem wahren Zielfunktionswert bestimmen. Dazu können Originalfunktion und Modellvorhersage auf einem Gitter ausgewertet und der entsprechende Wert berechnet werden. Verwenden Sie für die drei verschiedenen Funktionen jeweils `grid` und `true`, um die mittlere absolute Abweichung und die mittlere quadratische Abweichung zu berechnen und die Qualität der Modelle zu evaluieren.

Vergleichen Sie die Ergebnisse, die Sie erzielt haben. Verwenden Sie für den Vergleich sowohl Ihre Grafiken als auch die mittleren absoluten Abweichungen und mittleren quadratischen Abweichungen zwischen Modellvorhersagen und wahren Zielfunktionswerten.

Zu beachten:

- Beim RBF-Netz sollen die Anzahl der Neuronen sowie die Architektur des Netzes optimiert werden. Dazu sollten Modellbildung und Vorhersage unbedingt wiederholt werden, da die Ergebnisse stochastisch sind.
- Überlegen Sie sich, ob es sinnvoll ist einen Nugget-Effekt für das Kriging zu verwenden. Berücksichtigen Sie Ihre Überlegung, falls möglich, bei der Einstellung der Parameter.
- Bei der Funktion `km` sollen beispielsweise die Parameter `covtype`, `nugget.estim` und `multistart` für die Optimierung betrachtet werden.
- Wählen Sie, falls möglich, eine sinnvolle Methode für die Vorhersage aus.