Allgemeine Informationen: Dieses Aufgabenblatt enthält schriftliche und/oder Programmieraufgaben. Bitte kombinieren Sie alle Lösungen zu den schriftlichen Aufgaben zu einem einzelnen PDF Dokument, welches Sie nach folgendem Schema benennen:`{lastname}-written.pdf`. Sie können Ihre Lösungen auch scannen oder fotografieren. Achten Sie in diesem Fall auf die Lesbarkeit. Es werden JPEG/PNG Bilddateien akzeptiert welche wie folgt benannt werden müssen:`{exercisenumber}-{lastname}-written.{jpeg/png}`. Stellen Sie sicher, dass alle Rechenschritte nachvollziehbar sind und kombinieren Sie nicht zu viele kleine Schritte zu einem einzelnen. Die Programmieraufgaben müssen in *Julia* gelöst sein und Ihr Quellcode sollte nach folgendem Schema benannt sein: `{exercisenumber}-{lastname}.jl`.

(1) (2.5 Punkte) Wenden Sie die (gewöhnliche) Methode der kleinsten Fehlerquadrate (MKQ oder orindary least squares) auf das California Housing Datenset (california_housing.csv) an. Verwenden Sie dabei die Variablen median_income und median_house_value. Die Datei ols.jl enthält die nötigen Vorbereitungsschritte um die Daten der zwei Variablen auszuwählen, sie zu vereinfachen, umzuwandeln, und das Datenset in ein Training Set und ein Test Set aufzuteilen. Dabei wenden wir MKQ auf das Training Set an, und überprüfen die Approximationen mittels neuer Daten aus dem Test Set. Dazu wählen Sie eine der unten angefürten Approximationen aus, um die Werte der median_house_value Variable für das Test Set zu bestimmen.

a) (0.5 Punkte) Finden Sie mittels **linearer** MKQ eine approximierte Funktion für das Training Datenset in Julia. Folgen Sie den Anweisungen gegeben in ols.jl, plotten Sie das resultierende Polynom (mit originalen Datenpunkten), und berechnen Sie den Approximationsfehler. Wie gut beschreibt die Approximation die Daten?

b) (0.5 Punkte) Ähnlich zum vorigen Beispiel, finden Sie mittels **quadratischer** MKQ eine approximierte Beschreibung, plotten Sie das resultierende Polynom, und berechnen Sie den Approximationsfehler. Beschreibt die quadratische Approximation die Daten besser als die lineare MKQ Approximation?
Hinweis: Wir suchen die Koeffizienten a_0, a_1, und a_2 von $p(x) = a_0 + a_1 x + a_2 x^2$, die den Approximationsfehler $\sum_{j=1}^{9} |a_2 x_j^2 + a_1 x_j + a_0 - y_j|^2$ minimieren.

c) (0.5 Punkte) Benützen Sie die **Cramersche Regel** um durch **kubische** MKQ wiederum eine approximierte Beschreibung zu finden, das resultierende Polynom zu plotten, und den Approximationsfehler zu berechnen. Vergleichen Sie die kubische mit den anderen zwei Approximationen. Sie können die det() Funktion des Linear Algebra Pakets für die Cramersche Regel verwenden.
Hinweis: Ähnlich zur quadratischen MKQ suchen wir die Koeffizienten a_0, a_1, a_2, und a_3 von $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$, die den Approximationsfehler $\sum_{j=1}^{n} |a_3 x_j^3 + a_2 x_j^2 + a_1 x_j + a_0 - y_j|^2$ minimieren. Bemerken Sie die Ähnlichkeit zur Vandermonde Matrix.

d) (0.5 Punkte) Benützen Sie das **Test Set** um Inferenz auf einer Ihrer Approximationen durchzuführen und deren Generalisierung zu überprüfen. Plotten Sie dazu die Datenpunkte des Test Sets, und projezieren Sie die abhängige Variable median_income auf eine der zuvor bestimmten Kurven. Berechnen Sie den Approximationsfehler (für die ausgewählte Approximation auf den Test Daten) und beschreiben Sie ob die Approximation zu neuen Daten generalisiert.

e) (0.5 Punkte) Beschreiben Sie die Konzepte und Folgen von Overfitting und Underfitting zuerst generell, dann im Kontext dieser Aufgabe in wenigen Sätzen. One andere

Polynome auszuprobieren, welcher Grad würde zu Overfitting/Underfitting führen, und welcher Grad würde eine gute Beschreibung liefern?

(2) (2.5 Punkte) Wenden Sie baryzentrische Interpolation auf das Dreieck definiert in barycentric_interpolation.jl an. Folgen Sie dabei den Anweisungen, und benützen Sie keine vorgefertigten Julia Routinen. Sie können einige Hilfsfunktionen wie zum Beispiel cross(), norm(), sign() vom Linear Algebra Paket verwenden. Speichern Sie das entstandene Bild als barycentric_interpolation_result.png und geben Sie das Bild mit Ihren anderen Dateien ab. Ihr Dreieck sollte ähnliche Farben zu dem unten dargestellten Dreieck aufweisen.

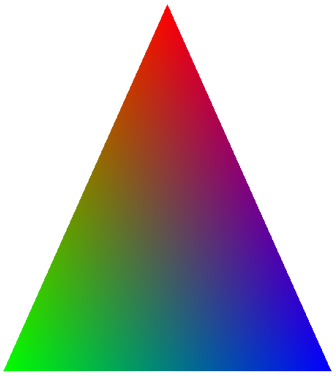


Abbildung 1: Dreieck Farben Interpolation

(3) (2.0 Punkte)

a) (0.5 Punkte) Bestimmen Sie das Polynom 3. Grades $p(x)$ sodass:

$$\begin{aligned} p(-1) &= 0 \\ p'(-1) &= -5 \\ p(1) &= 10 \\ p'(1) &= 20 \end{aligned}$$

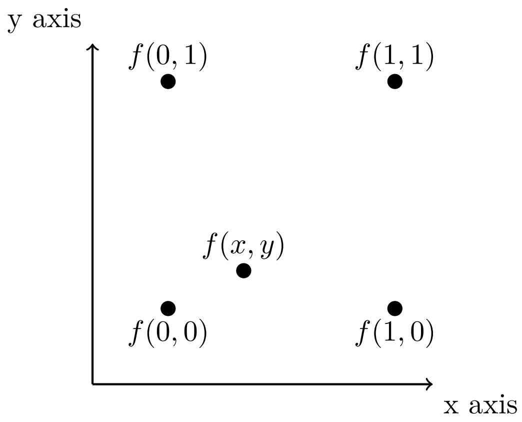
Benutzen Sie dazu die Hermite-Interpolation.

b) (0.5 Punkte) Bilden Sie das Polynom $p(x)$ für die Funktion $f(x) = \sin(\ln(x))$ mithilfe der Lagrange-Interpolation. Benutzen Sie dazu die Punkte $x_0 = 2$, $x_1 = 2.75$, $x_2 = 4$ und bestimmen Sie den Interpolationsfehler für dieses Polynom (siehe Vorlesungsfolien).

c) (1.0 Punkte) Beweisen Sie, dass die bilineare Interpolation separierbar ist. Dabei sollten Sie zeigen, dass die Formel für bilineare Interpolation in einem Einheitsquadrat:

$$f(x, y) = \begin{pmatrix} 1-x \\ x \end{pmatrix}^T \begin{pmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1) \end{pmatrix} \begin{pmatrix} 1-y \\ y \end{pmatrix}$$

hergeleitet werden kann, indem zuerst eine lineare Interpolation entlang der x-Achse und dann der y-Achse durchgeführt wird. Zeigen Sie auch, dass Sie das gleiche Ergebnis erhalten, wenn Sie zuerst entlang der y-Achse und danach an der x-Achse interpolieren.


y axis
$f(0,1)$
$f(1,1)$
$f(x,y)$
$f(0,0)$
$f(1,0)$
x axis