**General Information:** This assignment contains written and/or programming tasks. Combine all the answers to the written tasks in a single PDF document, named `{lastname}-written.pdf`. You can also scan or take pictures of (readable) handwritten papers. JPEG/PNG image files are accepted in this case and they should be named `{exercisenumber}-{lastname}-written.{jpeg/png}`. Make sure that we can follow the manual calculations. Do not combine too many small steps into one. The programming tasks have to be solved in *Julia* and the source code files have to be submitted using the following naming scheme: `{exercisenumber}-{lastname}.jl`.

(1) (2.5 Points) Perform ordinary least squares (OLS) on the California housing dataset (california_housing.csv) using the variables median_income and median_house_value. The file ols.jl contains preprocessing steps necessary to select data corresponding to these values, simplify, reshape, and split the data into a training and test set.
We will use the train set to conduct OLS, and check one approximation for novel data via the test set. To do so, we will use one of the obtained approximations to predict the median_house_value values for the test set.

a) (0.5 Points) Conduct **linear** OLS approximation using the preprocessed training set in Julia. Follow the instructions outlined in ols.jl, plot the resulting polynomial (w.r.t. the training data points), and calculate the approximation error. Does the approximation describe the data sufficiently?

b) (0.5 Points) Analogously to the prior task, perform **quadratic** OLS, plot the polynomial, and calculate the approximation error. Does the fitted curve capture the data better than its linear counterpart?
**Hint:** We are looking for the coefficients $a_0, a_1$, and $a_2$ of $p(x) = a_0 + a_1x + a_2x^2$ that minimize the approximation error $\sum_{j=1}^{n} |a_2x_j^2 + a_1x_j + a_0 - y_j|^2$.

c) (0.5 Points) Using **Cramer's rule**, conduct **cubic** OLS, plot the polynomial, and calculate the approximation error. Compare the resulting approximation to its predecessors. You are allowed to use det() from the Linear Algebra package for Cramer's rule.
**Hint:** Similar to quadratic OLS, we are looking for the coefficients $a_0, a_1, a_2$, and $a_3$ of $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$ that minimize the approximation error $\sum_{j=1}^{n} |a_3x_j^3 + a_2x_j^2 + a_1x_j + a_0 - y_j|^2$. Note that this is linked to the Vandermonde matrix.

d) (0.5 Points) Use the **test set** for inference and check the generalizability of (one of) your approximations. To this end, plot the actual data points and project the dependent variable median_income onto one of the previously obtained curves. Calculate the approximation error (w.r.t. the selected approximation) and reason whether or not the approximation generalizes to unseen data.

e) (0.5 Points) Briefly describe the concepts and consequences of overfitting and underfitting generally first, and then in this context. Without explicitly trying other polynomials, which order could lead to overfitting/underfitting, and which order would you deem a good fit?

Solution: See ols_solution.jl for a-d. For e: Generally: Overfitting: Model/approximation loses generalizability (mostly due to overshot complexity or prolonged training) and only fits already seen data/training set, but doesn't work for new observation. Underfitting: Model/approximation is too simplistic/does not have enough complexity to capture data sufficiently and is thus a bad fit. New, as well as already seen observations, are not represented well by the model/approximations which leads to poor performance/predictions. In this context: Overfitting would entail using a (high order) polynomial that captures almost every data point in the training set perfectly but yields a high inference error. The linear OLS approximation borders on underfitting the dataset since it doesn't capture the increasing curve formed by the data points or some of their variations. Polynomials of order 2 or 3 seem to be good fits. Everything above that might lead to overfitting.

(2) (2.5 Points) Apply (general) barycentric interpolation to the triangle defined in barycentric_interpolation.jl. Follow the outlined instructions, and don't use any routines readily available in Julia. You may use some functions e.g. cross(), norm(), sign() from the Linear Algebra package. Save the resulting image and submit it as barycentric_interpolation_result.png together with your other files. Your triangle should feature similar colors to the one depicted below.
Solution: See barycentric_interpolation_solution.jl

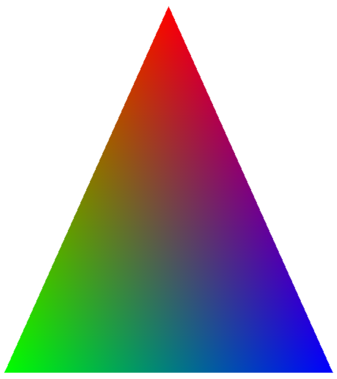


Figure 1: Triangle color interpolation

(3) (2.0 Points)

a) (0.5 Points) Determine the 3rd order polynomial $p(x)$ such that

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

using Hermite interpolation.

Solution: The polynomial is given by $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$ with $p'(x) = a_1 + 2a_2 x + 3a_3 x^2$. Enforcing the constraints we get a $4 \times 4$ linear system given by

$$\begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & -2 & 3 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 10 \\ -5 \\ 20 \end{pmatrix}$$

Solving the system (using Gaussian elimination) the coefficients are $a_0 = -1.25$, $a_1 = 3.75$, $a_2 = 6.25$, $a_3 = 1.25$.

b) (0.5 Points) Construct the Lagrange interpolating polynomial $p(x)$ for the function $f(x) = \sin(\ln(x))$ using the nodes $x_0 = 2$, $x_1 = 2.75$, $x_2 = 4$ and determine the error form for this polynomial (as given in the lecture slides).

Solution: We have

$$\begin{aligned} L_0(x) &= \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} = \frac{2}{3}(x-2.75)(x-4) \\ L_1(x) &= \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} = -\frac{1}{0.9375}(x-2)(x-4) \\ L_2(x) &= \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} = \frac{2}{5}(x-2)(x-2.75) \end{aligned}$$

Thus the interpolating polynomial is given by

$$\begin{aligned} p(x) &= \sum_{k=0}^{2} f(x_k)L_k(x) = f(x_0)L_0(x) + f(x_1)L_1(x) + f(x_2)L_2(x) \\ &= \frac{2\sin(\ln(2))}{3}(x-2.75)(x-4) - \frac{\sin(\ln(2.75))}{0.9375}(x-2)(x-4) \\ &+ \frac{2\sin(\ln(4))}{5}(x-2)(x-2.75) \end{aligned}$$

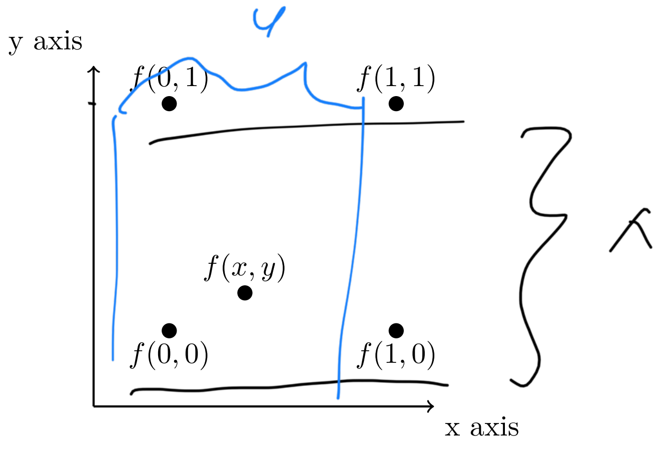
For the error we have that $f'(x) = \frac{1}{x}\cos(\ln(x))$ and $f''(x) = -\frac{1}{x^2}\cos(\ln(x)) - \frac{1}{x^2}\sin(\ln(x))$ and $f'''(x) = \frac{2}{x^3}\cos(\ln(x)) + \frac{1}{x^3}\sin(\ln(x)) + \frac{2}{x^3}\sin(\ln(x)) - \frac{1}{x^3}\cos(\ln(x)) = \frac{\cos(\ln(x))+3\sin(\ln(x))}{x^3}$. Thus the Lagrange polynomial has the error form:

$$\begin{aligned} e(x) &= \frac{f'''(\xi)}{3!}(x-2)(x-2.75)(x-4) \text{ for } \xi \in (2,4) \\ &= \frac{1}{6}\frac{\cos(\ln(\xi)) + 3\sin(\ln(\xi))}{\xi^3}(x-2)(x-2.75)(x-4) \end{aligned}$$

c) (1.0 Points) Prove that bilinear interpolation is separable. Specifically, show that the final formula for bilinear interpolation in a unit square:

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

can be derived by first performing a linear interpolation along the x-axis, then the y-axis. Further, prove that the opposite holds by showing that we get the same result if we interpolate along the y-axis first, and then the x-axis.



Solution: We perform interpolation along the x-axis first and get:

$$f(x,0) = f(0,0)(1-x) + f(1,0)x$$
$$f(x,1) = f(0,1)(1-x) + f(1,1)x$$

Performing interpolation along the y-axis then we get:

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

For the other way around by performing interpolation along the y-axis first we get:

$$f(0,y) = f(0,0)(1-y) + f(0,1)y$$
$$f(1,y) = f(1,0)(1-y) + f(1,1)y$$

Then along the x-axis we get:

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