Abteilung für
Angewandte Mathematik

Prof. Dr. Sören Bartels, Tatjana Schreiber April 23th, 2025

Algorithmic Aspects of Data Analytics and Machine Learning

SS 2025 — Sheet 1

https://aam.uni-freiburg.de/agba/lehre/ss25/algml/index.html

Due: May 2, 2025, 2 p.m.
Note: You must register for the exercises between April 23, 12 a.m. and April 25, 2 p.m. via HisInOne.

Task 1 (4 points)
Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $x, y \in \mathbb{R}^n$. Compute the derivative of the function

$$t \mapsto \|A(x+ty)-b\|_2^2, \quad t \in \mathbb{R},$$

and show the following:
If x minimizes $\|Ax-b\|_2^2$, then it satisfies

$$A^\top Ax = A^\top b.$$

This equation is called the *(Gauss) normal equation.*

Task 2 (4 points)
Given the data set $D = \{(1,2),(2,2),(2,3),(3,3),(4,3)\} \subseteq \mathbb{R}_x \times \mathbb{R}_y$. Calculate by hand

$$f^* = \arg\min \left\{ \sum_{i=1}^{n} (f(x_i)-y_i)^2 \,\middle|\, f:\mathbb{R}\to\mathbb{R} \text{ affine-linear} \right\}$$

and the mean values

$$\bar{x} = \frac{x_1+\cdots+x_n}{n} \quad \text{and} \quad \bar{y} = \frac{y_1+\cdots+y_n}{n}.$$

Verify that $f^*(\bar{x}) = \bar{y}$.
Sketch the data points and the regression line.

Task 3 (4 points)
Find a data set in which not all x-values are the same, such that the optimization problem

$$\arg\min \left\{ \sum_{i=1}^{n} |f(x_i)-y_i| \,\middle|\, f:\mathbb{R}\to\mathbb{R} \text{ affine-linear} \right\}$$

has more than one solution. Sketch the data points and at least two different regression lines.
Then, for your chosen data set, compute the unique solution to

$$\arg\min \left\{ \sum_{i=1}^{n} (f(x_i)-y_i)^2 \,\middle|\, f:\mathbb{R}\to\mathbb{R} \text{ affine-linear} \right\}.$$

Task 4 (4 points)
Show, using a contradiction argument, that the solution to

$$f^* = \arg\min \left\{ \sum_{i=1}^{n} (f(x_i)-y_i)^2 \,\middle|\, f:\mathbb{R}\to\mathbb{R} \text{ affine-linear} \right\}$$

is unique.
Hint: Assume, for contradiction, that there are two minimizer $f_1^*(x) = \lambda_1 x + b_1$ and $f_2^*(x) = \lambda_2 x + b_2$. Show that for $s \in (0,1)$

$$\sum_{i=1}^{n} (h(x_i)-y_i)^2 < \sum_{i=1}^{n} (f_1^*(x_i)-y_i)^2 = \sum_{i=1}^{n} (f_2^*(x_i)-y_i)^2,$$

where $h(x) := s f_1^*(x) + (1-s) f_2^*(x)$.