Worksheet 4 (March 15th, 2021)

Least squares and infections disease

Let us assume an infectious disease with the following reported new infections I_i on each day t_i, for $i = 1, \ldots, 10$.

Table 1: Number of new infections I_i on days t_i.

t_i:	1	2	3	4	5	6	7	8	9	10
I_i:	14	20	21	24	15	45	67	150	422	987

Using least squares fitting, we would like to understand the nature of this growth. We consider two models to describe the connection between time (i.e., days) t and the number of new infections, both with 3 unknown parameters (a, b, c):

$$I(t) = a + bt + ct^2 \qquad \text{(polynomial model)}$$

$$I(t) = a + bt + c\exp(t) \qquad \text{(exponential model)}$$

Our goal is to figure out which model describes the progression of the infections better, and we use least squares fitting to figure that out. Note that if a model would fit the data perfectly, $I(t_i) = I_i$ for all i. In general, you will not be able to find parameters that satisfy this, and thus have to use least squares fitting (sometimes this is also called *regression*).

1a. Formulate, assuming the polynomial model, the least squares problem for the parameters $\boldsymbol{x} = [a, b, c]^T$ by specifying the matrices A and the vector $\boldsymbol{b}$:

$$\min_{\boldsymbol{x}\in\mathbb{R}^3} \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2$$

1b. Same as above, but for the exponential model.

1c. Solve this problem in Matlab using three methods: backlash, QR factorization, and normal equations. Do you get the same result? Plot the data as points, as well as the model as a line (put more than 10 points on the line to make it a smooth curve).

1d. To decide which model describes the data better, we need to compute the difference between the model and the data points. Compute in Matlab the magnitude of the residual, i.e., $\|A\boldsymbol{x} - \boldsymbol{b}\|_2^2$, for both models, and decide which one is better. Even better, look at the plot from part 3 and see how well the model actually works, for example, does it work well for all times or only for early/later times?