1 AR system identification
You are given input/output measurements

\[(u(1), y(1)), \ldots, (u(N), y(N)),\]

where \(u(t)\) and \(y(t)\) are, respectively, the input and output of an unknown system at time \(t\). You believe that the system can be approximated by an autoregressive model:

\[\hat{y}(t) = a_0 u(t) + b_1 y(t - 1) + \cdots + b_n y(t - n).\]

(a) Explain how to choose the coefficients \(a_0, b_1, \ldots, b_n\) in order to minimize the sum of squared errors:

\[\sum_{t=n+1}^{N} (y(t) - \hat{y}(t))^2.\]

(b) Apply your method to the data given in \texttt{ar_system_identification_data.m}. Carry out the estimation for \(n = 1, \ldots, 35\). The relative error of the model is defined to be:

\[\bar{\epsilon} = \frac{\sum_{t=n+1}^{N} (y(t) - \hat{y}(t))^2}{\sum_{t=n+1}^{N} y(t)^2}.\]

Submit a plot of the relative error versus \(n\). What do you think is a good value of \(n\)?

(c) The file \texttt{ar_system_identification_data.m} contains a second set of data:

\[(u_{val}(1), y_{val}(1)), \ldots, (u_{val}(N), y_{val}(N)).\]

Compute the estimated output sequence

\[\hat{y}_{val}(t) = a_0 u_{val}(t) + b_1 y_{val}(t - 1) + \cdots + b_n y_{val}(t - n),\]

and the relative error on the validation set:

\[\bar{\epsilon}_{val} = \frac{\sum_{t=n+1}^{N} (y_{val}(t) - \hat{y}_{val}(t))^2}{\sum_{t=n+1}^{N} y_{val}(t)^2}.\]

Submit a plot of the relative error on the validation set versus \(n\). Briefly discuss your results. What do you think is a good value of \(n\)?

2 Approximate inductance formula
Figure 1 shows a planar spiral inductor implemented in CMOS for use in RF circuits.
The inductor is characterized by four parameters:

- $n$, the number of turns,
- $w$, the width of the wire,
- $d$, the inner diameter, and
- $D$, the outer diameter.

The inductance $L$ of the spiral inductor is a complicated function of the parameters $n$, $w$, $d$, and $D$. It is possible to compute $L$ by solving Maxwell’s equations, but this takes considerable computer time. An alternative approach is to fabricate many inductors with different parameter values, and then fit a model to the measurements. Suppose we fabricate $N$ inductors. Let $n_i$, $w_i$, $d_i$, and $D_i$ be the parameters of the $i$th inductor (measured in $\mu$m), and let $L_i$ be its inductance (measured in nH).

(a) Explain how to fit a model of the form

$$\hat{L}_i = \beta_1 n_i^\beta_2 w_i^\beta_3 d_i^\beta_4 D_i^\beta_5,$$

where $\beta_1, \ldots, \beta_5$ are parameters. Note that we have not specified the criterion used to choose the parameters: this is part of the problem.

(b) Apply your method to the data given in `inductance_formula_data.m`. Report your parameter estimates and the average percentage error, which is defined to be

$$\frac{1}{N} \sum_{i=1}^{N} \frac{100|\hat{L}_i - L_i|}{L_i}.$$

3 Least-squares quadratic extrapolation of a time series

Suppose we are given a time series $z(1), \ldots, z(t)$. We can predict $z(t + 1)$ as follows. First, we fit a quadratic function $f(t) = a_2 t^2 + a_1 t + a_0$ to $z(t), z(t - 1), \ldots, z(t - d)$, where $d \geq 2$ is a parameter. Then, we use the fitted function to extrapolate the value of the time series at $t + 1$: that is, our estimate is $\hat{z}(t+1) = f(t+1)$. 
(a) Find a constant vector $c \in \mathbb{R}^{d+1}$ such that

$$
\hat{z}(t+1) = c^T \begin{bmatrix}
  z(t) \\
  z(t-1) \\
  \vdots \\
  z(t-d)
\end{bmatrix}.
$$

Intuitively, this means that $\hat{z}(t+1)$ is a linear function of $z(t), z(t-1), \ldots, z(t-d)$, and the weights of the linear function do not depend on $t$.

(b) Consider the time series

$$
z(t) = 5 \sin\left(\frac{1}{10} t + 2\right) + \frac{1}{10} \sin(t) + \frac{1}{10} \sin(2t - 5).
$$

For $d = 2$ and $d = 10$, compute $\hat{z}(t)$ for $t = d+1, \ldots, T = 1000$. Submit a plot showing $z(t)$ for $t = 1, \ldots, T_{\text{plot}}$ and $\hat{z}(t)$ for $t = d+1, \ldots, T_{\text{plot}}$ on a single set of axes, where $T_{\text{plot}} = 100$. Report the relative root mean squared prediction error:

$$
\sqrt{\frac{1}{T-d+1} \sum_{t=d+1}^{T} (\hat{z}(t) - z(t))^2}.
$$

4 Estimating emissions from spot measurements

There are $n$ sources of a pollutant at known locations $s_1, \ldots, s_n \in \mathbb{R}^2$. The $j$th source emits the pollutant at $x_j > 0$. We measure the total pollutant at known locations $t_1, \ldots, t_m \in \mathbb{R}^2$. The total pollutant measured at the $i$th location is the sum of the contributions from the $n$ sources. The contribution from the $j$th source to the pollutant level at the $i$th measurement location is $y_i = \alpha x_j / \| s_j - t_i \|^2$, where $\alpha > 0$ is a known constant. In other words, the pollutant concentration from a source follows an inverse square law, and is proportional to the emission rate. We assume that the measurement locations do not coincide with the source locations: that is, $t_i \neq s_j$ for all $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Additionally, we assume that the measurement and source locations are distinct: that is, $t_{i_1} \neq t_{i_2}$ whenever $i_1 \neq i_2$, and $s_{j_1} \neq s_{j_2}$ whenever $j_1 \neq j_2$.

(a) Give an example of $n = 3$ source locations and $m = 4$ measurement locations such that it is impossible to determine the emission rates from the spot measurements. In this part of the problem, we ignore the issues of noise and sensor errors, and assume that the spot measurements are exactly as described above. To demonstrate the validity of your example, give two distinct sets of emission rates that yield the same set of measurements.

(b) The file `emissions_estimation_data.m` defines

- the scalar $\alpha$,
- a $2 \times n$ matrix $s$ whose $j$th column is the location of the $j$th source,
- a $2 \times m$ matrix $t$ whose $i$th column is the $i$th measurement location, and
• vectors \( y_1 \) and \( y_2 \) of length \( m \) containing two sets of noisy spot measurements.

Explain how to estimate the emission rates given the noisy spot measurements. Apply your method to the first set of noisy spot measurements, and report your estimates of the emission rates of the pollutant sources.

(c) Now suppose that one of the sensors is faulty. The corresponding measurement is likely to have a much larger error than the other measurements. Explain how to identify the faulty sensor, and estimate the emission rates using the other sensors. Apply your method to the second set of noisy spot measurements. Report your estimates of the emission rates and which sensor you identified as faulty. (Note that \( y_1 \) and \( y_2 \) do not necessarily correspond to the same vector of emission rates, so you should not be worried if your estimates are not similar.)