Homework 4: Least squares

Due date: 11:00pm on Monday February 27, 2017
See the course website for instructions and submission details.

1. [10 pts] Moving averages. There are many ways to model the relationship between an input sequence $\{u_1, u_2, \dots\}$ and an output sequence $\{y_1, y_2, \dots\}$. In class, we saw the *moving average* (MA) model, where each output is approximated by a linear combination of the k most recent inputs:

$$\text{MA:} \qquad y_t \approx b_1 u_t + b_2 u_{t-1} + \cdots + b_k u_{t-k+1}$$

We then used least-squares to find the coefficients $b_1, \dots, b_k$. What if we didn't have access to the inputs at all, and we were asked to predict future y values based *only* on the previous y values? One way to do this is by using an *autoregressive* (AR) model, where each output is approximated by a linear combination of the ℓ most recent outputs (excluding the present one):

$$\text{AR:} \qquad y_t \approx a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_\ell y_{t-\ell}$$

Of course, if the inputs contain pertinent information, we shouldn't expect the AR method to outperform the MA method!

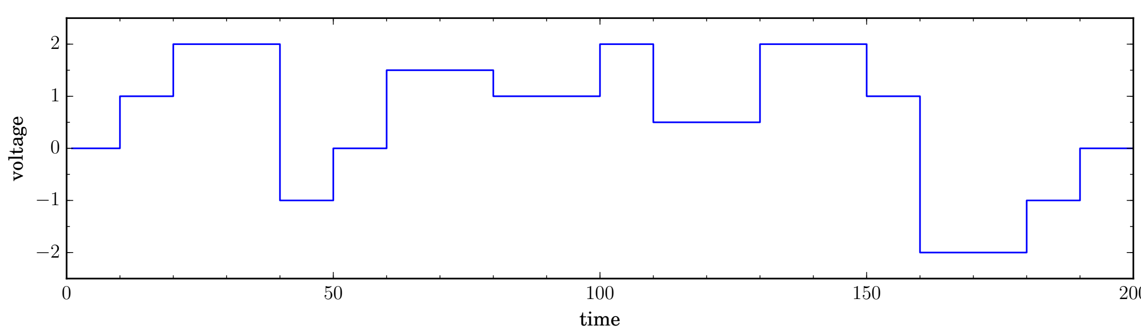
a) Using the same dataset from class `uy_data.csv`, plot the true y, and on the same axes, also plot the estimated $\hat{y}$ using the MA model and the estimated $\hat{y}$ using the AR model. Use $k = 5$ for both models. To quantify the difference between estimates, also compute $\|y - \hat{y}\|$ for both cases.

b) Yet another possible modeling choice is to combine both AR and MA. Unsurprisingly, this is called the *autoregressive moving average* (ARMA) model:

$$\text{ARMA:} \qquad y_t \approx a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_\ell y_{t-\ell} + b_1 u_t + b_2 u_{t-1} + \cdots + b_k u_{t-k+1}$$

Solve the problem once more, this time using an ARMA model with $k = \ell = 1$. Plot y and $\hat{y}$ as before, and also compute the error $\|y - \hat{y}\|$.

2. [10 pts] Voltage smoothing. We would like to send a sequence of voltage inputs to the manipulator arm of a robot. The desired signal is shown in the plot below (also available in `voltages.csv`)

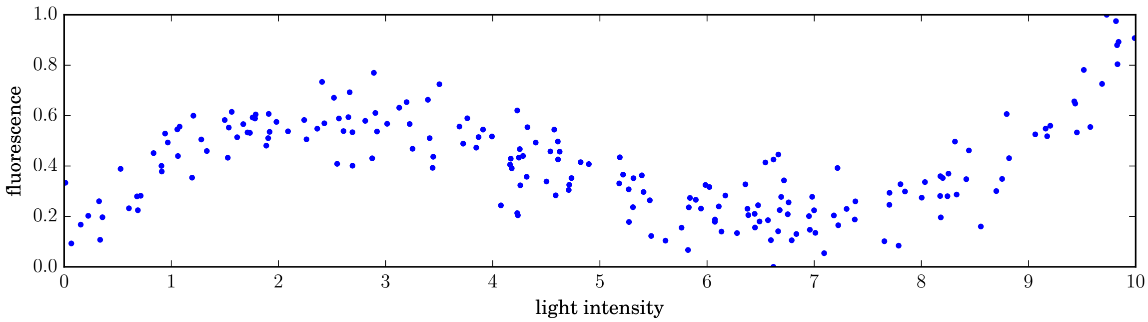


Unfortunately, abrupt changes in voltage cause undue wear and tear on the motors over time, so we would like to modify the signal so that the transitions are smoother. If the voltages above are given by $v_1, v_2, \dots, v_{200}$, one way to characterize smoothness is via the sum of squared differences:

$$R(v) = (v_2 - v_1)^2 + (v_3 - v_2)^2 + \cdots + (v_{200} - v_{199})^2$$

When $R(v)$ is smaller, the voltage is smoother. Solve a regularized least squares problem that explores the tradeoff between matching the desired signal above and making the signal smooth. Explain your reasoning, and include a plot comparing the desired voltages with your smoothed voltages.

3. [10 pts] Spline fitting. We are running a series of experiments to evaluate the properties of a new fluorescent material. As we vary the intensity of the incident light, the material should fluoresce different amounts. Unfortunately, the material isn't perfectly uniform and our method for measuring fluorescence is not very accurate. After testing 200 different intensities, we obtained the result below (also available in `xy_data.csv`). The intensities x_i and fluorescences y_i are recorded in the first and second columns of the data matrix, respectively.



The material has interesting nonlinear properties, and we would like to characterize the relationship between intensity and fluorescence by using an approximate model that agrees well with the trend of our experimental data. Although there is noise in the data, we know from physics that *the fluorescence must be zero when the intensity is zero.* This fact must be reflected in all of our models!

a) Polynomial fit. Find the best cubic polynomial fit to the data. In other words, look for a function of the form $y = a_1x^3 + a_2x^2 + a_3x + a_4$ that has the best possible agreement with the data. Remember that the model should have zero fluorescence when the intensity is zero! Include a plot of the data along with your best-fit cubic on the same axes.

b) Spline fit. Instead of using a single cubic polynomial, we will look for a fit to the data using two quadratic polynomials. Specifically, we want to find coefficients p_i and q_i so that our data is well modeled by the piecewise quadratic function:

$$y = \begin{cases} p_1x^2 + p_2x + p_3 & \text{if } 0 \le x < 4 \\ q_1x^2 + q_2x + q_3 & \text{if } 4 \le x < 10 \end{cases}$$

These quadratic functions must be designed so that:

- as in the cubic model, there is zero fluorescence when the intensity is zero.
- both quadratic pieces have the same value at $x = 4$.
- both quadratic pieces have the same slope at $x = 4$.

In other words, we are looking for a *smooth* piecewise quadratic. This is also known as a *spline* (this is just one type of spline, there are many other types!). Include a plot of the data along with your best-fit model.