

Homework 1

Due Date: Wednesday September 29, 2021

1. *Mostly zeros.* Consider the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- (a) Find a basis for $\text{range}(A)$ and find a basis for $\text{null}(A)$.
- (b) Find a vector b such that $Ax = b$ has no solutions, or explain why no such b can exist. Repeat for the case of exactly one solution, and the case of infinitely many solutions.
2. *Predicting tree age.* We want to estimate the age of maple trees in a particular forest based on the circumferences of their trunks. Finding the true age of a tree is invasive and time-consuming, so we only have data for $N = 10$ trees. Our data is points (x_i, y_i) for $i = 1, \dots, N$, where, x_i and y_i are the circumference and age of tree i , respectively. Here are the data:

x_i (inches)	y_i (years)
18.1	83
8.0	42
16.8	79
2.8	16
3.9	24
12.7	73
11.5	60
9.5	44
8.0	47
13.2	67

- (a) Suppose age is proportional to circumference, so $y \approx mx$ for some m (no intercept). Find the slope m that minimizes the squared norm of the residual: $\sum_{i=1}^N (mx_i - y_i)^2$.
- (b) Suppose we instead have the more complicated quadratic relationship $y \approx a_2x^2 + a_1x + a_0$. Find the choice of parameters a_0, a_1, a_2 that minimizes the squared norm of the residual.
- (c) Make a scatter plot of the data, with circumference on the x -axis and age on the y -axis. On your plot, include the curves of best fit from (a) and (b).
- (d) Suppose the data were changed so that $x_i = 10$ for all i , but the y_i are unchanged. Consider a linear-plus-intercept model $y \approx mx + b$. Show and explain why the least squares problem has infinitely many solutions in this case. Find a general expression for the set of all solutions, and make a plot showing what these solutions look like.

3. *Perp properties.* In this problem, we assume $S \subseteq \mathbb{R}^n$ is *any subset of* \mathbb{R}^n , not necessarily a subspace! In this case, the perp space S^\perp is defined as before:

$$S^\perp := \{x \in \mathbb{R}^n \mid \langle x, s \rangle = 0 \text{ for all } s \in S\}.$$

- (a) Prove that S^\perp is a subspace of \mathbb{R}^n .
- (b) Prove that $S \subseteq S^{\perp\perp}$, where we defined $S^{\perp\perp} := (S^\perp)^\perp$.

4. *Sums and intersections.* Consider the following two operations between sets $S, T \subseteq \mathbb{R}^n$.

$$\text{Intersection: } S \cap T := \{x \in \mathbb{R}^n \mid x \in S \text{ and } x \in T\}$$

$$\text{Sum: } S + T := \{s + t \mid s \in S \text{ and } t \in T\}.$$

Suppose S and T are subspaces of \mathbb{R}^n .

- (a) Prove that $S \cap T$ and $S + T$ are subspaces.
 - (b) Prove that $(S + T)^\perp = S^\perp \cap T^\perp$.
 - (c) Prove that $(S \cap T)^\perp = S^\perp + T^\perp$.
5. *Hovercraft rendez-vous.* You are in command of a hovercraft, which can move around in 2D through the use of two thrusters. The dynamics are described by the equations

$$x_{t+1} = x_t + v_t,$$

$$v_{t+1} = v_t + u_t,$$

where $x_t \in \mathbb{R}^2$, $v_t \in \mathbb{R}^2$, and $u_t \in \mathbb{R}^2$ are the position, velocity, and thruster input at time t . Our task is to reach the following waypoints at the following times:

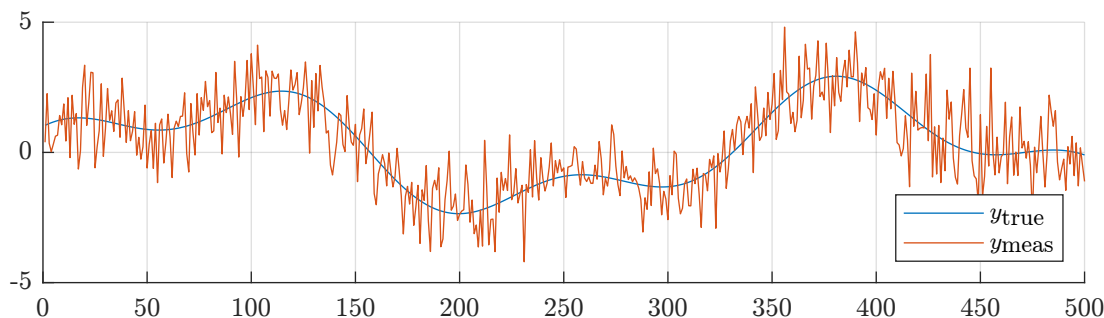
$$x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad x_{20} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \quad x_{35} = \begin{bmatrix} 7 \\ 0 \end{bmatrix}, \quad x_{60} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

There are many sequences of thruster commands u_0, \dots, u_{59} that achieve this. We want the solution that minimizes the total fuel consumption, which is given by

$$J = \sum_{t=0}^{59} \|u_t\|^2$$

- (a) Write the waypoint constraints as a large set of linear equations of the form $Au = b$, where u is the thruster input. Note that $u \in \mathbb{R}^{120}$ because $u_t \in \mathbb{R}^2$ and $t = 0, 1, \dots, 59$.
- (b) Solve the minimum norm optimization problem to find the optimal thruster input \hat{u} . Make a 2D plot of the optimal trajectory x_t as a function of t and verify that the trajectory passes through the waypoints at the appropriate times.
- (c) Plot the optimal thruster inputs u_t as a function of t . What do you observe about the shape of these functions?

6. *Smoothing via regularization.* We want to estimate a smooth signal $y_{\text{true}} \in \mathbb{R}^{500}$. We have access to a noisy measurement $y_{\text{meas}} = y_{\text{true}} + w$, where $w \in \mathbb{R}^{500}$ is unknown noise. Here is a plot of the true signal and the noisy measurement.



The Matlab code that produces y_{true} and y_{meas} is:

```
t = 1:500; % time points
ytrue = 2*sin(t/50) + cos(t/20); % true signal
rng(1); ymeas = ytrue + randn(1,500); % noisy measurement
```

To find a smooth estimate, we will use least squares with a regularizer that penalizes abrupt changes. One way to do this is by penalizing the norm squared of the first derivative. Our signal is discrete, so we'll use the approximation $y'(t) \approx y_t - y_{t-1}$ and solve the regularized least squares problem

$$y_{\text{est}} = \arg \min_{y \in \mathbb{R}^{500}} \|y - y_{\text{meas}}\|^2 + \lambda \sum_{t=2}^{500} (y_t - y_{t-1})^2,$$

where $\lambda > 0$ is a trade-off parameter.

- Transform the problem into an ordinary least squares problem $\min_x \|Ax - b\|^2$. What are A , b , and x , in this case?
- Experiment with λ values in the range $0 < \lambda < 100$ and plot/compare the regularized estimates y_{est} for different values of λ . What happens as λ is increased?
- Instead of penalizing the first derivative, let's instead penalize the norm squared of the second derivative (curvature penalty). For this, we'll use $y''(t) \approx y_{t+1} - 2y_t + y_{t-1}$. Repeat parts (a) and (b) using this new regularizer. Adjust the λ range as needed.