First of all, carefully reread Chapter 2, especially the sections on vector norms and operator norms.

Problem 01 Using MATLAB, do the following procedure.

(**a**) Download the data file

HW_02.mat

from Canvas to your working directory, and load it into your MATLAB session. Check what variables (i.e., arrays) are defined in this data file by running

```
>> whos
```

(**b**) Plot the data by typing

```
>> plot(x,y); grid;
```

(**c**) Create the Vandermonde matrix for polynomials of degree 1; (i.e., lines) by typing

```
>> A=[x.^0 x.^1];
```

(**d**) Compute the least squares line over the given data by typing

```
>> sol = inv(A'*A)*A'*y;
```

Then, overlay the least squares line over the current plot by typing

```
>> hold on; plot(x, sol(1)+sol(2)*x, '--');
```

Create the title and axis labels by typing

```
>> title('Least Squares Linear Fit'); xlabel('x'); ylabel('y');
```

Print out this plot and include a the PDF copy of the plot in you HW PDF file also with a *carefully written description* of how you obtained the plot and what it is.

(**e**) Finally, create a MATLAB file HW_02.m from the above commands with appropriate comments. Use the MATLAB listing in the Solutions to HW_01 as an example of good commenting practice.

(**f**) Write a detailed explanation of what this MATLAB program does and put it in your PDF file.

Problem 02 Prove that for a square matrix A, null(A) = {**0**} implies A is invertible.

Note that you are trying to prove a statement of the form P $\Longrightarrow$ Q, where

- P is "null(A) = {**0**}"
- Q is "A is invertible."

I would have preferred that your solution begin by only using the following fundamental definition. **Definition:** The matrix A is *invertible* if and only if the mapping

$$\mathrm{A} : \mathrm{Range}\,(\mathrm{A}) \longrightarrow \mathrm{Codomain}\,(\mathrm{A}) \tag{1}$$

is one-to-one. However, since I didn't explicitly state this in the problem statement I accepted any correct proof.

Solution to Problem 02: (By Contradiction) Let A be a square matrix. Suppose that null(A) = {**0**}. (This is the statement that P is true.) We will also assume that A is *not* invertible (this is the statement that Q is *not* true, which is written $\not{Q}$) and show that this contradicts P:

$$\mathrm{null}(\mathrm{A}) \;=\; \{\mathbf{0}\}\,. \tag{2}$$

Since A is not invertible, it does not have full rank. If A does not have full rank, then there exist vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ such that $x \neq y$ but $\mathrm{A}\mathbf{x} = \mathrm{A}\mathbf{y}$. We then can write

$$\mathrm{A}\mathbf{x} - \mathrm{A}\mathbf{y} \;=\; \mathrm{A}(\mathbf{x} - \mathbf{y}) \;=\; \mathbf{0}\,.$$

Then the vector $\mathbf{x} - \mathbf{y}$ is in the null space of A. However, this vector has to be nonzero. Thus, we have a contradiction.

Problem 03 Find the minimum value of $\|\mathbf{x}\|_1$ subject to $\|\mathbf{x}\|_2 = 1$ in $\mathbb{R}^2$. Which $\mathbf{x}$ achieves such minimum?

Solution to Problem 03: Define the function $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(\mathbf{x}) = \|\mathbf{x}\|_1 = \sum_{k=1}^{2} |x_k|$$

Next, define the region $\mathrm{R} \subset \mathbb{R}^2$ to be the set of vectors such that

$$\mathrm{R} = \{\, \mathbf{x} \;\mid\; \|\mathbf{x}\|_2 = 1 \,\}.$$

Then, in this problem we want to maximize the function $f(\mathbf{x})$ over the region R. We can translate the region from cartesian coordinates into polar coordinates. We know that for any $\mathbf{x} \in \mathrm{R}$, we can write

$$\mathbf{x} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$$

With $\theta \in [0, 2\pi]$. Then, to maximize the function $f(\mathbf{x})$ over the interval R, we can rewrite the function on the interval as follows

$$f(\theta) = |\cos(\theta)| + |\sin(\theta)|$$

Now, we want to maximize this function. To do this, we can use the first derivative test on $f(\theta)$ by breaking the interval $[0, 2\pi]$ into four sections.

Analyzing the function $f(\theta)$

Value of s	$\lvert\cos(\theta)\rvert$	$\lvert\sin(\theta)\rvert$	$f(\theta)$	$f'(\theta)$
$0 < \theta < \pi/2$	$\cos(\theta)$	$\sin(\theta)$	$\cos(\theta) + \sin(\theta)$	$-\sin(\theta) + \cos(\theta)$
$\pi/2 < \theta < \pi$	$-\cos(\theta)$	$\sin(\theta)$	$-\cos(\theta) + \sin(\theta)$	$\sin(\theta) + \cos(\theta)$
$\pi < \theta < 3\pi/2$	$-\cos(\theta)$	$-\sin(\theta)$	$-\cos(\theta) - \sin(\theta)$	$\sin(\theta) - \cos(\theta)$
$3\pi/2 < \theta < 2\pi$	$\cos(\theta)$	$-\sin(\theta)$	$\cos(\theta) - \sin(\theta)$	$-\sin(\theta) - \cos(\theta)$

To apply the first derivative test, we need to find the critical points of the function $f(\theta)$. Recall, that the critical points are the points $\theta \in [0, 2\pi]$ where $f'(\theta)$ either zero or does not exist. We know that for the points $\theta = 0, \pi/2, \pi, 3\pi/2$, the derivative of f does not exist. The minimum value of $f(\mathbf{x})$ is 1 and occurs at four different points in the region R

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \mathbf{x}_3 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \qquad \mathbf{x}_4 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

For the points $\theta = \pi/4, 3\pi/4, 5\pi/4, 7\pi/4$, the derivative f' is zero. We can evaluate $f(\theta)$ at these points and we find that the the maximum occurs when $\theta = \pi/4, 3\pi/4, 5\pi/4 7\pi/4$. Thus, the maximum value of $f(\mathbf{x})$ is $\sqrt{2}$ and is achieved for four different values of $\mathbf{x}$ including

$$\mathbf{x}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad \mathbf{x}_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad \mathbf{x}_3 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \qquad \mathbf{x}_4 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Problem 04 Let $\|\cdot\|$ denote any norm on $\mathbb{R}^m$ and also the induced matrix norm on $\mathbb{R}^{m\times m}$. Let $\rho(A)$ be the *spectral radius* of A; i.e., $\rho(A)|\lambda_i(A)|$, where $\lambda_i(A)$ is the ith eigenvalue of A. Prove $\rho(A) \le \|A\|$.

This problem has been removed from HW 02.

Problem 05 Let $A = \mathbf{u}\mathbf{v}^T$ where $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$. Prove that $\|A\|_2 = \|\mathbf{u}\|_2\|\mathbf{v}\|_2$.

Before we begin the solution itself recall that:

- From page one of Professor Saito's Lecture 05 we know that for any two vectors $\mathbf{v}, \mathbf{x} \in \mathbb{R}^n$, the angle θ between these vectors is given by

$$\cos(\theta) = \frac{\mathbf{v}^T\mathbf{x}}{\|\mathbf{v}\|_2\|\mathbf{x}\|_2}$$

- From page two of Professor Saito's Lecture 05 we know that

  A norm is a function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ that possesses the following properties. For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.

  (a) $\|\mathbf{x}\| \ge 0$ and $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$

  (b) $\|\mathbf{x} + \mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$

  (c) $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|$

Solution to Problem 05 Let $A = \mathbf{u}\mathbf{v}^T$ where $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$. Then, for any $\mathbf{x} \in \mathbb{R}^n$, consider

$$\begin{aligned} A\mathbf{x} &= \mathbf{u}\mathbf{v}^T\mathbf{x}, \\ &= \mathbf{u}\,\|\mathbf{v}\|_2\,\|\mathbf{x}\|_2\cos(\theta) \end{aligned}$$

where θ is the angle between the vector $\mathbf{v}$ and the vector $\mathbf{x}$. Then, by property Problem 05c above, we have

$$\|A\mathbf{x}\|_2 = \Big\|\mathbf{u}\,\|\mathbf{v}\|_2\,\mathbf{u}\|\mathbf{x}\|_2\cos(\theta)\Big\|_2 = \|\mathbf{u}\|_2\,\|\mathbf{v}\|_2\,\|\mathbf{x}\|_2\,|\cos(\theta)|.$$

Assuming $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\| = 1$,

$$\|A\mathbf{x}\| = \|\mathbf{u}\|_2\,\|\mathbf{v}\|_2\,|\cos(\theta)|.$$

Now note that to maximize the ratio $\|A\mathbf{x}\|$ the function $\cos(\theta)$ must be maximized. This will only happen if the angle between the two vectors $\mathbf{v}$ and $\mathbf{x}$ is a multiple of π or 2π. Thus, it must be that $\mathbf{x} \in$ span $\{\mathbf{v}\}$. Setting $\mathbf{x} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ is one maximizer of this function. Hence,

$$\|A\|_2 = \sup_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| = \|\mathbf{u}\|_2\,\|\mathbf{v}\|_2.$$

Problem 06 (**a**) Define the following matrix

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 1 & 3 \end{bmatrix},$$

in MATLAB. Then, compute the 2-norm by the `norm` function, and report the result in a long format (16 digits) via

```
>> format long
>> norm(A)
```

(b) Compute the 2-norm explicitly using the largest eigenvalue of $A^T A$ using the `eig` function, i.e.,

```
>> sqrt(max(eig(A'*A)))
```

Then, compare the result with that of Part (a). What is the relative error between the norm computed in Part (a) and that in Part (b)?

(c) Compute the 1-norm, ∞-norm, and Frobenius norm of A by hand using the formulas derived in the class. Then, using the `norm` function, compare the MATLAB outputs with your hand-computed results. You should check how to use the `norm` function using the `help` utility:

```
>> help norm
```

(d)] Let's load the MATLAB data file

```
>> load HW_01.mat
```

that you used for HW 01 again. It's located on both Piazza and Canvas Then, compute first the coefficient vector by

```
>> a = U'*x;
```

Now, compute $\|\mathbf{x}\|_p$ and $\|\mathbf{a}\|_p$, $p = 1, 2, \infty$, using the `norm` function, and report the results. Which value of p, you got $\|\mathbf{x}\|_p = \|\mathbf{a}\|_p$?

(e) Now, compute the matrix norms, $\|U\|_p$, $p = 1, 2, \infty$ as well as $\|U\|_F$ using the `norm` function, then report the results.

Problem 07 Linear Least Squares: You are meant to do this problem by hand calculation as you would on a test.

(a) Set up the *normal equation* for the linear least squares approximation for the data (1, −1), (2, 3), and (3, 1).

Solution to Problem 07 (a): The least squares approximation is given by the equation $y = c + dx$. Therefore, we have

$$\begin{aligned} c + d(1) &= -1 \\ c + d(2) &= 3 \\ c + d(3) &= 1 \end{aligned}$$

Translating this into a matrix equation yields:

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \tag{3}$$

(b) Solve for the least squares approximation from **Problem 07 (a)**.

Solution to Problem 07 (b): Multiplying both sides of equation (3) by A^T yields:

$$\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 3 \\ 8 \end{bmatrix}$$

Using Gauss-Jordan elimination or the equation for the inverse of a 2×2 matrix to solve this system gives:

$$d = 1 \quad \text{and} \quad c = -1\,.$$

Therefore, the linear least squares approximation, also known as the 'best fit line' to the data in part **(a)** is given by

$$y = -1 + x$$