ACM/IDS 104 APPLIED LINEAR ALGEBRA
PROBLEM SET 4

Please submit your solution as a single PDF file, that contains both the written-up and published code parts, via Gradescope by **9pm Tuesday, November 7**. An example of the submission process is shown here: `https://www.gradescope.com/get_started#student-submission`

- For theoretical problems, please use a pen, not a pencil: it is hard to read scanned submission written by a pencil.
- For coding problems, please convert your MATLAB livescripts (`.mlx`) to PDF by selecting `Live Editor` → `Save` → `Export to PDF` and merge them with the rest of your solution.
- After uploading your submission to Gradescope, please label all pages.

Problem 1. (10 points) Least Squares Solution
Find the least squares solution to the system $Ax = b$, where

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -2 & 3 \\ 1 & 5 & -1 \\ -3 & 1 & 1 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 0 \\ 5 \\ 6 \\ 8 \end{bmatrix} \tag{1}$$

What is the least squares error?

Problem 2. (10 points) Interpolation for Integration
Most numerical methods for evaluating integrals are based on interpolation. To compute $\int_a^b f(x)dx$, one chooses $n+1$ interpolation points $a \leq x_0 < x_1 < \ldots < x_n \leq b$ and replaces the integrand $f(x)$ by its interpolating polynomial $p_n(x)$ of degree n, leading to the approximation:

$$\int_a^b f(x)dx \approx \int_a^b p_n(x)dx, \tag{2}$$

where the polynomial integral can be done explicitly.

(a) (3 points) Find approximation (2) for $x_0 = a$ and $x_1 = b$ (trapezoid rule).
(b) (3 points) Find approximation (2) for $x_0 = a + \frac{1}{3}(b-a)$ and $x_1 = a + \frac{2}{3}(b-a)$ (open rule).
(c) (4 points) Test the trapezoid and open rules for accuracy on the following integrals:

$$\int_0^1 e^x dx \quad \text{and} \quad \int_0^\pi \sin x dx. \tag{3}$$

Problem 3.(10 points) Application of Least Squares to Data Fitting
The dataset `carbig`, a built-in MATLAB dataset[1] contains various characteristics for $m = 406$ automobiles from the 1970's and 1980's. Let x_1, x_2, and y be the `Weight`, `Horsepower`, and `MPG` (miles per gallon), respectively. Let's assume the following theoretical model for the data:

$$y = f(x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2. \tag{4}$$

Let $\beta^* = (\beta_0^*, \beta_1^*, \beta_2^*, \beta_3^*)^T$ denote the best fit to the data, i.e. the vector that minimizes the Euclidean norm of the residual vector $r = (r_1, \ldots, r_m)^T$, where

$$r_i = y^{(i)} - f(x_1^{(i)}, x_2^{(i)}). \tag{5}$$

(a) (5 points) Derive the system of normal equations on β^*.
(b) (5 points) Complete Problem 3 in `PS4.mlx`.

[1] Can be loaded by `load carbig`.

Problem 4.(10 POINTS) POLYNOMIAL INTERPOLATION
Complete Problem 4 in `PS4.mlx`.

Problem 5. (10 POINTS) STABILITY OF THE GRAM-SCHMIDT ALGORITHM
Complete Problem 5 in `PS4.mlx`.