

CSE 203B WI25 Homework 2

Due Time: 11:50 pm, Thursday, Jan. 23, 2025.

Submit to Gradescope: <https://gradescope.com/>

In this homework, we work on exercises from the textbook including midpoint convexity (2.3), Voronoi diagram (2.7), quadratic function (2.10), general sets (2.12), cones and dual cones (2.28, 2.31, 2.32), and separation of cones (2.39). Extra assignments are given on solving systems of linear equations and support vector machines.

Total points: 50.

Please make two separate submissions on Gradescope:

Exercises:

- Graded by completion, you may work in a group of up to four students.
- Submit a single PDF file and add all group members to the submission.
- Describe each member's contributions at the beginning of your report.

Assignments:

- Graded by content and must be completed individually.
- Submit a single PDF file.

I. Exercises from Textbook Chapter 2 (8 pts, 1pt for each problem)

2.3, 2.7, 2.10, 2.12, 2.28, 2.31, 2.32, 2.39.

II. Assignments (42 pts)

For each problem:

- Clearly show the steps involved in deriving the solution. Simply providing the final answer without explanation may result in point deductions.
- You may use any software tools (e.g., Python, MATLAB) to assist in solving the problems. However, you must explain how the tools were used to arrive at the solution.
- When showing numerical results with decimal values, please round up to at least four digits after the decimal point.

Part One (26 pts)

A matrix solver is one of the fundamental tools in linear algebra. Thus, it is important to grasp the basic concepts (e.g., condition number, machine precision, error, residual). In the following, we will start with a small case and then increase the size of the problem to test the concepts.

Given an equation $\mathbf{Ax} = \mathbf{b}$, we find the solution vector \mathbf{x} that satisfies the equation. Let's start with a small system of linear equations where

$$\mathbf{A} = \begin{bmatrix} 4 & 12 & -8 \\ 12 & 37 & -19 \\ -8 & -19 & 50 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 12 \\ 59 \\ 127 \end{bmatrix}.$$

1.1 Gaussian Elimination

Write down the augmented matrix, and apply Gaussian elimination to derive \mathbf{x} . (4 pts)

1.2 LU Decomposition

Derive the LU decomposition of \mathbf{A} , and apply the result to solve $\mathbf{Ax} = \mathbf{b}$. (4 pts)

1.3 Conjugate Gradient Method (CG)

Run CG with initial guess $\hat{\mathbf{x}}^{(0)} = \mathbf{0}$ and tolerance for the relative residual $\frac{\|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\|_2}{\|\mathbf{b}\|_2} \leq 10^{-8}$.

Report $\hat{\mathbf{x}}^{(k)}$ after each iteration and plot the relative residual curve in the log scale. (4 pts)

1.4 Generalized Minimal Residual Method (GMRES)

Run GMRES with initial guess $\hat{\mathbf{x}}^{(0)} = \mathbf{0}$ and tolerance for the relative residual $\frac{\|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\|_2}{\|\mathbf{b}\|_2} \leq 10^{-8}$.

Report $\hat{\mathbf{x}}^{(k)}$ after each iteration and plot the relative residual curve in the log scale. (4 pts)

Now, let's solve a large system of linear equations where matrix $\mathbf{A} \in R^{1000 \times 1000}$, and vectors \mathbf{x} and $\mathbf{b} \in R^{1000}$. The matrix \mathbf{A} and vector \mathbf{b} are provided in `A.csv` and `b.csv`, respectively. Use software tools to implement and run the four methods (Gaussian Elimination, LU Decomposition, CG, and GMRES) with the provided \mathbf{A} and \mathbf{b} . For the iterative methods, set the initial guess $\hat{\mathbf{x}}^{(0)} = \mathbf{0}$ and tolerance for the relative residual $\leq 10^{-8}$. Answer the following questions:

1.5 Large Matrix Equation (10 pts, 1 pt for each subproblem)

1.5.1 What is the Euclidean norm of the solution vector \mathbf{x} ?

1.5.2 What is the trace of matrix \mathbf{U} after LU decomposition?

1.5.3 Plot the relative residual curve in the log scale for CG.

1.5.4 Plot the relative residual curve in the log scale for GMRES.

1.5.5 Is the hardware you used 32-bit or 64-bit? What is the machine precision provided?

1.5.6 Calculate the condition number of matrix \mathbf{A} , which is defined as $\|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2$.

1.5.7 What is the meaning of the condition number? How to interpret it?

1.5.8 Briefly explain the relationship between relative error $\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2}$, relative residual $\frac{\|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\|_2}{\|\mathbf{b}\|_2}$, machine precision, and condition number.

1.5.9 Report the computation times (in seconds) for each approach. Which method runs the fastest?

1.5.10 Briefly discuss the pros and cons of each method.

Part Two (16 pts)

Support vector machine (SVM): Given a set of points $\{(x_i, y_i) \mid i = 1, \dots, m\}$, where $x_i \in R^n$, and $y_i \in \{-1, 1\}$. We find a hyperplane with vector $a \in R^n$ and bias $b \in R$ to minimize the following objective function.

$$\min_{a,b} \|a\|_2^2, \quad a \in R^n, \quad b \in R \quad (1)$$

$$\text{s.t. } y_i(a^T x_i - b) \geq 1, \quad i = 1, \dots, m \quad (2)$$

2.1 State the conditions with which the above formulation can have valid solutions. (3 pts)

2.2 Create a numerical example with a “feasible” solution. Let us set $m = 7$, $n = 2$.

Use software tools to derive the solution and demonstrate the result in a 2D plot. (5 pts)

2.3 Create a numerical example with an “infeasible” solution. Let us set $m = 7$, $n = 2$.

Plot the set of points you choose and explain why this example is infeasible. (3 pts)

2.4 Revise the formulation so that we can derive a solution for the case created in 2.3.

What is the revised formulation? (Hint: Soft Margin SVM) (2 pts)

Use software tools to derive the solution and demonstrate the result in a 2D plot. (3 pts)