Introduction

Real world problem (e.g ML problem)

$\downarrow$ Modeling

Opt problem

$\downarrow$ Linear Algebra techniques

Soln to opt problem

$\downarrow$ Python

Implement

Module 1 : Linear

Module 2 : Least Squares

1. Linear $\rightarrow$ Q1 , Q2, Q5

2. Optimization workflow
3. Least Squares + Modelling

$\}$ Q3, Q4

Midterm Solutions

Problem 1 (Linear Independence, Matrices) [6pts].

Consider three vectors $u, v, w \in \mathbb{R}^3$ that are linearly independent, and let $a, b, c \in \mathbb{R}^3$ be defined as

$$a = u + v,\ b = v + w,\ c = u + w,\ d = u - v.$$

a. Suppose vectors u and v have unit length. What is the angle between the vectors a and d?

b. Again suppose vectors u and v have unit length. Give an expression for a left-inverse of the matrix $A = \begin{bmatrix} a & d \end{bmatrix}$ (your expression can depend on a and d).

c. Show that nullspace of the matrix $B = \begin{bmatrix} a & b & c \end{bmatrix}$ contains only the zero vector $\mathbf{0} \in \mathbb{R}^3$.

(a) $a^T d = \|a\| \, \|d\| \cos\theta$

$$a^T d = (u+v)^T (u-v) = \overbrace{\cancel{u^T u}}^{1} + \overbrace{\cancel{v^T u} - \cancel{v^T u}}^{0} - \overbrace{\cancel{v^T v}}^{1}$$

$$= 0$$

$\cos\theta = 0$ $\quad \boxed{\theta = \frac{\pi}{2}}$

(b) $A = \begin{bmatrix} | & | \\ a & d \\ | & | \end{bmatrix}$

Approach 1: More interesting

$$XA = I$$

$$\begin{bmatrix} a^T / \|a\|^2 \\ d^T / \|d\|^2 \end{bmatrix} \begin{bmatrix} a & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Approach 2: Computational

$$XA = I$$

(c) $B = [\,a \quad b \quad c\,]$

Given u, v, w are LI

Goal a, b, c are LI

$$BX = 0$$

$$x_1 (u+v) + x_2 (v+w) + x_3 (u+w) = 0$$

$$(x_1+x_3)\, u + (x_1+x_2)\, v + (x_2+x_3)\, w = 0$$

$\Rightarrow \quad x_1 = x_2 = x_3$

Problem 2 (Matrix inverse & properties) [6pts]. Consider the following $n \times n$ matrix,

$$S = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \dots & 1 & 0 \\ 1 & 1 & \dots & 1 & 1 \end{bmatrix}$$

(a)
What does the matrix S do when applied to a vector? Find the inverse of S: Solve $SX = I$ for the unknown matrix X by writing the linear equations each column of X should satisfy. What is the interpretation of S^{-1}? (c)

(a) $$S x = \begin{bmatrix} x_1 \\ x_1 + x_2 \\ \vdots \\ x_1 + x_2 - - - x_n \end{bmatrix}$$ "Running Sum"

(b) $$SX = I$$

$$\begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \dots & 1 & 0 \\ 1 & 1 & \dots & 1 & 1 \end{bmatrix} \begin{bmatrix} | & & | & & | \\ x_1 & - - - & x_j & & x_n \\ | & & | & & | \end{bmatrix} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

$$S x_j = e_j$$

$$x_j = \begin{bmatrix} x_{1j} \\ \vdots \\ x_{jj} \\ \vdots \\ x_{nj} \end{bmatrix} \longrightarrow 1$$

$$X_{1j} = 0$$

$$\cancel{X_{1j}} + X_{2j} = 0$$

$$\vdots$$

$$X_{1j} + \dots X_{jj} = 1$$

$$\underbrace{X_{1j} + X_{2j}}_{0} + \underbrace{X_{jj}}_{1} + \underbrace{X_{j+1,j}}_{-1} = 0$$

$$X_j = \begin{array}{cl} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{array}{l} \\ \\ \\ j \\ (j+1)\text{th} \\ \\ \\ \\ \end{array} \end{array}$$

$$S^{-1} = \begin{bmatrix} 1 & 0 & \\ -1 & 1 & \\ | & -1 & \\ 0 & | & \\ & 0 & \\ 1 & 1 & \end{bmatrix}$$

(c) $$\begin{bmatrix} x_1 \\ -x_1 + x_2 \\ -x_2 + x_3 \end{bmatrix}$$ "Differencing matrix"

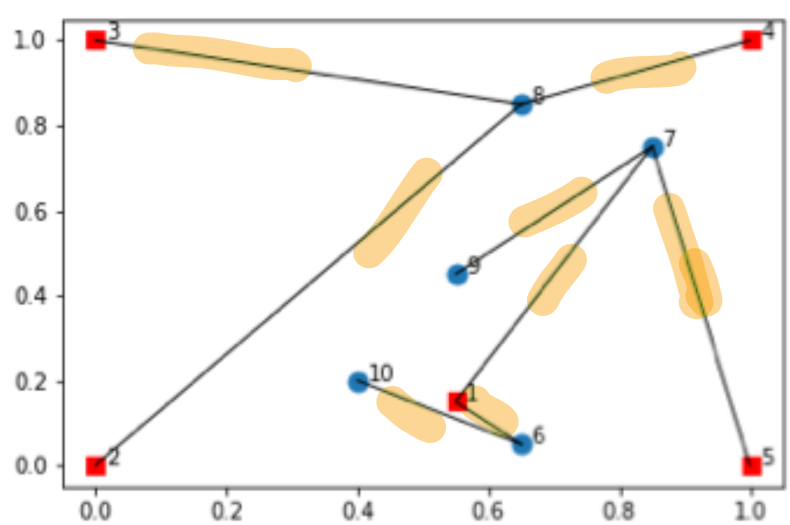
Problem 3 (Least Squares Placement) [18pts]. This problem has parts a–e. We have provided space below each problem for the solution.

The vectors $p_1, \ldots, p_N$, each in $\mathbb{R}^2$ represent the locations of N objects. There are two types of objects: factories and warehouses. The first K objects are factories, whose locations are fixed and given. Our goal in the placement problem to choose the locations of the last $N - K$ objects i.e the warehouses.

Our choice of the locations is guided by an undirected graph; an edge between two objects means we would like them to be close to each other. In least squares placement, we choose the locations $p_{K+1}, \ldots, p_N$ of warehouses so as to minimize the sum of the squares of the distances between objects connected by an edge, where the L edges of the graph are given by the set E. For a specific location of factories $p_1, \ldots, p_K$, we can frame our task as solving the following optimization:

$$g(p_1, \ldots p_K) = \min_{p_{K+1}, \ldots p_N} \sum_{(i,j) \in E} \|p_i - p_j\|^2$$

For illustration, see the figure below. The factories are denoted by red squares, and warehouses by blue circles; we want to move the blue circles so that the objective is minimized.



(a) (iv) E_1 = Edges from warehouse to warehouse

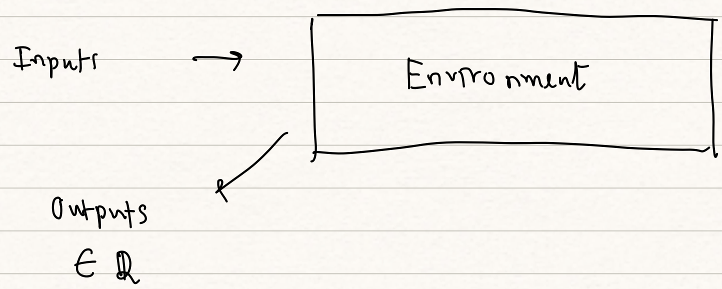
E_2 = " " factory to factory

$$\min_{\substack{\text{Locations of}\\ \text{warehouses}}} \sum_{e \in E} \| \quad \|^2$$

$$\min_{\substack{\text{Locations of}\\ \text{warehouses}}} \sum_{e \in E_1} \| - \|^2 + \sum_{e \in E_2} \| \quad \|^2$$

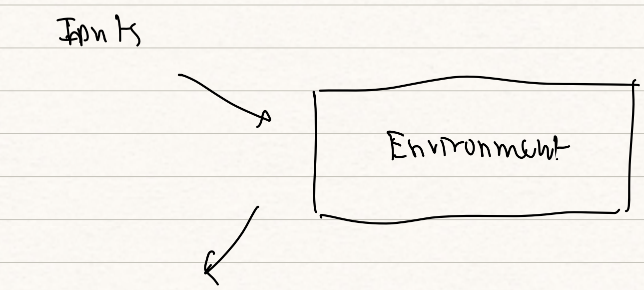
$$\sum_{e \in E_2} \quad + \quad \min_{\substack{\text{Locations of}\\ \text{warehouses}}} \sum_{e \in E_1} \| - \|^2 \rightarrow 0$$

CLASSIFICATION



$$\hat{f}(x) = \theta_0 + \theta_1 x_1 \cdots \theta_d x_d$$

$$\min \; \| \hat{y} - y \|^2$$



Outputs $\in \{1 \ldots K\}$

$y^{(i)} \in \{1 \ldots K\}$ <u>Assume</u> $y^{(i)} \in \mathbb{R}$

$$\min_{\hat{f}} \sum_{i=1}^{m} \mathbb{I}\left(\hat{f}(x^{(i)}) = y^{(i)} \right)$$