

IS5152 Data-driven Decision Making
SEMESTER 2 2023-2024
Assignment 2

- Upload a softcopy of your solution as a **pdf** document to Canvas.
 - Deadline: 11.59 pm, Friday, 1 March, 2024.
-

1. (10 points) Consider the following dataset with 2 feature attributes x_1 and x_2 and a target attribute Class:

ID #	x_1	x_2	Class
1	5	1	No
2	6	2	No
3	6	4	No
4	3	4	No
5	8	2	Yes
6	8	4	Yes
7	5	6	Yes
8	6	7	Yes

- (a) State explicitly the quadratic programming problem that needs to be solved to find the hyperplane with maximum margin of separation in the original feature space.
- (b) Plot the data points. Which data points are likely to be support vectors?
- (c) Check if the points that you identify in part (b) are indeed support vectors by (i) finding their corresponding Lagrange multipliers, and (ii) by showing that the KT optimality conditions are satisfied.
2. (10 points) The Ministry of Education wants to evaluate the efficiency of four elementary schools in the following towns: Clementi, Ang Mo Kio, Yishun and Bedok. The three outputs of the schools are defined to be:
- Output 1 = average reading score
 - Output 2 = average writing score
 - Output 3 = average science score

The three inputs to the schools are defined to be:

- Input 1 = average educational level of mothers (defined by highest grade completed; 12 = high school graduate, 16 = university graduate, and so on).
- Input 2 = number of parent visits to the school per child.
- Input 3 = teacher to student ratio.

The relevant information for the four schools is given in the table below.

School	Inputs			Outputs		
	1	2	3	1	2	3
1 (Clementi)	13	4	0.05	90	70	60
2 (Ang Mo Kio)	14	5	0.05	95	80	75
3 (Yishun)	11	6	0.06	90	76	80
4 (Bedok)	15	8	0.08	90	90	90

The Data Envelopment Analysis method is used to identify school(s) that could be inefficient. The Linear Programming outputs from SAS for each of the schools are as follows.

School 1:

The LP Procedure						
Constraint Summary						
Row	Constraint Name	Type	S/S		Activity	Dual Activity
1	z	OBJECTIVE	Col	Rhs	1	.
2	School1	GE	7	0	0	-1
3	School2	GE	8	0	0	0
4	School3	GE	9	0	0.26562	0
5	School4	GE	10	0	0.6859933	0
6	Input1	EQ	.	1	1	1

School 2:

The LP Procedure						
Constraint Summary						
Row	Constraint Name	Type	S/S		Activity	Dual Activity
1	z	OBJECTIVE	Col	Rhs	1	.
2	School1	GE	7	0	0.0541158	0
3	School2	GE	8	0	0	-1
4	School3	GE	9	0	0.2511358	0
5	School4	GE	10	0	0.6485758	0
6	Input2	EQ	.	1	1	1

School 3:

The LP Procedure						
Constraint Summary						
Row	Constraint Name	Type	S/S		Activity	Dual Activity
1	z	OBJECTIVE	Col	Rhs	1	.
2	School1	GE	7	0	0.0113119	0
3	School2	GE	8	0	0	0
4	School3	GE	9	0	0	-1
5	School4	GE	10	0	0.3461923	0
6	Input3	EQ	.	1	1	1

School 4:

The LP Procedure
Constraint Summary

Row	Constraint		S/S		Dual	
	Name	Type	Col	Rhs	Activity	Activity
1	z	OBJECTIVE	.	0	0.8761191	.
2	School1	GE	7	0	0.0637822	0
3	School2	GE	8	0	0	-0.060565
4	School3	GE	9	0	0	-1.120458
5	School4	GE	10	0	0.1238809	0
6	Input4	EQ	.	1	1	0.8781965

- (a) Determine which (if any) schools are inefficient.
- (b) For each school that is inefficient, explain in terms of the inputs and outputs, why it is not efficient.
3. (10 points) The Clementi Garment Company manufactures men shirts and women blouses. A shirt requires an average of 4 hours and a blouse an average 6 hours in the cutting and sewing center owned by the company. The workers at this center work a normal shift of 8 hours a day, 6 days a week. There are a total of 30 workers employed by the company. The unit profits for shirts and blouses are \$5 and \$10, respectively. The demands require that at least 200 shirts and 300 blouses be produced each week. For each unit of either shirt or blouse by which production falls short of demand, a penalty of \$2.50 is assessed. A \$10 penalty is incurred for each hour of overtime labor and a \$2 penalty is incurred for each hour of available labor that is unused.
- (a) Formulate a linear program that can be used to minimize the total penalty incurred by the company. Clearly define all your decision variables.
- (b) The company wishes to achieve the following goals, which are listed in order of importance (highest to lowest):
- Achieve the profit goal of at least \$3000 per week.
 - Meet demand for blouses.
 - Meet demand for shirts.
 - Avoid having any underutilization of production capacity.
 - Do not use any overtime.

Formulate and solve a preemptive goal programming model for this problem.