

CS 145 - Homework 3

Zoey Nguyen

zoeyn@ucla.edu

June 2, 2021

Question 1.

The support vectors are the ones with ID 1, 5, and 17. They are the only ones with nonzero Lagrangian multipliers, and only active points or support vectors will have nonzero multipliers, all others are zero.

Normal vector and bias.

$$\begin{aligned}w &= \sum_s \lambda_s y_s x_s \\&= (0.5084)(1)(0.91, 0.32) + (0.4625)(1)(0.41, 2.04) + (0.9709)(-1)(2.05, 1.54) \\&= (-1.338, -0.3889) \\b &= \sum_s \lambda_s y_s \\&= 0.5084 + 0.4625 - 0.9709 \\&= 0\end{aligned}$$

The label of x_{new} is the following.

$$\begin{aligned}y &= \text{sign}(w \cdot x_{new} + b) \\&= \text{sign}((-1.338, -0.3889) \cdot (2, -0.5)) \\&= -1\end{aligned}$$

Question 2.

String representations for each tree.

- T1: 5 1 5 -1 -1 6 2 -1 3
- T2: 5 1 -1 6
- T3: 5 1 -1 3
- T4: 5 1 -1 6 -1 3

T2 and T3 are embedded subtrees of T1, but T4 is not.

Only T2 is an embedded subtree of T1.

Question 3.

Frequent subtrees.

- 1-subtrees: A, B, C, D
- 2-subtrees: AA, AB, AC, AD, BA, BB, BC, BD, CA, CB
- 3-subtrees: ACB, BBC, BCA

Question 4.

Minimum DFS codes for each graph. $G3 < G2 < G1$

- G1: axc, cxf, cyb, cyd, cxa, dxe, cye
- G2: axb, byc, cxa, cyd, dxa, dxg, cye
- G3: axb, byc, cxa, axd, dxe, eyc, eyg, gxf

gSpan frequent subgraphs without single-node graphs. There are 20 frequent subgraphs in total.

- Two-node: ab x, ac x, ad x, bc y, cf x, cd y, ce y, de x
- Three-node: abc xxy, adc xxy, ade xx, ace xy, bcd yy, bce yy, cde yyx
- Four-node: abcd, abce, acde, bcde
- Five-node: abcde

Canonical adjacency matrix.

$$\begin{pmatrix} f & 0 & 0 & 0 & 0 & 0 \\ 0 & e & 0 & 0 & 0 & 0 \\ 0 & x & d & 0 & 0 & 0 \\ x & y & y & c & 0 & 0 \\ 0 & 0 & 0 & y & b & 0 \\ 0 & 0 & x & x & 0 & a \end{pmatrix}$$