

1. Naszkicuj na wykresie hiperpłaszczyznę separującą określoną wzorem  $x_1 + 2x_2 - 3 = 0$ . Dorysuj również hiperpłaszczyzny określone wzorem:  $x_1 + 2x_2 - 3 = 1$  oraz  $x_1 + 2x_2 - 3 = -2$ . Zaznacz na wykresie punkty dla których  $x_1 + 2x_2 - 3 > 0$ .
2. Naszkicuj poniższe obserwacje na płaszczyźnie, a następnie:
  - zaznacz hiperpłaszczyznę separującą klasy o największym marginesie,
  - podaj wzór na tę hiperpłaszczyznę,
  - nanieś na wykres kierunek wektora wag.

| $x_1$ | $x_2$ | $y$ |
|-------|-------|-----|
| 1     | 1     | +1  |
| 3     | 2     | +1  |
| 1     | 4     | +1  |
| 2     | 4     | -1  |
| 5     | 1     | -1  |
| 6     | 3     | -1  |
| 5     | 5     | -1  |
| 5     | 4     | -1  |

3. W zależności od liczby wymiarów  $d$ , jaka jest minimalna liczba obserwacji w zbiorze danych, aby można było określić unikalną hiperpłaszczyznę o maksymalnym marginesie?
4. Oblicz odległość poniższej obserwacji z klasy  $-1$  do hiperpłaszczyzny określonej przez  $w = [-4, 3]$  i  $b = -2$ .

| $x_1$ | $x_2$ | $y$ |
|-------|-------|-----|
| 2     | 3     | -1  |

5. Podaj wzór na wartość marginesu określonej hiperpłaszczyzny separującej o danym wektorze  $w$  i  $b$ .
6. Zdefiniuj problem optymalizacyjny (hard) SVM dla danych liniowoseparowalnych. Przyjmij następujący zbiór uczący:

| $x_1$ | $x_2$ | $y$ |
|-------|-------|-----|
| -1    | 7     | +1  |
| 2     | 3     | -1  |
| 4     | 2     | +1  |

7. Zdefiniuj problem optymalizacyjny (soft) SVM dla następującego zbioru uczącego:

| $x_1$ | $x_2$ | $y$ |
|-------|-------|-----|
| -1    | 7     | +1  |
| 2     | 3     | -1  |
| 4     | 2     | +1  |

8. Jaka jest interpretacja wartości współczynników Lagrange'a w formulacji dualnej problemu SVM? Odpowiedź uzasadnij odwołując się do warunków KKT.
9. W formulacji dualnej problemu SVM optymalizowane są zmienne, które zwykle oznaczamy jako  $\alpha_i$  i nie są to wagi. W jaki sposób zatem znajdowana jest hiperpłaszczyzna separująca?
10. Przekształć problem (soft) SVM do problemu bez ograniczeń. Jakie są podobieństwa i różnice pomiędzy klasyfikatorem regresji logistycznej a klasyfikatorem SVM?