Fundamentals of Linear Algebra and Optimization
Classification of Data Points: Terminology

Jean Gallier and Jocelyn Quaintance

CIS Department
University of Pennsylvania

jean@cis.upenn.edu

April 25, 2024

Classification of Data Points for $(\mathrm{SVM}_{s2'})$

In this module we introduce the concepts necessary to discuss a classification of the points u_i and v_j in terms of *Lagrange multipliers*.

If $(w, \eta, \epsilon, \xi, b)$ is an optimal solution of Problem $(\mathrm{SVM}_{s2'})$ with $w \neq 0$ and $\eta \neq 0$, then the complementary slackness conditions yield a classification of the points u_i and v_j in terms of the values of λ and μ.

Indeed, we have $\epsilon_i \alpha_i = 0$ for $i = 1, \ldots, p$ and $\xi_j \beta_j = 0$ for $j = 1, \ldots, q$.

Classification of Data Points for $(\mathrm{SVM}_{s2'})$

Also, if $\lambda_i > 0$, then the corresponding constraint is active, and similarly if $\mu_j > 0$. Since $\lambda_i + \alpha_i = K_s$, it follows that $\epsilon_i \alpha_i = 0$ iff $\epsilon_i(K_s - \lambda_i) = 0$, and since $\mu_j + \beta_j = K_s$, we have $\xi_j \beta_j = 0$ iff $\xi_j(K_s - \mu_j) = 0$.

Thus if $\epsilon_i > 0$, then $\lambda_i = K_s$, and if $\xi_j > 0$, then $\mu_j = K_s$.

Also, if $\lambda_i < K_s$, then $\epsilon_i = 0$ and u_i is *correctly classified*, and similarly if $\mu_j < K_s$, then $\xi_j = 0$ and v_j is *correctly classified*.

Definition of Support Vectors

Definition. The vectors u_i on the blue margin $H_{w,b+\eta}$ and the vectors v_j on the red margin $H_{w,b-\eta}$ are called *support vectors*. Support vectors correspond to vectors u_i for which $w^\top u_i - b - \eta = 0$ (which implies $\epsilon_i = 0$), and vectors v_j for which $w^\top v_j - b + \eta = 0$ (which implies $\xi_j = 0$).

Support vectors u_i such that $0 < \lambda_i < K_s$ and support vectors v_j such that $0 < \mu_j < K_s$ are *support vectors of type 1*.

Support Vectors of Type 1 and Type 2

Support vectors of type 1 play a special role so we denote the sets of indices associated with them by

$$I_\lambda = \{i \in \{1, \ldots, p\} \mid 0 < \lambda_i < K_s\}$$
$$I_\mu = \{j \in \{1, \ldots, q\} \mid 0 < \mu_j < K_s\}.$$

We denote their cardinalities by $numsvl_1 = |I_\lambda|$ and $numsvm_1 = |I_\mu|$.

Support vectors u_i such that $\lambda_i = K_s$ and support vectors v_j such that $\mu_j = K_s$ are *support vectors of type 2*.

Exceptional Support Vectors; Failing the Margin

Support vectors u_i such that $\lambda_i = 0$ and support vectors v_j such that $\mu_j = 0$ are *exceptional support vectors*.

The vectors u_i for which $\lambda_i = K_s$ and the vectors v_j for which $\mu_j = K_s$ are said to *fail the margin*.

The sets of indices associated with the vectors failing the margin are denoted by

$$K_\lambda = \{i \in \{1, \ldots, p\} \mid \lambda_i = K_s\}$$
$$K_\mu = \{j \in \{1, \ldots, q\} \mid \mu_j = K_s\}.$$

We denote their cardinalities by $p_f = |K_\lambda|$ and $q_f = |K_\mu|$.

Definition of Margin at Most δ

Definition. Vectors u_i such that $\lambda_i > 0$ and vectors v_j such that $\mu_j > 0$ are said to *have margin at most δ*.

The sets of indices associated with these vectors are denoted by

$$I_{\lambda>0} = \{i \in \{1, \ldots, p\} \mid \lambda_i > 0\}$$
$$I_{\mu>0} = \{j \in \{1, \ldots, q\} \mid \mu_j > 0\}.$$

We denote their cardinalities by $p_m = |I_{\lambda>0}|$ and $q_m = |I_{\mu>0}|$.

Definition of Strictly Failing the Margin

Vectors u_i such that $\epsilon_i > 0$ and vectors v_j such that $\xi_j > 0$ are said to *strictly fail the margin*.

The corresponding sets of indices are denoted by

$$E_\lambda = \{i \in \{1, \dots, p\} \mid \epsilon_i > 0\}$$
$$E_\mu = \{j \in \{1, \dots, q\} \mid \xi_j > 0\}.$$

We write $p_{sf} = |E_\lambda|$ and $q_{sf} = |E_\mu|$.

Classification of the Points

We have the inclusions $E_\lambda \subseteq K_\lambda$ and $E_\mu \subseteq K_\mu$.

The difference between the first sets and the second sets is that the second sets may contain support vectors of type 1 such that $\lambda_i = K_s$ and $\epsilon_i = 0$, or $\mu_j = K_s$ and $\xi_j = 0$.

We also have the equations $I_\lambda \cup (K_\lambda - E_\lambda) \cup E_\lambda = I_{\lambda>0}$ and $I_\mu \cup (K_\mu - E_\mu) \cup E_\mu = I_{\mu>0}$, and the inequalities $p_{sf} \leq p_f \leq p_m$ and $q_{sf} \leq q_f \leq q_m$.

Classification of the Points

The blue points u_i of index $i \in I_{\lambda>0}$ are classified as follows:

(1) If $i \in I_\lambda$, then u_i is a support vector of type 1 ($\lambda_i < K_s$).

(2) If $i \in K_\lambda - E_\lambda$, then u_i is a support vector of type 2 ($\lambda_i = K_s$).

(3) If $i \in E_\lambda$, then u_i strictly fails the margin, that is $\epsilon_i > 0$.

Similarly the red points v_j of index $j \in I_{\mu>0}$ are classified as follows:

(1) If $j \in I_\mu$, then v_j is a support vector of type 1 ($\mu_j < K_s$).

(2) If $j \in K_\mu - E_\mu$, then v_j is a support vector of type 2 ($\mu_j = K_s$).

(3) If $j \in E_\mu$, then v_j strictly fails the margin, that is $\xi_j > 0$.

Classification of the Points

Note that $p_m - p_f$ is the number of blue support vectors of type 1 and $q_m - q_f$ is the number of red support vectors of type 1.

The remaining blue points u_i for which $\lambda_i = 0$ are either exceptional support vectors or they are (strictly) in the open half-space corresponding to the blue side.

Similarly, the remaining red points v_j for which $\mu_j = 0$ are either exceptional support vectors or they are (strictly) in the open half-space corresponding to the red side.

Classification of the Points

In the example below (from the last lesson) , we have

$$numsvl1 = 2, numsvm1 = 1,\ p_{sf} = p_f = 2,\ q_{sf} = q_f = 3,\ p_m = 4,\ q_m = 4.$$

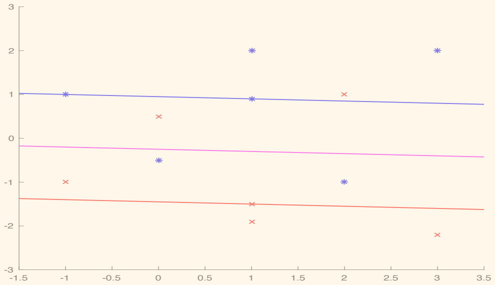

Figure 1: Soft margin ν-SVM for two sets of six points for $\nu = 0.6$.