Lecture Notes on Machine Learning: Minimum Enclosing Balls

Technical Report · February 2019

2 authors:

Christian Bauckhage
University of Bonn

Tiansi Dong
Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS

Lecture Notes on Machine Learning
Minimum Enclosing Balls

Christian Bauckhage and Tiansi Dong
B-IT, University of Bonn

The minimum enclosing ball problem is another example of a problem that can be cast as a constrained convex optimization problem. It thus provides us with another opportunity to work with the Karush-Kuhn-Tucker (KKT) conditions and to derive a Lagrangian dual which, as we shall see later, allows for a simple solution for the problem.

Setting the Stage

Given a data set $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \subset \mathbb{R}^m$, the ***minimum enclosing ball*** problem is to determine the smallest m-ball

$$\mathcal{B}(\boldsymbol{c}, r) = \left\{ \boldsymbol{x} \in \mathbb{R}^m \;\middle|\; \left\| \boldsymbol{x} - \boldsymbol{c} \right\| \leq r \right\} \tag{1}$$

that contains all the points in $\mathcal{X}$.

Since balls are defined in terms of their center $\boldsymbol{c} \in \mathbb{R}^m$ and radius $r \in \mathbb{R}$, the basic problem therefore is to determine optimal choices $\boldsymbol{c}^*$ and r^* for these two parameters.

In this note, we approach this problem from the point of view of constrained convex optimization.[1] Indeed, noting the equivalencies

$$\left\| \boldsymbol{x} - \boldsymbol{c} \right\| \leq r \quad \Leftrightarrow \quad \left\| \boldsymbol{x} - \boldsymbol{c} \right\|^2 \leq r^2 \quad \Leftrightarrow \quad \left\| \boldsymbol{x} - \boldsymbol{c} \right\|^2 - r^2 \leq 0, \tag{2}$$

we can immediately formalize the problem of finding $\boldsymbol{c}^*$ and r^* as an inequality constrained quadratic minimization problem, namely

$$\begin{aligned} \boldsymbol{c}^*, r^* = \operatorname*{argmin}_{\boldsymbol{c}, r} \quad & r^2 \\ \text{s.t.} \quad & \left\| \boldsymbol{x}_j - \boldsymbol{c} \right\|^2 - r^2 \leq 0 \qquad j = 1, \ldots, n. \end{aligned} \tag{3}$$

In plain English, this means that we are simultaneously searching for an appropriate center and radius. The radius should be as small as possible but not smaller. That is, the distance between each of the given data points and the center of the ball must be less than or equal to the radius of the ball.

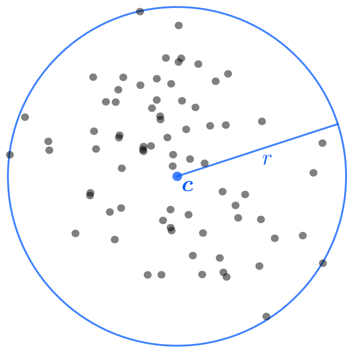

Figure 1: Minimum enclosing ball of a data set $\mathcal{X} \subset \mathbb{R}^2$.

[1] Note that there are various other ideas for how to tackle the minimum enclosing ball problem; Wikipedia lists some of them.

Our goal in this note is to show that the ***primal (minimization) problem*** in (3) has the following ***dual (maximization) problem***

$$\begin{aligned} \boldsymbol{\mu}^* = \operatorname*{argmax}_{\boldsymbol{\mu}} \; & \sum_{j=1}^{n} \mu_j \, \boldsymbol{x}_j^\intercal \boldsymbol{x}_j - \left[\sum_{j=1}^{n} \mu_j \, \boldsymbol{x}_j \right]^\intercal \left[\sum_{j=1}^{n} \mu_j \, \boldsymbol{x}_j \right] \\ \text{s.t.} \; & \begin{aligned} \boldsymbol{\mu}^\intercal \mathbf{1} &= 1 \\ \boldsymbol{\mu} &\succeq \mathbf{0} \end{aligned} \end{aligned} \tag{4}$$

where $\boldsymbol{\mu} \in \mathbb{R}^n$ is a vector of Lagrange multipliers and $\mathbf{1} \in \mathbb{R}^n$ denotes the vector of all ones.

Deriving the Dual Problem

To begin with, we write down the Lagrangian of the problem in (3). Letting μ_j denote the Lagrange multiplier for the j-th inequality in (3), we have

$$\mathcal{L}(\boldsymbol{c}, r, \boldsymbol{\mu}) = r^2 + \sum_{j=1}^{n} \mu_j \left(\left\| \boldsymbol{x}_j - \boldsymbol{c} \right\|^2 - r^2 \right) \tag{5}$$

$$= r^2 + \sum_{j=1}^{n} \mu_j \left(\boldsymbol{x}_j^\intercal \boldsymbol{x}_j - 2\, \boldsymbol{c}^\intercal \boldsymbol{x}_j + \boldsymbol{c}^\intercal \boldsymbol{c} - r^2 \right) \tag{6}$$

$$= r^2 + \sum_{j=1}^{n} \mu_j\, z_j - 2\, \boldsymbol{c}^\intercal \left[\sum_{j=1}^{n} \boldsymbol{x}_j\, \mu_j \right] + \boldsymbol{c}^\intercal \boldsymbol{c} \sum_{j=1}^{n} \mu_j - r^2 \sum_{j=1}^{n} \mu_j \tag{7}$$

Observe that, when expanding the expression in (6) to the one in (7), we introduced the shorthand

$$z_j = \boldsymbol{x}_j^\intercal \boldsymbol{x}_j \tag{8}$$

Note: Henceforth, we work with the substitution

$$z_j = \boldsymbol{x}_j^\intercal \boldsymbol{x}_j.$$

just to reduce notational clutter.

In fact, we can write our Lagrangian even more compactly. To this end, we note that

$$\sum_{j=1}^{n} \mu_j = \boldsymbol{\mu}^\intercal \mathbf{1} \tag{9}$$

$$\sum_{j=1}^{n} \mu_j\, z_j = \boldsymbol{\mu}^\intercal \boldsymbol{z} \tag{10}$$

and

$$\sum_{j=1}^{n} \boldsymbol{x}_j\, \mu_j = \boldsymbol{X}\, \boldsymbol{\mu} \tag{11}$$

Note: Henceforth, we gather the $\boldsymbol{x}_j \in \mathcal{X}$ in an $m \times n$ data matrix

$$\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n].$$

where $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n] \in \mathbb{R}^{m \times n}$ is a data matrix whose columns correspond to the data points in $\mathcal{X}$. Hence, the Lagrangian in (7) can also be written as

$$\mathcal{L}(\boldsymbol{c}, r, \boldsymbol{\mu}) = r^2 + \boldsymbol{\mu}^\intercal \boldsymbol{z} - 2\, \boldsymbol{c}^\intercal \boldsymbol{X}\, \boldsymbol{\mu} + \boldsymbol{c}^\intercal \boldsymbol{c} \cdot \boldsymbol{\mu}^\intercal \mathbf{1} - r^2\, \boldsymbol{\mu}^\intercal \mathbf{1}. \tag{12}$$

Lagrangian of the minimum enclosing ball problem

Given this Lagrangian, we next evaluate the Karush-Kuhn-Tucker conditions. Considering the KKT 1 conditions (stationarity), we have

$$\mathbf{0} \stackrel{!}{=} \frac{\partial \mathcal{L}}{\partial \boldsymbol{c}} = -2\, \boldsymbol{X}\, \boldsymbol{\mu} + 2\left(\boldsymbol{\mu}^\intercal \mathbf{1}\right) \boldsymbol{c} \qquad \Rightarrow \qquad \boldsymbol{c} = \frac{\boldsymbol{X}\, \boldsymbol{\mu}}{\boldsymbol{\mu}^\intercal \mathbf{1}} \tag{13}$$

and

$$0 \stackrel{!}{=} \frac{\partial \mathcal{L}}{\partial r} = r - r\, \boldsymbol{\mu}^\intercal \mathbf{1} = r\left(1 - \boldsymbol{\mu}^\intercal \mathbf{1}\right) \qquad \Rightarrow \qquad \boldsymbol{\mu}^\intercal \mathbf{1} = 1. \tag{14}$$

Plugging the result in (14) into the one in (13), the latter simplifies to

$$\boldsymbol{c} = \boldsymbol{X}\,\boldsymbol{\mu}. \tag{15}$$

Plugging the expressions in (14) and (15) back into the Lagrangian in (12), we find the Lagrangian dual, because

$$\begin{aligned}
\mathcal{L}(\boldsymbol{c}, r, \boldsymbol{\mu}) &= r^2 + \boldsymbol{\mu}^\intercal \boldsymbol{z} - 2\,\boldsymbol{c}^\intercal \boldsymbol{X}\,\boldsymbol{\mu} + \boldsymbol{c}^\intercal \boldsymbol{c} \cdot \boldsymbol{\mu}^\intercal \mathbf{1} - r^2\,\boldsymbol{\mu}^\intercal \mathbf{1} && (16)\\
&= r^2 + \boldsymbol{\mu}^\intercal \boldsymbol{z} - 2\,\boldsymbol{c}^\intercal \boldsymbol{X}\,\boldsymbol{\mu} + \boldsymbol{c}^\intercal \boldsymbol{c} - r^2 && (17)\\
&= \boldsymbol{\mu}^\intercal \boldsymbol{z} - 2\,\boldsymbol{\mu}^\intercal \boldsymbol{X}^\intercal \boldsymbol{X}\,\boldsymbol{\mu} + \boldsymbol{\mu}^\intercal \boldsymbol{X}^\intercal \boldsymbol{X}\,\boldsymbol{\mu} && (18)\\
&= \boldsymbol{\mu}^\intercal \boldsymbol{z} - \boldsymbol{\mu}^\intercal \boldsymbol{X}^\intercal \boldsymbol{X}\,\boldsymbol{\mu} && (19)\\
&= \mathcal{D}(\boldsymbol{\mu}). && (20)
\end{aligned}$$

Lagrangian dual of the minimum enclosing ball problem

Now, if we further note that the KKT 3 conditions (dual feasibility) demand $\boldsymbol{\mu} \succeq \mathbf{0}$, all of this is to say that

$$\begin{aligned}
\boldsymbol{\mu}^* = \underset{\boldsymbol{\mu}}{\operatorname{argmax}} \;\; & \boldsymbol{\mu}^\intercal \boldsymbol{z} - \boldsymbol{\mu}^\intercal \boldsymbol{X}^\intercal \boldsymbol{X}\,\boldsymbol{\mu} \\
\text{s.t.} \;\; & \boldsymbol{\mu}^\intercal \mathbf{1} = 1 \\
& \boldsymbol{\mu} \succeq \mathbf{0}
\end{aligned} \tag{21}$$

Exercise: Convince yourself that

$$\boldsymbol{\mu}^\intercal \boldsymbol{X}^\intercal \boldsymbol{X}\,\boldsymbol{\mu} = \left[\sum_{j=1}^{n} \mu_j\,\boldsymbol{x}_j\right]^\intercal \left[\sum_{j=1}^{n} \mu_j\,\boldsymbol{x}_j\right].$$

is indeed the dual of the primal problem in (3).

SOLVING THE DUAL in (21) provides $\boldsymbol{\mu}^*$ which, in turn, allows us to compute the center of the minimum enclosing ball. According to (15), we find

$$\boldsymbol{c}^* = \boldsymbol{X}\,\boldsymbol{\mu}^*. \tag{22}$$

center of the minimum enclosing ball

But what about the corresponding radius r^*?

THE RADIUS r^* of the minimum enclosing ball can be obtained from considering the KKT 4 conditions (complementary slackness). These dictate that, at a solution, we must have

$$\mu_j^* \left(z_j - 2\,\boldsymbol{c}^{*\intercal} \boldsymbol{x}_j + \boldsymbol{c}^{*\intercal} \boldsymbol{c}^* - r^{*2} \right) = 0 \qquad j = 1, \ldots, n. \tag{23}$$

Summing these equations over all j then yields

$$\sum_{j=1}^{n} \mu_j^* \left(z_j - 2\,\boldsymbol{c}^{*\intercal} \boldsymbol{x}_j + \boldsymbol{c}^{*\intercal} \boldsymbol{c}^* - r^{*2} \right) = 0 \tag{24}$$

for which it is fairly easy to see (considering our computations above) that some additional algebra leads to

$$r^{*2} = \boldsymbol{\mu}^{*\intercal} \boldsymbol{z} - \boldsymbol{\mu}^{*\intercal} \boldsymbol{X}^\intercal \boldsymbol{X}\,\boldsymbol{\mu}^*. \tag{25}$$

In other words, we find that the radius of the minimum enclosing ball for the given data amounts to

$$r^* = \sqrt{\mathcal{D}(\boldsymbol{\mu}^*)}. \tag{26}$$

radius of the minimum enclosing ball

Summary and Outlook

In this note, we were concerned with the minimum enclosing ball problem which asks for the smallest Euclidean m-ball that contains all the columns of a given data matrix $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n] \in \mathbb{R}^{m \times n}$.

We discussed that the basic problem is to compute the parameters $\boldsymbol{c}^*$ and r^* of an appropriate m-ball and saw that this problem can be formalized as the problem of solving the inequality constrained quadratic minimization problem in (3).

Given this primal problem, we then derived the corresponding Lagrangian dual problem and found it to be given by

$$\begin{aligned} \boldsymbol{\mu}^* = \underset{\boldsymbol{\mu}}{\operatorname{argmax}} \;\; & \boldsymbol{\mu}^\intercal \boldsymbol{z} - \boldsymbol{\mu}^\intercal \boldsymbol{X}^\intercal \boldsymbol{X} \boldsymbol{\mu} \\ \text{s.t.} \;\; & \boldsymbol{\mu}^\intercal \mathbf{1} = 1 \\ & \boldsymbol{\mu} \succeq \mathbf{0} \end{aligned} \tag{27}$$

where $\boldsymbol{\mu} \in \mathbb{R}^n$ is a vector of Lagrange multipliers, $\mathbf{1} \in \mathbb{R}^n$ is the vector of all ones, and the entries of $\boldsymbol{z} \in \mathbb{R}^n$ are given by $z_j = \boldsymbol{x}_j^\intercal \boldsymbol{x}_j$.

We finally saw that, once the solution $\boldsymbol{\mu}^*$ of (27) is available, the center and radius of the minimum enclosing ball we are after can be computed as

$$\boldsymbol{c}^* = \boldsymbol{X} \boldsymbol{\mu}^* \tag{28}$$

$$r^* = \sqrt{\boldsymbol{\mu}^{*\intercal} \boldsymbol{z} - \boldsymbol{\mu}^{*\intercal} \boldsymbol{X}^\intercal \boldsymbol{X} \boldsymbol{\mu}^*}. \tag{29}$$

What we did not discuss, however, is how to actually solve the dual problem in (27). Yet, there is a rather simple algorithm for this purpose and we shall study it in an upcoming note.

Finally, we should point out that the problem we considered in this note is not of purely academic nature. In fact, m-balls provide an interesting data structure for representation learning[2,3] and the minimum enclosing ball problem is closely related to the problem of training certain one-class support vector machines[4]. Practical applications like these will be studied in great detail later on.

[2] T. Dong, Z. Wang, J. Li, C. Bauckhage, and A.B. Cremers. Triple Classification Using Regions and Fine-Grained Entity Typing. In *Proc. AAAI*, 2019b

[3] T. Dong, C. Bauckhage, H. Jin, J. Li, O. Cremers, D. Speicher, A.B. Cremers, and J. Zimmermann. Imposing Category Trees onto Word-Embeddings Using a Geometric Construction. In *Proc. ICLR*, 2019a

[4] D.M.J. Tax and R.P.W. Duin. Support Vector Data Description. *Machine Learing*, 54(1), 2004

Acknowledgments

This material was prepared within project P3ML which is funded by the Ministry of Education and Research of Germany (BMBF) under grant number 01/S17064. The authors gratefully acknowledge this support.

References

T. Dong, C. Bauckhage, H. Jin, J. Li, O. Cremers, D. Speicher, A.B. Cremers, and J. Zimmermann. Imposing Category Trees onto Word-Embeddings Using a Geometric Construction. In *Proc. ICLR*, 2019a.

T. Dong, Z. Wang, J. Li, C. Bauckhage, and A.B. Cremers. Triple Classification Using Regions and Fine-Grained Entity Typing. In *Proc. AAAI*, 2019b.

D.M.J. Tax and R.P.W. Duin. Support Vector Data Description. *Machine Learing*, 54(1), 2004.