Homework 8

1. Exercise A6.5.

2. [Pronzato] Consider a collection of m points $a_1, \ldots, a_m$ in $\mathbf{R}^n$. The smallest ball that contains the m points can be computed by solving the optimization problem

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \|a_i - x\|_2^2 \leq t, \quad i = 1, \ldots, m, \end{array} \tag{1}$$

with variables $x \in \mathbf{R}^n$ and $t \in \mathbf{R}$.

(a) State the optimality (Karush–Kuhn–Tucker) conditions for this problem.

(b) We denote the radius and center of the smallest covering ball by R and c. In other words, $t = R^2$ and $x = c$ are optimal in (1). Let $\bar{a} = (1/m)(a_1 + \cdots + a_m)$ be the mean of the points, and suppose $\lambda_1, \ldots, \lambda_m$ are optimal dual multipliers in the Lagrange dual of (1). Show that

$$\begin{aligned} R^2 + \|c - \bar{a}\|_2^2 &= \sum_{i=1}^m \lambda_i \|a_i - c\|_2^2 + \|c - \bar{a}\|_2^2 & \text{(2a)} \\ &= \sum_{i=1}^m \lambda_i \|a_i - \bar{a}\|_2^2 & \text{(2b)} \\ &\leq \max_{i=1,\ldots,m} \|a_i - \bar{a}\|_2^2 & \text{(2c)} \end{aligned}$$

and that

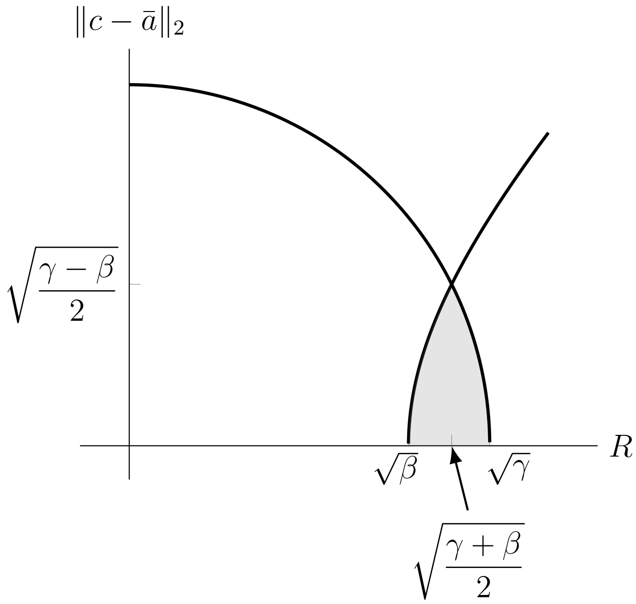
$$\begin{aligned} R^2 - \|c - \bar{a}\|_2^2 &\geq \frac{1}{m} \sum_{i=1}^m \|a_i - c\|_2^2 - \|c - \bar{a}\|_2^2 & \text{(2d)} \\ &= \frac{1}{m} \sum_{i=1}^m \|a_i - \bar{a}\|_2^2. & \text{(2e)} \end{aligned}$$

The inequalities (2c) and (2e) give bounds on R and $\|c - \bar{a}\|_2$ in terms of two quantities

$$\beta = \frac{1}{m} \sum_{i=1}^m \|a_i - \bar{a}\|_2^2, \qquad \gamma = \max_{i=1,\ldots,m} \|a_i - \bar{a}\|_2^2,$$

that are easily computed from the problem data. This is illustrated in the figure below. The shaded region in the figure shows the values of R and $\|c - \bar{a}\|_2$ that satisfy two inequalities

$$R^2 + \|c - \bar{a}\|_2^2 \leq \gamma, \qquad R^2 - \|c - \bar{a}\|_2^2 \geq \beta.$$



3. Exercise A12.12.

4. Exercise A8.1.

5. Exercise A8.7. *Hint.* The Hessian of the cost function can be written as

$$H = I + A^T(\mathbf{diag}(z) - zz^T)A$$

where z is a positive vector with $\mathbf{1}^T z = 1$. Show that

$$H = I + B^T \mathbf{diag}(z)^{-1} B,$$

where $B = (\mathbf{diag}(z) - zz^T)A$, and follow the approach of page 10–30 of the lecture slides.