THREE-WAY LINEAR CLASSIFICATION (10 points possible)

We are given data

$$x^{(1)}, \ldots, x^{(N)}, \quad y^{(1)}, \ldots, y^{(M)}, \quad z^{(1)}, \ldots, z^{(P)},$$

three nonempty sets of vectors in $\mathbf{R}^n$. We wish to find three affine functions on $\mathbf{R}^n$,

$$f_i(z) = a_i^T z - b_i, \quad i = 1, 2, 3,$$

that satisfy the following properties:

$$\begin{aligned} f_1(x^{(j)}) &> \max\{f_2(x^{(j)}), f_3(x^{(j)})\}, \quad j = 1, \ldots, N, \\ f_2(y^{(j)}) &> \max\{f_1(y^{(j)}), f_3(y^{(j)})\}, \quad j = 1, \ldots, M, \\ f_3(z^{(j)}) &> \max\{f_1(z^{(j)}), f_2(z^{(j)})\}, \quad j = 1, \ldots, P. \end{aligned}$$

In words: f_1 is the largest of the three functions on the x data points, f_2 is the largest of the three functions on the y data points, f_3 is the largest of the three functions on the z data points. We can give a simple geometric interpretation: The functions f_1, f_2, and f_3 partition $\mathbf{R}^n$ into three regions,

$$\begin{aligned} R_1 &= \{z \mid f_1(z) > \max\{f_2(z), f_3(z)\}\}, \\ R_2 &= \{z \mid f_2(z) > \max\{f_1(z), f_3(z)\}\}, \\ R_3 &= \{z \mid f_3(z) > \max\{f_1(z), f_2(z)\}\}, \end{aligned}$$

defined by where each function is the largest of the three. Our goal is to find functions with $x^{(j)} \in R_1$, $y^{(j)} \in R_2$, and $z^{(j)} \in R_3$.

Pose this as a convex optimization problem. You may not use strict inequalities in your formulation.

Solve the specific instance of the 3-way separation problem given in `sep3way_data.m`, with the columns of the matrices X, Y and Z giving the $x^{(j)}$, $j = 1, \ldots, N$, $y^{(j)}$, $j = 1, \ldots, M$ and $z^{(j)}$, $j = 1, \ldots, P$. To save you the trouble of plotting data points and separation boundaries, we have included the plotting code in `sep3way_data.m`. (Note that `a1`, `a2`, `a3`, `b1` and `b2` contain arbitrary numbers; you should compute the correct values using CVX.)

Which of the following statements most accurately describe the correct plot?

- ○ A small number of points are misclassified.
- ○ All points are correctly classified.

Show Answer

You have used 0 of 2 submissions

Help

FITTING A SPHERE TO DATA (10 points possible)

Consider the problem of fitting a sphere $\{x \in \mathbf{R}^n \mid \|x - x_c\|_2 = r\}$ to m points $u_1, \ldots, u_m \in \mathbf{R}^n$, by minimizing the error function

$$\sum_{i=1}^{m} \left(\|u_i - x_c\|_2^2 - r^2 \right)^2$$

over the variables $x_c \in \mathbf{R}^n$, $r \in \mathbf{R}$.

Formulate this problem as a convex or quasiconvex optimization problem. The simpler your formulation, the better. (For example: a convex formulation is simpler than a quasiconvex formulation; an LP is simpler than an SOCP, which is simpler than an SDP.)

Use your method to solve the problem instance with data given in the file `sphere_fit_data.m`, with $n = 2$. Plot the fitted circle and the data points.

Typesetting math: 100%

What is the optimal radius $r^\star$, rounded to two decimal places?

Show Answer You have used 0 of 2 submissions

LEARNING A QUADRATIC PSEUDO-METRIC FROM DISTANCE MEASUREMENTS (30 points possible)

We are given a set of N pairs of points in $\mathbf{R}^n$, $x_1, \ldots, x_N$, and $y_1, \ldots, y_N$, together with a set of distances $d_1, \ldots, d_N > 0$.

The goal is to find (or estimate or learn) a quadratic pseudo-metric d,

$$d(x, y) = \left((x - y)^T P (x - y) \right)^{1/2},$$

with $P \in \mathbf{S}_+^n$, which approximates the given distances, i.e., $d(x_i, y_i) \approx d_i$. (The pseudo-metric d is a metric only when $P \succ 0$; when $P \succeq 0$ is singular, it is a pseudo-metric.)

To do this, we will choose $P \in \mathbf{S}_+^n$ that minimizes the mean squared error objective

$$\frac{1}{N} \sum_{i=1}^{N} (d_i - d(x_i, y_i))^2.$$

(a) Fill in the blanks of the following exposition to show how to find P using convex optimization.

Solution.

The problem is

$$\text{minimize} \quad \frac{1}{N}\sum_{i=1}^{N}(d_i - d(x_i, y_i))^2$$

with variable $P \in \mathbf{S}^n_+$. This problem can be rewritten as

$$\text{minimize} \quad \frac{1}{N}\sum_{i=1}^{N}(d_i^2 - 2d_i d(x_i, y_i) + d(x_i, y_i)^2),$$

with variable P (which enters through $d(x_i, y_i)$). The objective is convex because each term of the objective can be written as (ignoring the $1/N$ factor)

$$d_i^2 - 2d_i\left((x_i - y_i)^T P(x_i - y_i)\right)^{1/2} + (x_i - y_i)^T P(x_i - y_i),$$

which is convex in P. To see this, note that the first term is _______ and the third term is _______. The middle term is convex because it is _______. ☐

The correct phrases that fill the blank spots are

- constant; quadratic in P; the negation of the composition of a concave function (square root) with a linear function of P
- quadratic; quadratic in P; the composition of a square root with a quadratic function of P and is thereby linear in P
- quadratic; linear in P; the composition of a square root with a quadratic function of P and is thereby linear in P
- constant; linear in P; the negation of the composition of a concave function (square root) with a linear function of P.

(b) Carry out the method of part (a) with the data given in `quad_metric_data_norng.m`. The columns of the matrices `X` and `Y` are the points x_i and y_i; the row vector `d` gives the distances d_i. Give the optimal mean squared distance error rounded to two decimal places.

We also provide a test set, with data `X_test`, `Y_test`, and `d_test`. Report the mean squared distance error on the test set (using the metric found using the data set above) rounded to two decimal places.

Show Answer

You have used 0 of 2 submissions

MAXIMUM VOLUME RECTANGLE INSIDE A POLYHEDRON (10 points possible)

Find the rectangle

$$\mathcal{R} = \{x \in \mathbf{R}^n \,|\, l \preceq x \preceq u\}$$

of maximum volume, enclosed in a polyhedron $\mathcal{P} = \{x \,|\, Ax \preceq b\}$. The variables are $l, u \in \mathbf{R}^n$. Your formulation should not involve an exponential number of constraints. Solve a specific instance of this problem given in `max_vol_box.m`, with data given as `A` and `b`.

Hint. maximizing `sum(log(x))` is equivalent to maximizing `geo_mean(x)`. The latter is SDP representable and therefore is more stable in CVX.

What is the maximum volume? *Hint.* Don't forget to convert the objective value to volume.

Show Answer *You have used 0 of 2 submissions*