Optimization and Algorithms
Project report

Group 42
José Neves 89683, Leonardo Pedroso 89691, and Gustavo Bakker 100660

1 Part 3

1.1 Task 1

The dataset in file `data_opt.csv` is loaded and the corresponding matrix D is computed according to $D_{mn} = ||\mathbf{x_m} - \mathbf{x_n}||_2$. The following MATLAB script solves task 1

```matlab
%% Part 3 - Task 1
%% Load dataset from file data_opt.csv
X = csvread("./data/data_opt.csv");
N = size(X,1); % Get number of datapoints
%% Compute matrix D
D = zeros(N); % Initialize matrix D
for m = 1:N % Four each off-diagonal pair of coordinates
    for n = m+1:N
       D(m,n) = norm(X(m,:)-X(n,:),2); % Compute D_{mn}
       D(n,m) = D(m,n); % D_{nm}=D_{mn}
    end
end
%% Check results
% Find maximum value of distance and repective indices
Dmax = max(max(D));
[mDmax,nDmax] = find(D==Dmax);
% Output results
fprintf("------------------------ Task 1 ------------------------\n");
fprintf("D(2,3) = %g | D(4,5) = %g.\n", D(2,3),D(4,5));
fprintf("max{D(m,n)} = %g for (m,n) = {(%d,%d),(%d,%d)}.\n",...
    Dmax,mDmax(1),nDmax(1),mDmax(2),nDmax(2))
%% Save data
save("./data/distancesTask1.mat",'D','Dmax','nDmax','mDmax','N');
```

obtaining

$$D_{2,3} = 5.8749, \quad D_{4,5} = 24.3769$$

and

$$\max(D_{mn}) = 83.003 \quad \text{for} \quad (m,n) \in \{(134,33),(33,134)\}\,.$$

1.2 Task 2

One has

$$f(\mathbf{y}) = \sum_{m=1}^{N} \sum_{n=m+1}^{N} (||\mathbf{y_m} - \mathbf{y_n}|| - D_{mn})^2 = \sum_{m=1}^{N} \sum_{n=m+1}^{N} f_{mn}(\mathbf{y})^2 \,, \tag{1}$$

where $\mathbf{y_m} \in \mathbb{R}^m$, k is the dimension of the target space, $\mathbf{y} = \mathrm{col}(\mathbf{y_1}, \ldots, \mathbf{y_N}) \in \mathbb{R}^{Nk}$ is the optimization variable, and

$$f_{mn}(\mathbf{y}) := ||\mathbf{y_m} - \mathbf{y_n}|| - D_{mn} \,. \tag{2}$$

Note that one can write $\mathbf{y_m} = \mathbf{E_m}\mathbf{y}$, where $\mathbf{E_m} \in \mathbb{R}^{k \times Nk}$ is defined as

$$\mathbf{E_m} := \begin{bmatrix} \mathbf{0}_{k\times k(m-1)} & \mathbf{I}_{k\times k} & \mathbf{0}_{k\times k(N-m)} \end{bmatrix} \,,$$

thus, it is possible to rewrite (2) as

$$f_{mn}(\mathbf{y}) = ||\mathbf{E_m}\mathbf{y} - \mathbf{E_n}\mathbf{y}|| - D_{mn} = \sqrt{\mathbf{y}^T(\mathbf{E_m} - \mathbf{E_n})^T(\mathbf{E_m} - \mathbf{E_n})\mathbf{y}} - D_{mn} \,. \tag{3}$$

Taking the jacobian of (3), one obtains

$$\begin{aligned} D_{\mathbf{y}} f_{m,n}(\mathbf{y}) &= D_u(\sqrt{u})\Big|_{u=\mathbf{y}^T(\mathbf{E_m}-\mathbf{E_n})^T(\mathbf{E_m}-\mathbf{E_n})\mathbf{y}} D_{\mathbf{y}}(\mathbf{y}^T(\mathbf{E_m} - \mathbf{E_n})^T(\mathbf{E_m} - \mathbf{E_n})\mathbf{y}) \\ &= \mathbf{y}^T \frac{(\mathbf{E_m} - \mathbf{E_n})^T(\mathbf{E_m} - \mathbf{E_n})}{\sqrt{\mathbf{y}^T(\mathbf{E_m} - \mathbf{E_n})^T(\mathbf{E_m} - \mathbf{E_n})\mathbf{y}}} \end{aligned}$$

therefore the gradient $\nabla_{\mathbf{y}} f_{mn}(\mathbf{y}) = (D_{\mathbf{y}} f_{mn}(\mathbf{y}))^T$ is given by

$$\nabla_{\mathbf{y}} f_{mn}(\mathbf{y}) = \frac{(\mathbf{E_m} - \mathbf{E_n})^T(\mathbf{E_m} - \mathbf{E_n})}{f_{mn}(\mathbf{y}) + D_{mn}}\mathbf{y} \,. \tag{4}$$

Similarly taking the jacobian of (1), one obtains

$$D_{\mathbf{y}} f(\mathbf{y}) = \sum_{m=1}^{N} \sum_{n=m+1}^{N} D_u(u^2)\Big|_{u=f_{mn}(\mathbf{y})} D_{\mathbf{y}} f_{mn}(\mathbf{y}) = \sum_{m=1}^{N} \sum_{n=m+1}^{N} 2 f_{mn}(\mathbf{y}) D_{\mathbf{y}} f_{mn}(\mathbf{y})$$

therefore the gradient $\nabla_{\mathbf{y}} f(\mathbf{y}) = (D_{\mathbf{y}} f(\mathbf{y}))^T$ is given by

$$\nabla_{\mathbf{y}} f(\mathbf{y}) = \sum_{m=1}^{N} \sum_{n=m+1}^{N} 2 f_{mn}(\mathbf{y}) \nabla_{\mathbf{y}} f_{mn}(\mathbf{y}) \tag{5}$$

In conclusion, $f(\mathbf{y})$, $f_{mn}(\mathbf{y})$,$\nabla_{\mathbf{y}} f_{mn}(\mathbf{y})$, and $\nabla_{\mathbf{y}} f(\mathbf{y})$ can be computed making use of (1),(2), (4), and (5), respectively.

For the implementation of the Levenberg-Marquardt (LM) method is is required to compute,for each new iteration, matrix $\mathbf{A}$ and vector $\mathbf{b}$, defined by

$$\mathbf{A} := \begin{bmatrix} \nabla_{\mathbf{y}} f_{1,1}(\mathbf{y}) \\ \nabla_{\mathbf{y}} f_{1,2}(\mathbf{y}) \\ \vdots \\ \nabla_{\mathbf{y}} f_{N-1,N}(\mathbf{y}) \\ \sqrt{\lambda}\mathbf{I}_{Nk\times Nk} \end{bmatrix} \quad \text{and} \quad \mathbf{b} := \begin{bmatrix} \nabla_{\mathbf{y}} f_{1,1}(\mathbf{y})^T\mathbf{y} - f_{1,1}(\mathbf{y}) \\ \nabla_{\mathbf{y}} f_{1,2}(\mathbf{y})^T\mathbf{y} - f_{1,2}(\mathbf{y}) \\ \vdots \\ \nabla_{\mathbf{y}} f_{N-1,N}(\mathbf{y})^T\mathbf{y} - f_{N-1,N}(\mathbf{y}) \\ \sqrt{\lambda}\mathbf{y} \end{bmatrix}. \tag{6}$$