Exercises Week 12: Optimization on nonsmooth sets through lifts

Instructor: Nicolas Boumal
TA: Andreea Musat (andreea.musat@epfl.ch)

May 22, 2025

We consider the optimization problem defined over the product manifold $(\mathbb{S}^{d-1})^n$, where each $x_i \in \mathbb{S}^{d-1} \subset \mathbb{R}^d$, and the cost function is given by

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}\, \varphi(x_i^\top x_j),$$

where $w_{ij} \geq 0$ are weights and $\varphi\colon [-1, 1] \to \mathbb{R}$ is a smooth function. In this exercise, we assume $w_{ij} = 1$ for all i, j.

When φ is increasing, the global maximizers of f are so-called "synchornized configurations", where all vectors are equal, i.e., $x_1 = x_2 = \cdots = x_n$. However, since the domain is non-convex, it is not a given that optimization algorithms always converge to this configuration. Our goal is to investigate whether Riemannian optimization methods can recover the global maximizer in practice, and under what conditions.

Implement this optimization task using (Py)Manopt, and experiment with different settings. Consider dimensions $d \in \{2, 3, 4\}$ and several different values for n, and the following choices of φ:

(1) $\varphi(t) = t$

(2) $\varphi(t) = t^2$

(3) $\varphi(t) = t^3$

(4) $\varphi(t) = \exp(\beta t)$, with several values of β

Use a very small gradient norm tolerance (eg: 10^{-12}) as the stopping criterion. For each configuration, perform multiple runs with random initializations. Visualize how the point configuration evolves, and analyze the optimization outcome.

At convergence, examine the norm of the Riemannian gradient and the spectrum of the Hessian (available via `hessianspectrum` in Manopt). This allows you to verify whether the algorithm terminated at a local maximum.

What do you observe in your experiments? Are all points collapsing to a single direction (consensus)? When consensus fails, what kind of configurations emerge? Does φ affect the convergence behavior? How does increasing β in the exponential case change results? What happens with the points during the optimization process, do you observe any patterns?

You can use the following plotting utilities:

```
%% Plotting function for visualization
function hdots = plot_dots(X)
    hold all;
    dotscolors = X(1, :) + X(2, :);

    if size(X, 1) == 3
        [Sx, Sy, Sz] = sphere(50);
        surf(Sx, Sy, Sz, 'FaceAlpha', 0.25, 'FaceColor', 2 * [0.1, 0.2, 0.3]);
        hdots = scatter3(X(1, :)', X(2, :)', X(3, :)', 200, dotscolors, 'filled');
    elseif size(X, 1) == 2
        t = linspace(0, 2*pi, 501);
        plot(cos(t), sin(t), 'k-', 'LineWidth', 2);
        hdots = scatter(X(1, :)', X(2, :)', 200, dotscolors, 'filled');
    end

    axis equal;
    axis off;
    set(gcf, 'Color', 'w');
end

%% Function to update plot and track stats
function stats = update_dots(hdots, X, time, stats)
    set(hdots, 'XData', X(1, :));
    set(hdots, 'YData', X(2, :));

    if size(X, 1) == 3
        set(hdots, 'ZData', X(3, :));
    end

    drawnow;

    if exist('time', 'var') && ~isempty(time)
        pause(time);
    end

    if ~exist('stats', 'var')
        stats = [];
    end
end
```

Listing 1: Functions for plotting and updating point configurations