Minimum Volume Ellipsoid

Kathleen Chang

Stanford University

May 30, 2022

Outline

Minimum Volume Ellipsoid

If $C \in \mathbb{R}^n$ is bounded with a nonempty interior, the minimum volume ellipsoid $\mathcal{E}$ covers C. $\mathcal{E}$ is also called the Löwner-John ellipsoid.

$$\mathcal{E} = \{v \,|\, ||Av + b||_2 \, \leq 1 \,\} \text{ where } A \in \mathcal{S}^n_{++}$$

The convex optimization problem to solve for the minimum volume ellipsoid is (Boyd, 2011):

$$\begin{aligned} \min \quad & \text{log det}\, A^{-1} \\ \text{s.t.} \quad & \sup_{v \in C} ||Av + b||_2 \leq 1 \end{aligned} \tag{1}$$

This selects the minimum volume ellipsoid from points in C, knowing vol $\mathcal{E}$ is proportional to det A^{-1}.

Minimum Volume Ellipsoid from a Finite Set

If C is a **finite set** where $C = \{x_1, ..., x_m\} \in \mathbb{R}^n$, the problem is formulated as:

$$\begin{aligned} \min \quad & \text{log det}\, A^{-1} \\ \text{s.t.} \quad & ||Ax_i + b||_2 \leq 1 \quad \text{for} \quad i = 1, ..., m \end{aligned} \tag{2}$$

Outline

Dual Problem

The dual is obtained by lifting the primal problem to another dimension (Moshtagh, 06).

$$X_{dual} = X_{primal} \cup H$$

$$H = \{(x, x_{n+1}) \in \mathbb{R}^{n+1} | x_{n+1} = 1\}$$

The dual is

$$\begin{aligned} \max \quad & g(u) = \log \det(XUX^T) \\ \text{s.t.} \quad & \mathbf{1}^T u = 1 \\ & u \geq 0 \end{aligned} \tag{3}$$

where $X = X_{dual} = [x_1, ..., x_m]$, $U = Diag(u)$

Mahalanobis Distance

$$M(x) = \sqrt{(\vec{x} - \vec{\mu})^T\, \mathbf{S}^{-1}\, (\vec{x} - \vec{\mu})}$$

- Generalization to multidimensional space of a point's position $\vec{x}$ in a distribution using covariance matrix **S**
- Distribution $\vec{\mu}$ has a Mahalanobis distance is 0

$$M(z, \mathcal{E}) = \textbf{inf}\{\, t \geq 0 \,|\, z \in c + t(\mathcal{E} - c)\,\}$$

- Factor to scale an ellipsoid $\mathcal{E}$ about its center such that a point z lies on the boundary of the ellipsoid

Dual Intuition

- u lies on the unit simplex making it a probability distribution over the set X
- By solving the dual problem, we can find the Mahalanobis distance for each point x_k as $x_k(XUX^T)^{-1}x_k$

Outline

Minimum Volume Ellipsoid

Dual problem

Outlier identification

Motivation

- Minimum-volume ellipsoid and its dual are convex optimization problems
- Outlier detection techniques presented remove 'most' binding ellipsoid points and drastically reduce ellipsoid volume
- Useful for statistical analysis of outliers
- Operates in multivariate space as a ranking tool

Ellipsoid Peeling

Remove or peel points from set C by solving the minimum ellipsoid problem until h points remain where h $\leq$ N (Ahipaşaoğlu, 2015)

---

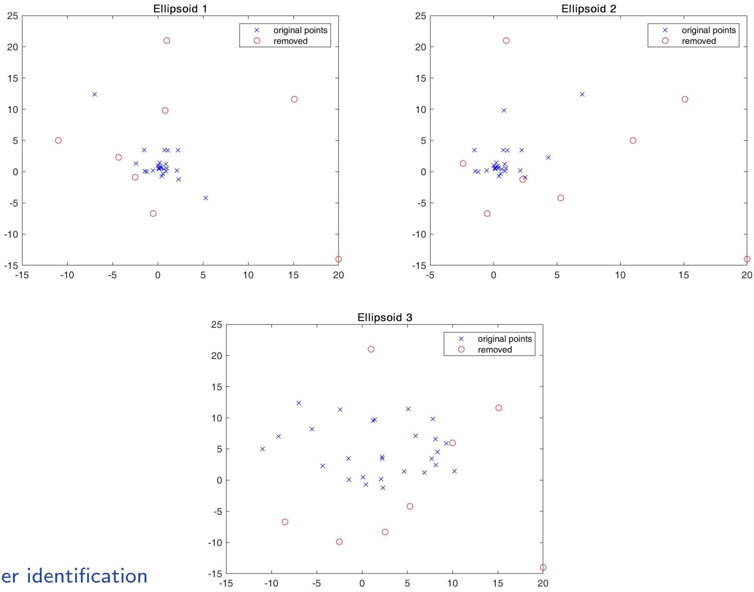
given $X = \{x_1, ..., x_N\} \in \mathbb{R}^n$

repeat

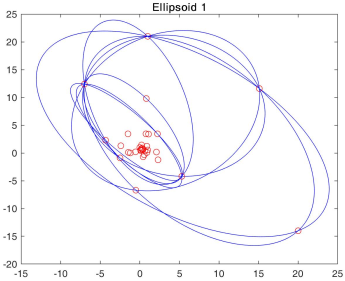
1. Solve for minimum volume ellipsoid $\mathcal{E}$
2. **break if** h points remain
3. Remove the point with the largest dual variable for the constraint $x_i \in \mathcal{E}$

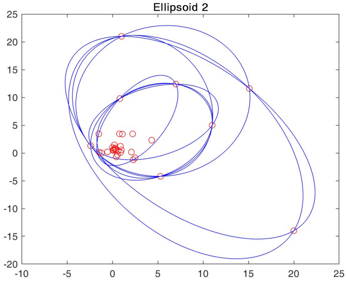
return $X = \{x_1, ..., x_h\}$ and $\mathcal{E}$.

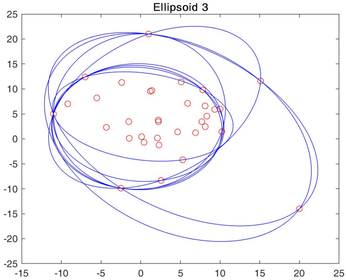
Ellipsoid Peeling



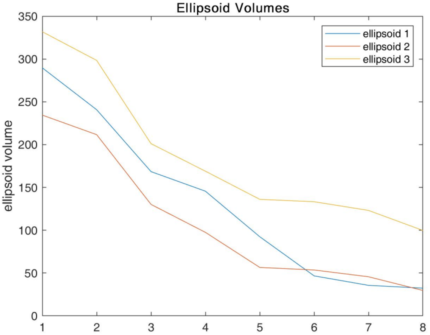
Ellipsoid Peeling







Ellipsoid Peeling



Ellipsoid Peeling

Ellipsoid $\in \mathbb{R}^2$	Original Volume	Reduced Volume	% Reduction
1	289.65	32.36	88.8%
2	234.44	29.69	87.3%
3	331.81	99.76	69.9%

- Remove 8 points $\in \mathbb{R}^2$ from a set of 35
- Not as effective on uniformly distributed points (ellipsoid 3)
- Similar performance regardless of outlier position (ellipsoid 1, uniformly placed outside center cloud vs 2, right skew)

Ellipsoid Ordering

Remove points associated with the Mahalanobis distance until h points remain where h $\leq$ N (Ahipaşaoğlu, 2015)

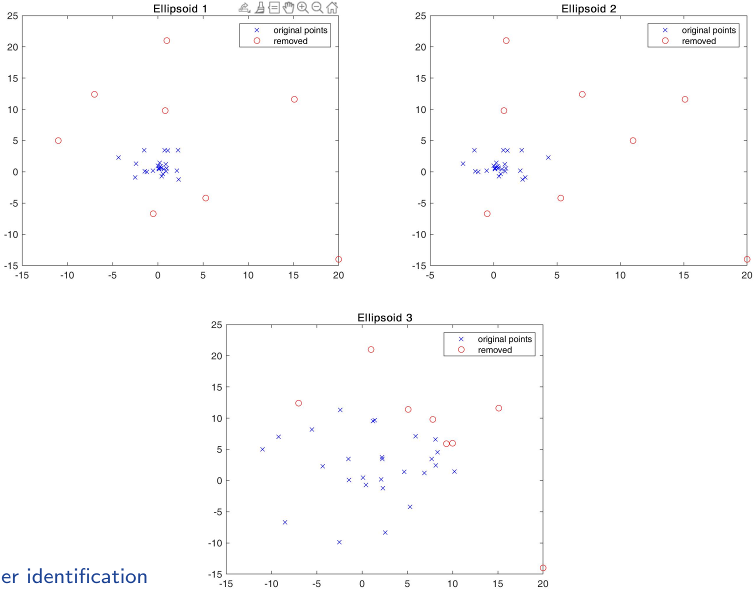
---

given $X = \{x_1, ..., x_N\} \in \mathbb{R}^n$

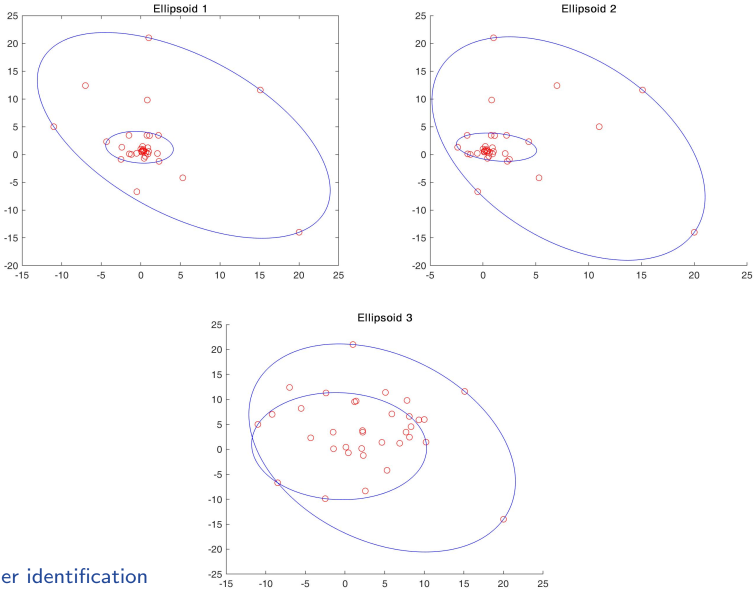
1. Solve the minimum volume ellipsoid dual problem.
2. Evaluate $\xi_k = x_k(XUX^T)x_k$ for all points.
3. Remove the points corresponding to the $N - h$ largest values of ξ_k.

return $X = \{x_1, ..., x_h\}$ and $\mathcal{E}$.

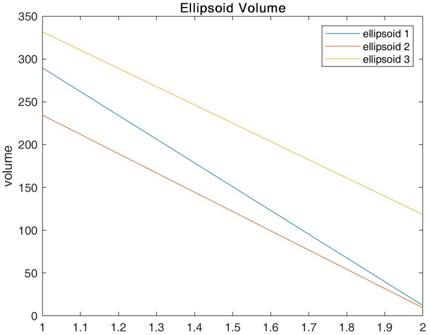
Ellipsoid Ordering



Ellipsoid Ordering



Ellipsoid Ordering



Ellipsoid Ordering

Ellipsoid $\in \mathbb{R}^2$	Original Volume	Reduced Volume	% Reduction
1	289.65	12.12	95.8%
2	234.44	9.38	96%
3	331.81	118.34	64.3%

- Remove 8 points $\in \mathbb{R}^2$ from a set of 35
- More effective than ellipsoid peeling at reducing volume
- Removing points based on a distance proportional to ellipsoid volume

Random Ellipsoid Ordering

Add N-h or remove h points starting out with random points (Ahipaşaoğlu, 2015)

---

given $X = \{x_1, ..., x_N\} \in \mathbb{R}^n$

Select z = n + 1 points to span $\mathbb{R}^n$, include these in the minimum volume ellipsoid

1. Solve the minimum volume ellipsoid dual problem.
2. Evaluate $\xi_k = x_k(XUX^T)x_k$ for all points.
3. **if** z $<$ h Add h - z x_k with the smallest Mahalanobis distance
   if z $>$ h Remove the z - h x_k with the largest Mahalanobis distance

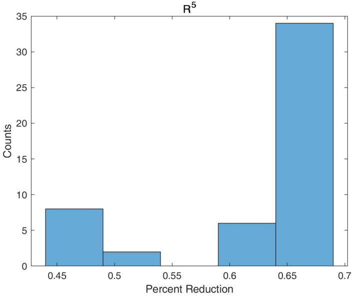
return $X = \{x_1, ..., x_h\}$ and $\mathcal{E}$.

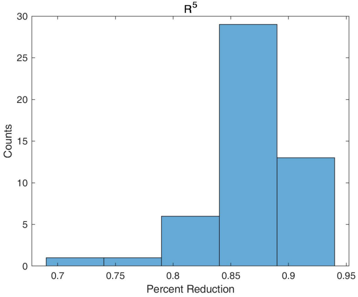
Random Ellipsoid Ordering

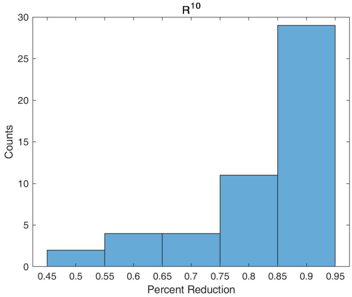
- Perform 50 trials of random ellipsoid ordering (h = 27) on 35 points in $\mathbb{R}^2, \mathbb{R}^5$, and $\mathbb{R}^{10}$
- Trade-off between dimensionality of data and range of percent volume reduction
- Effectiveness of reducing ellipsoid volume depends on initial basis vectors

Ellipsoid	$\mathbb{R}^2$	$\mathbb{R}^5$	$\mathbb{R}^{10}$
Mean % Reduction	60.81%	86.41%	83.03%

Random Ellipsoid Ordering







Random Ellipsoidal Peeling/Building

Add or remove points beginning with a random basis (Ahipaşaoğlu, 2015)

---

given $X = \{x_1, ..., x_N\} \in \mathbb{R}^n$

Select z = n + 1 points to span $\mathbb{R}^n$, include these in the minimum volume ellipsoid
repeat

1. Solve the minimum volume ellipsoid dual problem.
2. **break if** z == h

    if z < h Add the x_k with the smallest $\xi_k = x_k(XUX^T)x_k$

    z = z + 1

    if z > h Remove the x_i with the largest dual value

    z = z - 1

return $X = \{x_1, ..., x_h\}$ and $\mathcal{E}$.

Results

- Ellipsoid ordering reduces volume more than ellipsoid peeling
- Random ellipsoid ordering does not outperform either ellipsoid peeling or ordering in $\mathbb{R}^2$ case
- Larger volume reduction in non-uniformly distributed sets

Citations

- Moshtagh, Nima. (2006). Minimum volume enclosing ellipsoid. http://www.mathworks.com/matlabcentral/fileexchange/9542
- Ahipaşaoğlu, S.D. Fast algorithms for the minimum volume estimator. J Glob Optim 62, 351–370 (2015). https://doi.org/10.1007/s10898-014-0233-8
- Boyd, S., & Vandenberghe, L. (2022). Additional Exercises for Convex Optimization.
- Boyd, S. P., & Vandenberghe, L. (2011). Convex optimization. Cambridge Univ. Pr.