

Optimal Petunin Ellipsoids in the Design Space

Klyushin D.A., Lyashko S.I., Tymoshenko A.A.
 Taras Shevchenko National University of Kyiv
 Kyiv, Ukraine
 dokmed5@gmail.com

Abstract—The problem of construction of optimal Petunin ellipsoids in design space with linear constrains is considered. The algorithms of construction of Petunin ellipsoids without constrains and with linear constrains are described. The issues of the complexity and statistical effectiveness are analyzed.

Keywords—Petunin ellipsoids, minimum volume ellipsoid, minimum trace ellipsoid, constrained optimization, design space.

INTRODUCTION

Modern computer-aided design is impossible without data analysis that describes the parameters of objects and the results of experiments in the form of multidimensional vectors in the design space. The complexity of the problem is due to the fact that the number of such parameters is very large and therefore it is desirable to single out among them those that are important from a certain point of view (for example, statistically significant) and discard points that can be interpreted as outliers. Moreover, each parameter, as a rule, is subject to certain restrictions. It is either limited to the minimum and maximum values, or is related to other parameters by linear relationships. The statistical nature of the source data dictates the choice of the appropriate tool.

In particular, similar problems were considered in [1], where a comparative analysis of several methods for constructing optimal descriptions of design data was carried out, in particular, an ellipsoid of minimum volume, an ellipsoid with a minimum trace, and the Monte Carlo method. Despite the effectiveness of the methods demonstrated, they have a big drawback - they do not take into account the statistical nature of the parameters, focusing on geometric properties. The Petunin ellipse described in papers [3], [4], [6] allows us to elegantly solve this problem.

Let $X = \{x_1, \dots, x_n\}$ is a set of random points, where $x_i \in \mathbb{R}^d$, and $Ax \leq b$ is a system of s linear inequalities, where $A = \{a_{ij}\}_{i,j=1}^{s,d}$. The problem is to find the optimal ellipsoid containing at least given number of points and having minimal volume among the similar ellipsoids and lying within the polygon corresponding to the linear constrains $Ax \leq b$. To describe the algorithm of the construction of Petunin ellipsoids let us consider two cases: $d = 2$ and $d > 2$ without constrains.

PETUNIN ELLIPSE AND ELLIPSOID

Case $d = 2$. Construct a convex hull of $X = \{(x_1, y_1), \dots, (x_n, y_n)\}$ Find a diameter of the convex

hull and its ends (x_k, y_k) and (x_l, y_l) . Construct a segment L connecting these ends. Find points lying at the most distance from L : (x_r, y_r) and (x_q, y_q) .

Construct segments L_1 and L_2 that are parallel to L , pass through (x_r, y_r) and (x_q, y_q) , and have the length equals to the length of L . Construct segments L_3 and L_4 that are orthogonal to L and pass through (x_k, y_k) and (x_l, y_l) . Segments L_1, L_2, L_3 and L_4 form a rectangle Π . Denote a short side by a and a long side by b (fig. 1). Make translation, rotation and

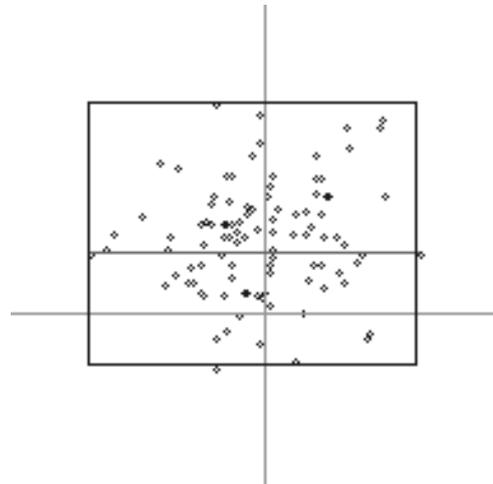


Fig. 1. Petunin rectangle Π

shrinking with a coefficient $\alpha = \frac{a}{b}$ so that Π transforms to a square Π' with a center (x'_0, y'_0) . The random points are transformed to points $(x'_1, y'_1), (x'_2, y'_2), \dots, (x'_n, y'_n) \in \Pi'$. Find distances r_1, r_2, \dots, r_n from (x'_0, y'_0) to $(x'_1, y'_1), (x'_2, y'_2), \dots, (x'_n, y'_n) \in \Pi'$. Let $R = \max(r_1, r_2, \dots, r_n)$.

Construct a circle C with the center (x'_0, y'_0) and radius R . Thus, $(x'_1, y'_1), (x'_2, y'_2), \dots, (x'_n, y'_n) \in C$ (fig. 2). Performing inverse transformations of this circle with obtain an ellipse E (fig. 3). The computational complexity of this algorithm is equal to computational complexity of the construction of a convex hull $O(n \lg n)$. Case $d > 2$. Construct the convex hull of $X = \{x_1, \dots, x_n\}$. Find a diameter of the convex hull and its ends (x_k, y_k) and (x_l, y_l) . Using rotation and translation let us align the diameter along to Ox'_1 . Project all the points $(x'_1, y'_1), (x'_2, y'_2), \dots, (x'_n, y'_n)$ to the orthogonal complement of Ox'_1 .

Using convex hull construction, rotation and translation until the orthogonal complement becomes two-

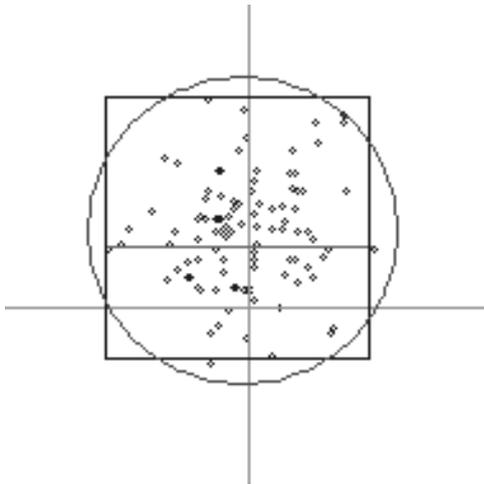


Fig. 2. Petunin circle C

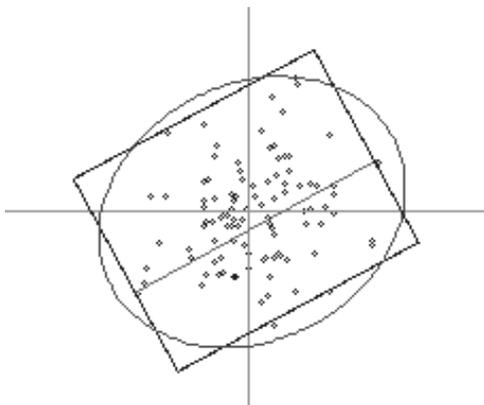


Fig. 3. Petunin ellipse E

dimensional we obtain a rectangle Π . Perform the algorithm for case $d = 2$. Construct a minimum volume axis-aligned parallelogram in d dimensional space containing the images x'_1, \dots, x'_n of input points. Transform this parallelogram to hypercube.

Denote its center by x_0 and find distances r_1, r_2, \dots, r_n from it to x'_1, \dots, x'_n . Let $R = \max(r_1, r_2, \dots, r_n)$. Construct a hypersphere with the center x_0 and radius R . Performing inverse transformations we obtain an ellipsoid in the input space.

The Petunin ellipsoid has some wonderful properties. First, it uniquely arranges the points according to their statistical depth because by construction at the surface of Petunin ellipsoid only one point lays. Second, the probability that random points from the same distribution lays within Petunin ellipse is equal $\frac{n-1}{n+1}$. Thus, we have not only ellipse containing a given set of n random points with the exact probability, but also may find outliers. Peeling these ellipsoids we may construct an influential set of points in design space.

The main complexity problem of the construction of the Petunin ellipsoid in a space of high dimension is the procedure of finding of two most distant points (ends of the set diameter). As it was noted in [4], the complexity of the construction of a convex hull in \mathbb{R}^d is $O(n^{\lfloor \frac{d}{2} \rfloor})$. But there are some effective and fast procedures [2], [5] that are not optimal but extremely smooth the problem.

PETUNIN ELLIPSES AND LINEAR CONSTRAINS

The above algorithm does not take into account the linear constrains. It constructs an ellipse that just contains all random points with given probability. But the design of this algorithm allow easy modification so that we can select some Petunin ellipse with the given confidence level that lays within the polygon formed by the constrains $Ax \leq b$.

To do this we use a procedure proposed in [1]. Let us return to the step in case $d = 2$ when we construct the concentric circles. These circles may not lie within the polygon. Now, finding the side that intersects with the outer circles (with radius R) we may construct the normal vector to this side and shrink the circle such that it become tangent to this side. Repeating this procedure as many times as the linear constrains are violated we obtain a desired ellipse satisfying the linear constrains. Only drawback is that this ellipse may have a lower confidence level since this level depends on the points covered by the ellipse. This procedure may be extended to the multidimensional case easily.

REFERENCES

- [1] A. Bedrintsev and V. Chepyzhov, Description of the design space by extremal ellipsoids in data representation problems, *J. Commun. Technol. El.* vol. 61, pp. 688-694, 2016.
- [2] S. Har-Peled, A practical approach for computing the diameter of a point-set, In: *Symposium on Computational Geometry (SOCG'2001)*, pages 177-186, 2001.
- [3] S. Lyashko, D. Klyushin, and V. Alexeenko. Multivariate ranking using elliptical peeling. *Cybern. Syst. Anal.* 49, pp. 511-516, 2013.
- [4] S. Lyashko, B. Rublev. Minimal ellipsoids and maximal simplexes in three-dimensional Euclidean space. *Cybern. Syst. Anal.* 39, pp. 831-834, 2003.
- [5] G. Malandain and J.-D. Boissonnat, Computing the Diameter of a Point Set. *International Journal of Computational Geometry and Applications, IJCGA*, vol. 12, pp. 197-208, 2002
- [6] Yu. Petunin and B. Rublev, Pattern recognition with the help of quadratic discriminant function, *J. Math. Sci.*, vol. 97, pp. 3959-3967, 1999.