Piecewise Linear Approximation of Plane Curves with Restrictions in Computer-Aided Design of Railway Routes

Valery I. Struchenkov

Moscow State University of Radio Engineering, Electronics and Automation, Moscow, Russia
*Corresponding Author: str1942@mail.ru

Copyright © 2014 Horizon Research Publishing All rights reserved.

Abstract Under study is approximation problems of plane curves defined by a sequence of points. These problems arise in computer-aided design planning of linear structures. The approximating curve consists of elements of a certain type (straight-line segments, arcs or 2nd degree parabolas, and also clothoids). The parameters of these elements are constrained. Moreover, the number of elements is unknown. The article deals with the simplest of the tasks of the per-element approximation, in which the elements are line segments with presence of a number of restrictions. The problem is solved by dynamic programming. An algorithm for solving the problem under the restrictions of special kind is considered. This algorithm for new CAD of railways gives decrease of construction cost when designing the optimal longitudinal profile.

Keywords Approximation, Restrictions, Dynamic Programming, Longitudinal Profile

1. Introduction

A railway line is a three-dimensional curve satisfying a number of restrictions.

The traditionally required three-dimensional curve is represented by two plane curves: horizontal and vertical alignment (later : plan and profile) (Figure 1). The plan curve is a projection of the 3D curve to the coordinate plane XOY, and the longitudinal profile is the coordinate z as a function of the length s of the curve in the plan.

Project line is broken, which must meet a number of technical restrictions. In the current computer aided design software (CAD) the planning is carried out manually [1,2].

In the new CAD software the optimal longitudinal profile is designed by computer for each alternative in the plan [3]. The project line is a broken line (polyline), which must meet a number of technical restrictions. The number of elements of the project polyline is unknown. Therefore, we believe that its nodes and the nodes of the longitudinal ground profile have the same abscissa. The longitudinal ground profile is always represented by a polyline with an uneven step between vertices, and this assumption makes it possible to fix the number of elements: n (the dimension of the problem) and s_i - the length of the elements (Figure 1). Using nonlinear programming we obtain a broken line with more elements than we need, but its deviation from the required $Z(s)$ is small because the project line is broken and must meet a number of technical restrictions [4]. The idea is to find the broken line by solving the optimization problem, and then transform it into a line with elements whose length is no shorter than the predefined value, thereby defining the real dimension of the problem and the initial approach, and in the final step, perform an optimization taking into account all the restrictions and necessary revisions to the objective function. Therefore the design of the vertical alignment should be carried out in three stages. On the first and the third stages the problem is solved using a new non-linear programming algorithm [4]. The objective function is the minimum construction cost. The mathematical model takes into account the structure of the cross-section profiles of the roadbed, the presence of culverts, bridges, earth works (cuts and fills), methods of excavation, etc. In the past, a heuristic algorithm of sequential local approximation was used on the second stage. The new algorithm solves this approximation problem using dynamic programming [5, 6]. The presentation of this algorithm is the purpose of this article.

Figure 1. Railway line representation: a) plan, b) longitudinal profile

2. Statement of the Problem

Approximated curve is given by a number of points in the Cartesian coordinates $L_1 = \{x_i, y_i\}$ ($i = 1, 2, \dots, n$), i.e. represented as a broken line. It's required to find another broken line with the nodes $L_2 = (X_j, Y_j)$ ($j = 1, 2, \dots, k$), with the following conditions:

$$\min_{L_2} K(L_1, L_2) \text{ and } f_r(X, Y) \leq 0, r = 1, 2, \dots$$

where $K(L_1, L_2)$ predefined quantitative criterion of deviation L_2 from L_1 , and the system $f_r(X, Y) \leq 0$ defines the restrictions on the parameters of the elements of L_2 . The number of restrictions is defined by the number of elements $(k-1)$, which is unknown. The first point (X_1, Y_1) and the last point (X_k, Y_k) are given. Additionally, the direction at start and end point may be specified (elements of contiguity). If L_1 and L_2 correspond to single-valued functions $L_1(x)$ and $L_2(x)$, then

$$K(L_1, L_2) = \int_{x_1}^{x_n} (L_1(x) - L_2(x))^2 dx \text{ where } X_1 = x_1 \text{ и } X_k = x_n \text{ and under restrictions:}$$

1. $c_j(X_j, X_{j+1}) \leq (Y_{j+1} - Y_j) / (X_{j+1} - X_j) \leq d_j(X_j, X_{j+1})$
2. $p_j(X_j) \leq (Y_{j+1} - Y_j) / (X_{j+1} - X_j) - (Y_j - Y_{j-1}) / (X_j - X_{j-1}) \leq q_j(X_j)$
3. $l_{\min} \leq X_{j+1} - X_j \leq l_{\max}$

Here j is number of the nodes of the designed broken line. Additionally we can set limits at certain points in which the abscissa x_r may not coincide with the abscissae of the given points.

4. $L_2(x_r) - L_1(x_r) \leq v_r$ and / or $w_r \leq L_2(x_r) - L_1(x_r)$. Here $x_1 < x_r < x_n$.

In the design of the longitudinal profile the restriction of type 1 is a discrete analogue of the restriction on the first derivative of the approximating function.

The values $c_j(X_j, X_{j+1})$ and $d_j(X_j, X_{j+1})$ depend on the j -th element position relatively to the ends of the curves in the route plan. They define the limits of the slopes.

The values $p_j(X_j)$ and $q_j(X_j)$ depend on the position of X_j relatively to the transition curves (clothoids) in the route plan. They define the limits of the differences of the slopes adjacent elements of the longitudinal profile respectively over the "hump" and in the "pits". The inequality 3 together with 1 defines the minimum and maximum length of the elements. The inequality 4 limits the required extreme deviations of the sought broken line at some points. In particular, the fixed point can be given. Values $l_{\min}$, $l_{\max}$, v_r and w_r are given.

$X_{j+1} - X_j$ changed discretely, for example in steps of 10 m.

Restrictions do not allow to take the given broken line $L_1(x)$ as a solution.

Other criteria, in particular $\min \max |L_2(x) - L_1(x)|$, may be

$$L_2 \quad x$$

used instead of a quadratic criterion.

In the design of the longitudinal profile of railway the criterion may be the total amount of earthwork. If the soils of cuts are not used for the construction of fills, the criterion could be the cost of construction.

3. The Method of Solution

The unknown number of elements k does not allow to use the methods of nonlinear programming. Additionally specifying a deviation band (search area) the problem can be solved using dynamic programming.

Regarding the original curve within the search grid is broken variation (Figure 2). The transition from a given initial point A to an end point B in the final set is considered as a multi-step process. In the first step it considers all the restrictions on the allowable connections between the starting point A and the grid nodes in the range of the first element ($l_{\min} \leq X_2 - X_1 \leq l_{\max}$), the second step - all valid connections between the end of the first element and the following nodes of the grid, so that

$$l_{\min} \leq X_3 - X_2 \leq l_{\max}, \text{ etc.}$$

If for the next j -th element $x_n - X_{j+1} \leq l_{\min}$, then it takes $X_{j+1} = x_n$ and $Y_{j+1} = Y_k$.

Figure 2. First step of piecewise linear approximation

To avoid the exhaustive search, it is necessary to introduce the concept of "state of the system." In accordance with Bellman's principle [2, 3] of all the acceptable paths that lead to a common state of the system, we may keep only one, for which the value of the criterion is minimal.

Because of the restrictions of type 2 we can't compare paths that converge at a single node of the grid. Therefore, as the "state of the system" for all paths we consider the last element (but not a last node) and then we may compare the possible paths of achievement the current node for which the last element is the same. The number of elements of the paths (trajectories) which are compared can be different.

4. Algorithm

1. The range of possible values of the abscissas of the ends of all elements $[x_1 + l_{\min}, x_n - l_{\min}]$ is divided into equal parts

with specified increments dx . For example $dx = 10$ m. As a result we get an array of integers $\{X_d\}$ and appropriate verticals.

2. Ordinates $\{Y_d\}$ of the original broken line L_1 are calculated using linear interpolation on each vertical.

3. We set the increment (i. e. discrete) dy on the verticals. Then up and down from each point of the original broken line we distribute an equal number of discretely, $kd/2$. The restrictions of type 3 are not taken into account when constructing the grid.

4. The resulting grid points are numbered sequentially, beginning with the bottom of the first vertical and ending with the highest point of the last vertical. Only the coordinates of the points of the original broken line are stored. It is easy to calculate the Cartesian coordinates of a point with number N , using the number N and the specified dx , dy , and kd .

$$x_N = l_{\min} + [(N-1)/(k_d+1)]dx; \quad y_N = y_q + (N-1 - [(N-1)/(k_d+1)])(k_d+1) \cdot (k_d/2)dy.$$

Here $[]$ is the integer part, q - point number in the array $\{Y_d\}$. $q = [(N-1)/k_d] + 1$.

This method of numbering states require two bytes for each of them, if the maximum number of points less than 65535. Otherwise it is convenient to number the points on each vertical separately.

All points of the grid are considered as the possible boundaries of the elements.

5. All possible variants of the first element are built and their estimations (the value of objective function) are stored. If the neighboring element on the left is given, then we check the restriction of type 2 for all variants of the first element and eliminate the unsuitable variants.

6. Instead of element wise construction of trajectories each point inside $x_1 + l_{\min} \leq x \leq x_n - l_{\min}$ is considered as the beginning of the next element, no matter how many element the incoming path contains. Iteratively we construct for each point a set of admissible variants of next element (Figure 3. "Outgoing fan").

At $x > x_n - 2l_{\min}$ this "fan" contains only one element and its end is point B. "Outgoing fan" is built only from the points with a nonzero number of incoming elements. For each element of "outgoing fan" only one of the incoming elements is remembered. For this pair (incoming and outgoing elements) the restriction of type 2 is executed and the value of the objective function is minimal. This value is added to the estimate of the element under consideration.

Thus at each point in the $x \geq x_1 + 2l_{\min}$ we generate a lot of incoming elements (Figure 3. "Incoming fan") and for each of them we calculate the estimate of the path which begins at point A and ends at the current point. The number of the input elements is remembered. Any "system state" is given by a pair of integers (begin and end of the element), and for each state one integer (link), and one real number (path estimate) are stored.

7. If the neighboring element on the right is given then at the last step when connecting to the point B we take it into

account. In this case we check the restriction of type 2 in the point B and we eliminate the unsuitable elements.

8. The optimal trajectory is restored by moving in reverse.

Figure 3. Common step of piecewise linear approximation

5. Approximation of Multi-valued Functions

The described algorithm can be used for other purposes, including approximation of multi-valued functions.

For multi-valued functions instead of deviations on the graph ordinates we can use the deviations of the normals to the original curve. (Figure 4).

Figure 4. Approximation with normals

At each node of the original broken line we calculate direction to the center of the circle connecting the three adjacent nodes. We identify this direction with the direction of the normal to the original line in the given node.

Calculation of the deviations of the next element of the approximating line from the original line is reduced to alternately calculating the coordinates of the intersection of this element and the normals to the original line at the nodes which are captured by this element. In Figure 4 there are the normals n_1 , n_2 , n_3 .

If the criterion of the quality of the approximation is the maximum deviation, then the calculation of the deviation in other points of the constructed element is not needed. Indeed, because of the linearity of elements the deviation in the

intermediate points of the element from the original broken line cannot exceed the maximum deviation at the ends. Therefore, we calculate the deviations only on the normals and at the ends of the element and we take the maximum of them. (Figure 4).

Approximation of curves that are not graphs of a single-valued functions, can be made with restrictions. Thus, the restrictions of type 4 go to the limit for deviations on the normals at some points (in particular, they can be the fixed points). Restriction of type 1 does not have much sense, and the restrictions of type 2 go to the limit on the angle at the node of the designed line. Except this we do not need to change anything in the above approximation algorithm.

In theory, as well as the approximation of single-valued functions with a slight complication of the algorithm we can solve the problem under an additional condition of adding circles with a given radius in the breaks of the designed line. Moreover, we can use several values of radius and elect the better of them. Such problems arise by projecting roads, pipelines, etc. However, in this case, we have to consider significantly larger deviations of designed line from the given line. The computational complexity is high if the number of discretized on the normals is large.

6. Results

It is known that the practical application of dynamic programming for a large number of states at each step involves dealing with significant computational complexity which dramatically increases with the number of states [5,6]. This phenomenon is called "curse of dimensionality."

The purpose of experimental calculations was to establish the possibility of practical application of the algorithm for designing the longitudinal profile of new railways within haul on public computers with the search area width 0.5-0.6 m

Due to the restrictions of type 1 and 2 we cannot perform calculations with a grid with large cells, followed by splitting them into a finer grid depending on the solution obtained. Therefore, when the width of the search area of 0.5-0.6 m we have 60 points or more at each vertical.

It was established experimentally that when the number of discretized on each vertical is from 40 to 100 and the length is from 2,500 to 3,000 there aren't any significant computational difficulties.

This enables to use the proposed algorithm not only in the second stage of the above described scheme of longitudinal profile design in conditions of a rugged relief, but also in the single stage of design under flat relief conditions. In the latter case, we design the fill whose height is not less than the given value. Therefore the grid is not divided according to the ground profile but to the appropriate line.

If we want to use dynamic programming to design of the longitudinal profile on a rugged relief without the use of nonlinear programming, then the width of the search area is equal to 10 m and more and computational difficulties are

insuperable.

7. Discussion

The method of solving the problem is based on the pioneering work of R. Bellman, who 50 years ago solved the problem of piecewise linear approximation of planar curves[7]. We used the key concepts of "step process", the "system state", etc. as is common in his works [5-7]. The difference is in presence of restrictions (1-4 types). These differences demanded to change the formalization of these concepts ("step process" and "state of the system") and led to a significant complication of the problem.

The proposed algorithm is one element of a multi-stage technology computer-aided design of the longitudinal profile of the railways. This technology is based on complex optimization methods and has clear advantages over the interactive elaboration of design solutions.

The use of complex mathematical models can solve design problems with their relationship and create intelligent CAD. Designs of longitudinal profile using optimization algorithms give objectively comparison the various route options and choose the most appropriate one. This approach is more promising than the various heuristic algorithms for railways routes design.

Practical tasks were solved on old computer with RAM =512 Mb, CPU 2.0 GHz. For length of the longitudinal profile 25 km the process time was a few minutes. Therefore practical tasks may be solved for a suitable time on modern computer.

The algorithm approximation of discretely defined plane curves with restrictions may find other applications that have nothing to do with designing of railways routes.

8. Conclusion

To apply dynamic programming it is necessary to identify the key concepts: the "step process", the "system state" and give the rule how to construct a set of states at each step, how to reject unpromising paths leading to the current state, and what must be stored to recover the optimal trajectory.

All this is done in this article with respect to the piecewise linear approximation of the plane curve given sequence of points.

Due to restrictions we can not take a single point on the vertical as the "system state". It is necessary to add the value of slope (derivative) in it. Search of the optimal number of elements and their lengths with all the restrictions led to a significant complication of the problem. But such search provides improvement of the quality of design solutions, because any small change in the alignment, will result in a significant change in the total cost [8, 9]. This is the main difference a new CAD software from previously developed [10-12].

The described algorithm is used in two CAD systems. The

first is for the design of new railway lines, and the second - for the design of the reconstruction of railways. A new algorithm assumes the use of a piecewise parabolic approximation in CAD of roads. This algorithm will be published in a future article.

Acknowledgement

The author would like to thank Žiga Ramšak for his help in correcting the English version of this article.

REFERENCES

-
- [1] CARD/1. URL: <http://www.card-1.com/en/home/>
 - [2] Topomatic Robur. URL: <http://www.topomatic.ru>
 - [3] V. I. Struchanov. "Optimization Methods in Applied Problems", Solon-Press, Moscow, 2009.
 - [4] V.I. Struchanov "Mathematical Models and Optimization in Line Structure Routing: Survey and Advanced Results", International Journal Communication, Network and System Sciences. Special Issue : Models and Algorithms for Application, 2012, 5
 - [5] R. Bellman and S. Dreyfus, "Applied Dynamic Programming", Princeton University Press, Princeton, 1962
 - [6] E.S. Wentzel "Operations Research: Challenges, principles, metodologiya." KnoRus, Moscow, 2010
 - [7] R. Bellman "On the approximation of curves by line segments using dynamic programming". Communications of the ACM vol.4, №6, 1961, 284-286
 - [8] Y. Shafahi and M. J. Shahbazi, "Optimum Railway Alignment," 2012.
http://www.uic.org/cdrom/2001/wcrr2001/pdf/sp/2_1_1/210.pdf
 - [9] M. K. Jha, P. M. Schonfeld, J.-C. Yong and E. Kim, "In-telligent Road Design," WIT Press, Southampton, 2006.
 - [10] J. Kurilko and V. Chesheva, "Geonics ZHELDOR- CAD," CADmaster, Moscow, Vol. 1, No. 36, Moscow, 2007.
 - [11] Bentley Rail Track. <http://www.bentley.com/>
 - [12] C. Parkholup "The system for design of transport highways "Trimble Quantm", ComputerPress, CAD and Grafics, No 3, Moscow, 20